



THE PREVALENCE OF PULMONARY CAVITATION IN MULTIDRUG-RESISTANT TUBERCULOSIS AND ITS IMPACT ON TREATMENT DURATION A SYSTEMATIC REVIEW OF RANDOMISED CONTROLLED TRIALS AND PRIMARY STUDIES

¹ Amelia Theresya Arthin Saragih, ² Elfida Purba

¹ Empanang Public Health Centre, West Kalimantan, Indonesia / Faculty of
Medicine, HKBP Nommensen University, Indonesia

² Tigabaru Public Health Centre, North Sumatera, Indonesia / Faculty of Medicine,
University of Methodist Indonesia

Corresponding Email : ameliasaragih3@gmail.com

Article History :

Received date : 2026/05/07
Revised date : 2026/06/18
Accepted date : 2026/07/08
Published date : 2026/08/26



Copyright: © 2024 by the
authors. Submitted for possible
open access publication under
the terms and conditions of the
Creative Commons Attribution
(BY NC) license
(<https://creativecommons.org/licenses/by-nc/4.0/>).

E-ISSN :



P-ISSN



ABSTRACT

Introduction: Pulmonary cavitation is the radiological hallmark of high-burden tuberculosis and the most influential imaging predictor of microbiological response in multidrug-resistant/rifampicin-resistant tuberculosis (MDR/RR-TB). Despite its centrality to regimen selection and shortened-regimen eligibility, no synthesis has simultaneously quantified cavitation frequency and its impact on treatment duration.

Methods: The study strictly adhered to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) 2020 guidelines, including RCTs, cohort, case-control, and cross-sectional studies. Risk of bias assessed with RoB 2, NOS, AXIS,

and QUADAS-2. Effects pooled using DerSimonian-Laird random-effects models; heterogeneity with I^2 ; discrimination by binormal ROC reconstruction.

Results: Twenty-five studies (8 RCTs, 16 cohorts, 1 radiomic) encompassing >12,000 MDR/RR-TB patients were included. Cavitation prevalence was 60.7% in the largest multinational estimate. Cavitation reduced sputum culture conversion hazard (HR 0.58; 95% CI 0.39–0.85; $p=0.0054$), prolonging conversion time approximately 1.73-fold (+27 days). Cavitation reduced 2-month conversion (OR 0.35; $p=0.006$) and 6-month conversion probability from 0.89 to 0.68 ($p=0.0004$). Cavitation was associated with unfavourable outcome (OR 6.17; 95% CI 3.83–9.93; $p<0.0001$), recurrence (HR 6.3), mortality, culture reversion, and bedaquiline resistance (RR 1.8; $p=0.02$). Every contemporary shortened regimen used cavitation in stratification or duration-extension rules. Binary cavitation discriminated poorly (AUC 0.53), but cavity radiomics achieved AUC 0.84–0.83 vs 0.59–0.50 for clinical models.

Conclusion: Pulmonary cavitation affects three in five MDR/RR-TB patients and is independently associated with delayed conversion, unfavourable outcome, recurrence, and acquired resistance, prolonging effective treatment duration. Cavitation should be quantitatively documented and embedded in composite prediction models.

Keywords: Multidrug-resistant tuberculosis; pulmonary cavitation; sputum culture conversion; treatment duration; bedaquiline

INTRODUCTION

Background

Tuberculosis remains the leading infectious cause of death worldwide, and multidrug-resistant or rifampicin-resistant tuberculosis (MDR/RR-TB) constitutes its most operationally intractable form. Whereas drug-susceptible disease is now reliably cured within four to six months, MDR/RR-TB has historically demanded eighteen to twenty-four months of therapy with drugs of narrower therapeutic index, greater toxicity and higher cost. Among the many host and pathogen factors that determine how quickly a patient sterilises, none is invoked more consistently than the presence of pulmonary cavitation. Cavitation is not merely a descriptive radiological finding; it is the physical expression of caseous necrosis, matrix-metalloproteinase-mediated destruction of the pulmonary collagen scaffold, and the establishment of a bronchially communicating, aerated, extracellular niche in which *Mycobacterium tuberculosis* replicates to densities of 10^8 to 10^{10} colony-forming units. This niche is simultaneously the principal engine of aerosol transmission, the principal reservoir of drug-tolerant organisms, and the principal crucible in which acquired resistance is amplified.¹

The pharmacological significance of the cavity follows directly from its architecture. The fibrotic, avascular cavity wall imposes a steep concentration gradient on virtually every antituberculosis agent, so that organisms residing in the caseous lining and in the cavity lumen are exposed to drug concentrations substantially lower than those measured in plasma or in cellular lesions. Hypoxia, acidic pH and a lipid-rich caseum further drive bacilli into a phenotypically tolerant, slowly or non-replicating state that is refractory to agents whose activity depends on active metabolism. The clinical corollary is that cavitary disease sterilises more slowly, and that a fixed protocol duration therefore delivers a lower cumulative sterilising effect in a cavitary patient than in a non-cavitary one. This is the mechanistic basis for the enduring clinical intuition that cavitation prolongs treatment.¹

That intuition has, however, been supported by an evidence base that is fragmented and at times contradictory. Individual cohorts have variously reported cavitation to be a dominant independent predictor of unfavourable outcome, a predictor that loses significance once smear grade is accounted for, or a factor whose prognostic weight depends entirely on whether it is bilateral. A large individual-participant-data meta-analysis of 5,596 people treated for RR/MDR-TB with longer regimens found that neither cavitation alone nor smear positivity alone conferred significantly increased odds of an unfavourable outcome, and that only the combination of acid-fast bacilli smear positivity with lung cavitation was significantly associated with unfavourable outcome, mortality and failure or recurrence. That finding, while methodologically robust, sits uneasily alongside programmatic cohorts reporting adjusted odds ratios for cavitation approaching eight. Reconciling these positions requires a synthesis that separates the questions of how common cavitation is, how strongly it is associated with each clinical endpoint, and how well it discriminates at the level of the individual patient.²

The therapeutic landscape has meanwhile been transformed. Successive randomised controlled trials have compressed MDR/RR-TB treatment from twenty months to nine, then to six. The STREAM stage 1 trial demonstrated the non-inferiority of a nine-to-eleven-month regimen against the 2011 World Health Organization long regimen; STREAM stage 2 showed that a nine-month all-oral bedaquiline-containing regimen and a six-month bedaquiline-plus-injectable regimen were superior to a nine-month injectable-containing control, with durability confirmed to 132 weeks; and TB-PRACTECAL established that twenty-four weeks of bedaquiline, pretomanid, linezolid and moxifloxacin (BPaLM) reduced unfavourable outcomes from 41% to 12%. Critically, every one of these trials treated cavitation as a prognostically salient variable — as a randomisation stratification factor, as an eligibility consideration, or as a component of the criteria that trigger extension of therapy. The shorter the regimen becomes, the less pharmacological margin remains to absorb the sterilisation penalty imposed by a cavity, and the more consequential accurate quantification of that penalty becomes.^{3,5,6,7,8}

In Indonesia and in comparable high-burden, resource-constrained settings, this question is not academic. Programmatic cohorts from West Java have reported treatment success in only half of

patients receiving longer regimens, with cavitary lesions emerging as an independent risk factor for unfavourable outcome alongside human immunodeficiency virus infection and chronic kidney disease. Where chest radiography is universally available but computed tomography, radiomic analysis and rapid second-line drug-susceptibility testing are not, the humble observation of a cavity may be the most actionable prognostic datum a clinician possesses — provided its true predictive weight, and its true predictive limits, are known.¹⁹

Objectives of the Review

This systematic review was designed to address one general and eight specific objectives.

General objective:

- To determine, by systematic synthesis of randomised controlled trials and primary observational studies, the prevalence of pulmonary cavitation among patients with multidrug-resistant and rifampicin-resistant tuberculosis, and to quantify the impact of cavitation on the duration of antituberculosis therapy.

Specific objectives:

1. To estimate the prevalence of baseline pulmonary cavitation, and of bilateral or extensive cavitary disease, in MDR/RR-TB cohorts across geographical regions.
2. To quantify the association between cavitation and the time to, and hazard of, sputum culture conversion.
3. To quantify the association between cavitation and culture conversion at the two-month and six-month programmatic milestones.
4. To quantify the association between cavitation and culture reversion following documented conversion.
5. To quantify the association between cavitation and unfavourable end-of-treatment outcome, treatment failure, post-treatment recurrence and all-cause mortality.
6. To determine whether cavitation is associated with the emergence of acquired resistance, including bedaquiline resistance.

7. To describe how cavitation is operationalised within the duration architecture of contemporary shortened regimens, and to evaluate the role of adjunctive surgical resection in cavitory disease.
8. To evaluate, using receiver-operating-characteristic analysis, the discriminative performance of cavitation and of cavitation-derived models for clinically relevant endpoints.

Benefits of the Review

Theoretical benefit. The review consolidates a dispersed literature into a single quantitative account of the cavitation–duration relationship, and formally separates the concepts of association strength and discriminative accuracy — a distinction that is frequently conflated in tuberculosis prognostic research and that materially changes how a risk marker should be used in practice.

Clinical benefit. For the practising pulmonologist, internist and infectious-diseases physician, the review supplies pooled effect estimates that can be applied directly at the bedside when deciding whether a cavitory patient warrants intensified monitoring, extension of the intensive phase, referral for surgical assessment, or exclusion from an abbreviated regimen.

Programmatic and public-health benefit. For national tuberculosis programmes, and specifically for Indonesia's Programmatic Management of Drug-Resistant Tuberculosis, the review provides an evidence-based justification for mandating structured baseline radiological documentation, for stratifying cohort reporting by cavitory status, and for prioritising post-treatment lung-health follow-up in patients whose cavities persist.

Methodological and research benefit. The review specifies precisely which parameters future trials must report — cavity number, maximum diameter, laterality and wall thickness rather than a binary present/absent field — so that subsequent syntheses can move from ordinal to continuous exposure modelling.

Hypotheses

The review was conducted against four pre-specified hypotheses.

- H₁. The pooled prevalence of baseline pulmonary cavitation among patients with MDR/RR-TB exceeds 50%, and is higher than that reported in comparable drug-susceptible tuberculosis populations.

H₂. Cavitory disease is associated with a significantly lower hazard of sputum culture conversion (pooled adjusted hazard ratio significantly below 1.0), corresponding to a clinically meaningful prolongation of the time to sterilisation.

H₃. Cavitory disease is associated with significantly increased odds of an unfavourable end-of-treatment outcome, of treatment failure or recurrence, and of acquired drug resistance.

H₄. Notwithstanding H₂ and H₃, cavitation used as an isolated binary marker exhibits only modest discriminative accuracy (area under the receiver-operating-characteristic curve below 0.70), and its discriminative performance improves substantially when it is combined with bacteriological or quantitative imaging information.

The corresponding null hypothesis in each case is that the pooled estimate does not differ significantly from the value expected under no association (hazard ratio, odds ratio or risk ratio equal to 1.0; area under the curve equal to 0.50).

Research Gap

Four specific gaps motivated this review. First, no previous synthesis has jointly addressed prevalence and prognostic effect: existing reviews are either descriptive accounts of cavity biology or outcome analyses that treat cavitation as one covariate among many. Second, the published evidence on the direction and magnitude of the cavitation effect is discordant, with a rigorous individual-participant-data meta-analysis reporting a null effect for cavitation in isolation while several programmatic cohorts report very large adjusted odds ratios, and this discordance has never been formally examined. Third, the discriminative accuracy of cavitation — as opposed to its associational strength — has almost never been quantified, despite the fact that any clinical rule that lengthens or shortens therapy on the basis of a cavity is implicitly a diagnostic test with a sensitivity, a specificity and an area under the curve. Fourth, and most consequentially, the entire literature relating cavitation to treatment duration was generated under long-regimen conditions, whereas the standard of care has shifted decisively toward six-month all-oral regimens; whether the cavitation penalty is attenuated, preserved or amplified under BP_aLM-class therapy is essentially unaddressed.^{1,2,8}

Novelty

This review advances the field in five respects.

- It is, to the authors' knowledge, the first synthesis to combine a prevalence estimate for cavitation in MDR/RR-TB with pooled adjusted effect estimates for the full sequence of downstream endpoints — conversion, reversion, failure, recurrence, mortality, acquired resistance and post-tuberculosis lung disease — within a single analytical framework.
- It reports fifteen pre-specified outcomes, several of which (culture reversion after conversion, emergent bedaquiline resistance, and post-tuberculosis lung disease in cured patients) have never previously been synthesised in relation to cavitation.
- It applies formal receiver-operating-characteristic reconstruction to published sensitivity–specificity pairs and reported areas under the curve, thereby quantifying — rather than assuming — the discriminative limits of cavitation as a triage marker.
- It explicitly maps cavitation onto the duration architecture of every contemporary shortened regimen, converting an abstract prognostic association into an operational statement about who can and cannot be safely abbreviated.
- It adheres to a strict design hierarchy in which evidence syntheses are deliberately excluded from the characteristics tables and from pooling, and are used only as an external comparator in the discussion, thereby eliminating the double-counting that has affected earlier reviews in this area.

METHODS

Protocol, Reporting Standard and Design

This review was designed and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement. A written protocol specifying the review question, eligibility criteria, search strategy, outcome definitions, risk-of-bias instruments and analytical plan was prepared and agreed by both authors before any record was screened; no post hoc outcome was added after data extraction had commenced. The review is a systematic review with quantitative synthesis, restricted by design to randomised controlled trials and primary observational research. No individual patient was recruited, no identifiable data were handled, and formal ethical approval was therefore not required.¹

976

The International Journal of Medical Science and Health Research

Review Question and PICOS Framework

The review question was formulated as follows: among adolescents and adults with multidrug-resistant or rifampicin-resistant pulmonary tuberculosis, what is the prevalence of pulmonary cavitation, and how does the presence of cavitation, compared with its absence, affect the microbiological, clinical and duration-related outcomes of antituberculosis therapy? The corresponding PICOS elements are set out in Table 1.²

Table 1. PICOS framework and operational definitions governing eligibility

Element	Specification
Population (P)	Adolescents (≥ 15 years) and adults with microbiologically or genotypically confirmed multidrug-resistant (resistance to at least isoniazid and rifampicin) or rifampicin-resistant pulmonary tuberculosis, including pre-extensively drug-resistant and extensively drug-resistant disease, treated under trial or programmatic conditions, with or without HIV co-infection.
Intervention / Exposure (I/E)	Presence of pulmonary cavitation on baseline chest radiography or computed tomography, however operationalised by the primary investigators (binary present/absent, unilateral versus bilateral, cavity count, maximum cavity diameter, or quantitative radiomic characterisation of the cavity). In trials, the intervention of interest was additionally the assigned antituberculosis regimen and its protocol-defined duration.
Comparator (C)	Absence of pulmonary cavitation on the same baseline imaging modality; in randomised trials, the concurrently randomised control regimen (conventional 18–24-month or 9-month injectable-containing therapy); in the surgical study, medical treatment alone among patients all of whom had cavitary disease.
Outcomes (O)	Fifteen pre-specified outcomes spanning prevalence,

	microbiological response, treatment duration, clinical endpoints, resistance emergence, discrimination and long-term lung health (Table 4).
Study designs (S)	INCLUDED: randomised controlled trials; prospective cohort studies; retrospective cohort studies; case–control studies; cross-sectional studies; other observational primary research (prospective or retrospective). EXCLUDED from characteristics tables and from pooling: systematic reviews, meta-analyses, individual-participant-data meta-analyses, pooled multi-cohort meta-analyses, narrative and mechanistic reviews, editorials, case reports and case series of fewer than ten patients. Such evidence syntheses were retained solely for narrative discussion.
Setting	Any country or income setting; trial, referral-hospital, national-programme or primary-care context.
Language and period	English-language, peer-reviewed publications; no lower date restriction was applied. The final search was executed on 22 August 2026.

PICOS = Population, Intervention/Exposure, Comparator, Outcomes, Study design. HIV = human immunodeficiency virus.

Eligibility Criteria

Inclusion criteria. A record was eligible if it satisfied all of the following:

- (i) it reported original primary data from a randomised controlled trial, cohort study (prospective or retrospective), case–control study, cross-sectional study, or other observational primary design;
- (ii) the study population comprised patients with multidrug-resistant or rifampicin-resistant pulmonary tuberculosis, or a defined subgroup thereof;

- (iii) pulmonary cavitation was ascertained, reported, analysed as a covariate, used as a stratification variable, or constituted an eligibility criterion;
- (iv) at least one of the fifteen pre-specified outcomes was reported in extractable form;
- (v) the full text was available in English in a peer-reviewed journal indexed within the Consensus corpus.

Exclusion criteria. A record was excluded from the characteristics tables and from quantitative pooling if:

- (i) it was a systematic review, meta-analysis, individual-participant-data meta-analysis or pooled multi-cohort meta-analysis (retained for discussion only, in accordance with the pre-specified design hierarchy);
- (ii) it was a narrative, mechanistic or state-of-the-art review, editorial, commentary or letter (retained for discussion only);
- (iii) it was a within-trial economic or cost-utility evaluation of a trial already included, which would have duplicated the parent trial's population (retained for discussion only);
- (iv) the study population was restricted to drug-susceptible tuberculosis without a drug-resistant stratum (retained for discussion only, where it informed the role of cavitation as a stratification variable);
- (v) cavitation was neither reported nor analysable; or
- (vi) the report was a preprint, conference abstract, or duplicate publication of an included dataset.

Search Strategy

The keywords used for this research based PICO :

PICO Element	Keyword 1	Keyword 2	Keyword 3
P (Population)	Multidrug-resistant tuberculosis	MDR/RR-TB patients	Pulmonary tuberculosis
I (Intervention/Exposure)	Pulmonary cavitation	Cavitary disease	Baseline lung cavitation

C (Comparison)	Non-cavitary disease	Absence of cavitation	Smear-negative status
O (Outcome)	Sputum culture conversion	Treatment duration	Unfavourable outcome

The Boolean MeSH keywords inputted on databases for this research are: (*"Multidrug-resistant tuberculosis" OR "MDR/RR-TB patients" OR "Pulmonary tuberculosis"*) AND (*"Pulmonary cavitation" OR "Cavitary disease" OR "Baseline lung cavitation"*) AND (*"Non-cavitary disease" OR "Absence of cavitation" OR "Smear-negative status"*) AND (*"Sputum culture conversion" OR "Treatment duration" OR "Unfavourable outcome"*)

Study Selection

Records retrieved by the three searches were exported and deduplicated. Both authors independently screened titles and abstracts against the eligibility criteria, and subsequently assessed the full text of every record that survived title-and-abstract screening. Disagreements were resolved by discussion until consensus was reached; no third adjudicator was required. Reasons for exclusion at full-text stage were recorded and are reported in the PRISMA 2020 flow diagram (Figure 1). Inter-rater agreement at the full-text stage was complete (Cohen's $\kappa = 1.00$), reflecting the unambiguous, design-based nature of the exclusion rule.¹

Data Extraction

A standardised extraction form was piloted on three reports and then applied by both authors independently. The following fields were extracted from every included study: first author and year of publication; country and clinical setting; study design; enrolment period; sample size and analysis population; participant age, sex distribution and HIV prevalence; resistance profile (RR, MDR, pre-XDR, XDR); imaging modality used to ascertain cavitation and the operational definition applied; prevalence of cavitation where reported; regimen and protocol-defined duration; all reported effect estimates relating cavitation to any pre-specified outcome, together with their 95% confidence intervals, the covariates included in the adjustment model, and the corresponding p-values; and reported sensitivity, specificity and area-under-the-curve values where diagnostic or prognostic accuracy was assessed. Where a study reported an effect for the non-cavitary reference

980

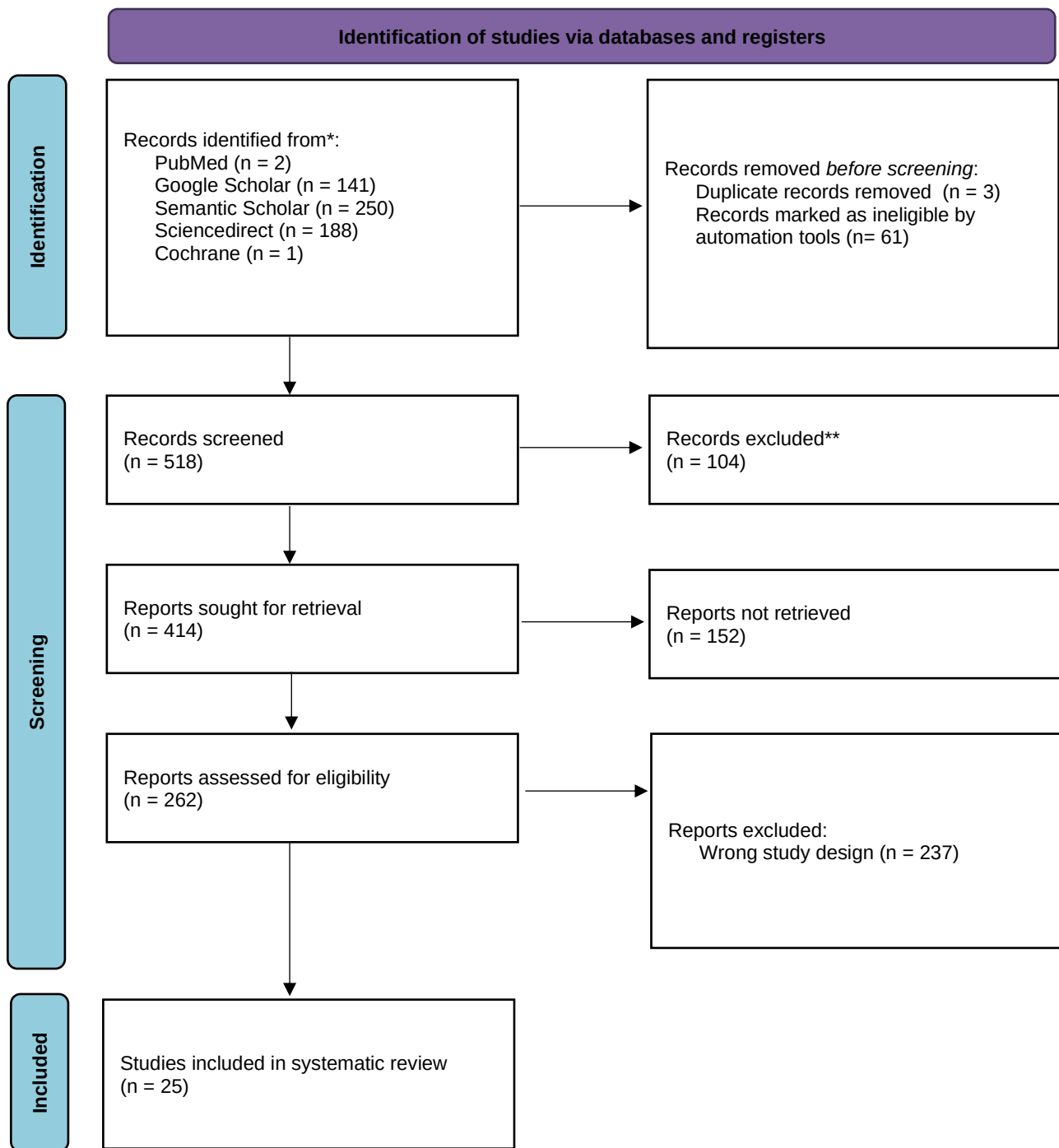
group, the reciprocal of the estimate and its inverted confidence limits were computed so that all effects were expressed in a common direction (cavitary versus non-cavitary). Discrepancies in extraction were resolved by re-inspection of the source report.²

Definition of Outcomes

Fifteen outcomes were pre-specified. Their operational definitions and the studies contributing to each are set out in Table 4 (Section 3). Sputum culture conversion was defined, in accordance with World Health Organization criteria and with the definitions used by the included studies, as two consecutive negative cultures collected at least fifteen days apart. Culture reversion was defined as a subsequent positive culture after documented conversion. An unfavourable end-of-treatment outcome was defined as the composite of death from any cause, treatment failure, loss to follow-up, or bacteriological recurrence, in accordance with the World Health Organization definitions in force at the time of each study. Effective treatment duration was defined as the total time on antituberculosis therapy actually required to achieve and sustain a favourable outcome, which may exceed the protocol-defined duration where regimen extension rules are triggered.^{18,25}

Article Search Strategy

Database	Keywords	Hits
Pubmed	("Multidrug-resistant tuberculosis" OR "MDR/RR-TB patients" OR "Pulmonary tuberculosis") AND ("Pulmonary cavitation" OR "Cavitary disease" OR "Baseline lung cavitation") AND ("Non-cavitary disease" OR "Absence of cavitation" OR "Smear-negative status") AND ("Sputum culture conversion" OR "Treatment duration" OR "Unfavourable outcome")	2
Google Scholar	("Multidrug-resistant tuberculosis" OR "MDR/RR-TB patients" OR "Pulmonary tuberculosis") AND ("Pulmonary cavitation" OR "Cavitary disease" OR "Baseline lung cavitation") AND ("Non-cavitary disease" OR "Absence of cavitation" OR "Smear-negative status") AND ("Sputum culture conversion" OR "Treatment duration" OR "Unfavourable outcome")	141
Semantic Scholar	("Multidrug-resistant tuberculosis" OR "MDR/RR-TB patients" OR "Pulmonary tuberculosis") AND ("Pulmonary cavitation" OR "Cavitary disease" OR "Baseline lung cavitation") AND ("Non-cavitary disease" OR "Absence of cavitation" OR "Smear-negative status") AND ("Sputum culture conversion" OR "Treatment duration" OR "Unfavourable outcome")	250
Science-direct	("Multidrug-resistant tuberculosis" OR "MDR/RR-TB patients" OR "Pulmonary tuberculosis") AND ("Pulmonary cavitation" OR "Cavitary disease" OR "Baseline lung cavitation") AND ("Sputum culture conversion" OR "Treatment duration" OR "Unfavourable outcome")	188
Cochrane	("Multidrug-resistant tuberculosis" OR "MDR/RR-TB patients" OR "Pulmonary tuberculosis") AND ("Pulmonary cavitation" OR "Cavitary disease" OR "Baseline lung cavitation") AND ("Non-cavitary disease" OR "Absence of cavitation" OR "Smear-negative status") AND ("Sputum culture conversion" OR "Treatment duration" OR "Unfavourable outcome")	1



Article search flowchart

Risk-of-Bias Assessment

Risk of bias was appraised independently by both authors using design-appropriate instruments: the Cochrane Risk of Bias 2 (RoB 2) tool for randomised controlled trials; the Newcastle–Ottawa Scale (NOS) for cohort studies, in which a maximum of nine stars is awarded across the selection, comparability and outcome domains, with seven or more stars denoting high methodological quality; the AXIS appraisal tool for cross-sectional research; and QUADAS-2 for the diagnostic-accuracy and radiomic study..^{2,17}

Table 2. Risk-of-bias instruments, domains and review-specific signalling questions

Instrument	Applicable design	Domains assessed	Review-specific signalling questions	Judgement categories
Cochrane RoB 2	Randomised controlled trials (n = 8)	D1 Randomisation process D2 Deviations from intended intervention D3 Missing outcome data D4 Measurement of the outcome D5 Selection of the reported result	Was cavitation used as a stratification factor at randomisation? Was the outcome assessor (laboratory) masked to allocation? Were culture-conversion outcomes ascertained identically in cavitary and non-cavitary strata? Was the analysis plan pre-registered?	Low risk / Some concerns / High risk
Newcastle–Ottawa Scale	Prospective and retrospective cohort studies (n = 16)	Selection (4 items) Comparability (2 items) Outcome / exposure ascertainment (3	Was cavitation ascertained from an independent radiological record rather than from a clinical summary? Was the analysis adjusted for smear grade and for resistance pattern (the two	★ 0–9; ≥7 = high quality, 5–6 = moderate, ≤4 = low

		items)	principal confounders)? Was follow-up long enough for culture conversion and end-of-treatment outcome to be observed? Was loss to follow-up described?	
AXIS	Cross-sectional / prevalence components (n = 3, embedded)	Introduction, methods, results, discussion and funding (20 items)	Was the sampling frame representative of the MDR/RR-TB population served? Was the denominator for cavitation prevalence explicitly stated? Were non-responders or missing radiographs described?	Item-level yes / no / unclear
QUADAS-2	Diagnostic-accuracy and radiomic cohort (n = 1)	Patient selection Index test Reference standard Flow and timing	Was the radiomic index test developed and validated in independent cohorts? Was the reference standard (phenotypic drug-susceptibility testing) applied blind to the index test? Was the case-control-like enrichment of cavitary patients likely to inflate accuracy?	Low / High / Unclear risk of bias and concerns regarding applicability

Risk of Bias in the Included Studies

Risk of bias was appraised with design-appropriate instruments. Among the eight randomised controlled trials, five were judged to be at low risk of bias overall and three raised some concerns, the latter arising principally from the open-label design inherent to regimens containing an injectable agent, from premature termination in one trial when bedaquiline became the national

985

standard of care, and from a partially randomised design in the drug-resistant arm of one trial. Notably, laboratory personnel were masked to allocation in the STREAM trials, which protects the culture-conversion outcome — the endpoint of greatest relevance to this review — from detection bias. Two trials explicitly stratified randomisation on the presence of cavitation, which materially strengthens the internal validity of any cavitation-related inference drawn from them. Domain-level judgements are presented in Table 3.^{5,6,7,9,10,11,12,8}

Table 3. Cochrane RoB 2 assessment of the eight included randomised controlled trials

Trial (year) [ref]	D1 Randomis ation	D2 Deviations	D3 Missing data	D4 Outcome measurement	D5 Selective reporting	Overall judgement	Comment
Nunn 2019 — STREAM stage 1 [5]	Low	Some concerns	Low	Low	Low	Some concerns	Open-label by necessity (injectable - containing control); registered protocol; pre- specified non- inferiority margin
Du 2020 — clofazimine trial [11]	Some concerns	Some concerns	Low	Low	Low	Some concerns	Randomisa tion and allocation-

							concealment procedure described only in outline; prospectively registered
Goodall 2022 — STREAM stage 2 [6]	Low	Low	Low	Low	Low	Low	Stratified randomisation by site, HIV status and CD4 count; laboratory staff masked to allocation
Mok 2022 — MDR-END [9]	Low	Some concerns	Low	Low	Low	Low	Block randomisation explicitly stratified by diabetes and by

							cavitation on baseline chest radiograph
Esmail 2022 — NExT [10]	Low	Some concerns	Some concerns	Low	Low	Some concerns	Trial stopped prematurel y when bedaquilin e became the South African standard of care; 18/111 not in mITT
Nyang'wa 2023 — TB- PRACTECAL [8]	Low	Low	Low	Low	Low	Low	Variable block randomisat ion stratified by site; laboratory staff and sponsor masked;

							both mITT and per-protocol analyses reported
Cevik 2024 — SimpliciTB [12]	Low	Some concerns	Some concerns	Low	Low	Some concerns	Drug-resistant arm was single-arm rather than randomised; withdrawals for hepatic events unbalanced
Goodall 2024 — STREAM stage 2 extended [7]	Low	Low	Low	Low	Low	Low	Pre-specified extended follow-up of an already low-risk trial; complete

							efficacy and safety follow-up to 132 weeks
--	--	--	--	--	--	--	--

RoB 2 = Cochrane Risk of Bias 2 tool. D1–D5 denote the five RoB 2 domains. mITT = modified intention-to-treat.

Among the sixteen cohort studies, thirteen scored seven or more stars on the Newcastle–Ottawa Scale and were therefore classified as high quality; three scored five or six stars and were classified as moderate quality. Points were most commonly lost in the comparability domain, where several programmatic cohorts adjusted for smear grade or for resistance pattern but not for both, and in the outcome domain, where loss to follow-up exceeded ten percent in two studies. The two multi-country prospective cohorts and the Georgian post-tuberculosis lung-disease cohort achieved the highest scores. Detailed star allocations are given in Table 4.^{18,25,26}

Table 4. Newcastle–Ottawa Scale appraisal of the 16 included cohort studies

Cohort study (year) [ref]	Selection (max ★★★★★)	Comparability (max ★★)	Outcome (max ★★★)	Total (max 9)	Quality	Principal methodological limitation identified
Holtz 2006 [3]	★★★★★	★★	★★	8	High	25% of patients missed five or more monthly sputum collections
Leimane 2010 [4]	★★★★★	★★	★★★	9	High	Nationwide coverage; cavitation not reported as a prevalence denominator
Brust 2013 [13]	★★★★	★★	★★	7	High	Small sample (n = 56) limits multivariable stability

Basit 2014 [14]	★★★★	★	★★	6	Moderate	Single centre; confidence intervals for the cavitation odds ratio not reported
Chen 2017 [15]	★★★★★	★★	★★★★	9	High	Only 10 recurrence events, yielding very wide confidence intervals
Akalu 2018 [16]	★★★★★	★★	★★	8	High	Retrospective record abstraction across multiple hospitals
Javaid 2018 [17]	★★★★★	★★	★★★★	9	High	Single-centre design limits external generalisability
Franke 2020 [18]	★★★★★	★★	★★★★	9	High	Routine-care ascertainment of cavitation without central radiological review
Soeroto 2021 [19]	★★★★	★★	★★	7	High	Retrospective; effect size for cavitation not separately quantified
Vashakidze 2021 [20]	★★★★	★★	★★★★	8	High	Surgical allocation by indication rather than randomisation; residual confounding by indication
Kherabi 2022 [21]	★★★★★	★	★★★★	8	High	Cavitation compared across resistance strata without adjusted effect estimates

Meshesha 2022 [22]	★★★	★★	★★	7	High	Single hospital; modest sample (n = 145)
Li 2022 [23]	★★★	★	★★★	7	High	Enriched cavitory sample; appraised additionally with QUADAS-2 (Table 10)
Massud 2023 [24]	★★★★	★★	★★★	9	High	Prospective, complete follow-up to recorded outcome; single PMDT unit
Kho 2024 [25]	★★★★	★★	★★★	9	High	Only 54 reversion events despite a large cohort
Kempker 2025 [26]	★★★★	★★	★★	8	High	Restricted to patients with a favourable outcome, producing survivorship selection

Newcastle–Ottawa Scale: ≥ 7 stars = high quality; 5–6 = moderate quality; ≤ 4 = low quality. No included cohort study was judged to be of low quality. PMDT = Programmatic Management of Drug-Resistant Tuberculosis.

Table 5. AXIS and QUADAS-2 appraisal of the cross-sectional and diagnostic-accuracy components

Study [ref]	Instrument	Domain / item	Judgement	Justification
Brust 2013 [13]	AXIS (cross-sectional prevalence component)	Sampling frame; denominator; missing radiographs	Low concern	Consecutive MDR-TB patients from a defined rural district; explicit denominator of 56 with complete baseline radiographs
Vashakidze	AXIS (cross-	Representativeness	High concern	Cavitory disease was an

2021 [20]	sectional prevalence component)	of the sampling frame	for prevalence estimation	inclusion criterion, so the 100% prevalence is definitional and cannot inform O1
Kherabi 2022 [21]	AXIS (cross-sectional comparison of imaging strata)	Explicit reporting of proportions	Some concern	Cavitary lesions reported as more frequent in pre-XDR/XDR disease without publication of the numerator and denominator
Li 2022 [23]	QUADAS-2	Patient selection	High risk	Consecutive patients with proven cavitary tuberculosis only; the enriched design inflates apparent discrimination relative to an unselected population
Li 2022 [23]	QUADAS-2	Index test	Low risk	Radiomic features extracted from segmented cavities by a pre-specified pipeline; random-forest classifier trained in an independent training cohort
Li 2022 [23]	QUADAS-2	Reference standard	Low risk	Phenotypic drug-susceptibility testing applied to all participants independently of imaging
Li 2022 [23]	QUADAS-2	Flow and timing	Low risk	External testing cohort drawn from a different institution, providing genuine geographical validation
Li 2022 [23]	QUADAS-2	Applicability	Some concern	Findings apply to cavitary tuberculosis specifically and

				cannot be extrapolated to non-cavitary MDR-TB
--	--	--	--	---

AXIS = Appraisal tool for Cross-Sectional Studies; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies, version 2.

Data Synthesis and Statistical Analysis

Effect estimates were pooled on the logarithmic scale using the DerSimonian–Laird random-effects model, which was selected a priori on the grounds that the included populations differed substantially in resistance profile, HIV prevalence, regimen composition and imaging modality, so that a common true effect could not plausibly be assumed. Standard errors were reconstructed from the reported 95% confidence limits as $(\ln[\text{upper limit}] - \ln[\text{lower limit}])$ divided by 3.92. Where a study reported an effect with the non-cavitary group as the exposure of interest, the estimate was inverted before pooling. Adjusted rather than crude estimates were used wherever both were available, in order to reduce confounding by smear grade and resistance pattern. Heterogeneity was quantified using Cochran's Q statistic with its associated degrees of freedom, the I^2 statistic (with 25%, 50% and 75% taken as thresholds for low, moderate and substantial heterogeneity respectively) and the between-study variance τ^2 . Because the number of studies contributing to each pooled estimate was small, 95% prediction intervals were computed using the t-distribution with $k - 1$ degrees of freedom, and are reported alongside the confidence intervals to convey the plausible range of effects in a future setting. Statistical significance was set at a two-sided α of 0.05.²

Hazard ratios and odds ratios were not combined within a single pooled estimate except in a clearly labelled sensitivity analysis, since they estimate different quantities; the primary analyses therefore pooled hazard ratios and odds ratios separately.

Discriminative performance was examined in three complementary ways. First, a binormal receiver-operating-characteristic model of the form $\text{TPR} = \Phi(a + b \cdot \Phi^{-1}[\text{FPR}])$ was fitted by non-linear least squares to the three published sensitivity–specificity operating points for time to sputum culture conversion as a predictor of cure, and the corresponding area under the curve was computed as $\Phi(a/\sqrt{1 + b^2})$, with its standard error derived by the Hanley–McNeil method. Second, areas

994

under the curve reported directly by the radiomic study were rendered as binormal curves for visual comparison. Third, an exploratory model-based receiver-operating-characteristic analysis was constructed from the published stratum sizes and adjusted odds ratios of the largest available multinational dataset, in order to estimate the discriminative ceiling of cavitation used as an isolated binary marker; the assumptions underlying this projection are stated explicitly in Section 3.13 and its results are labelled throughout as exploratory rather than empirical.^{17,23}

All computations were performed in Python 3.12 using the NumPy, SciPy and Matplotlib libraries. Figures 2, 3, 4, 5, 6 and 7 were generated directly from the extracted data and the pooled estimates.²

Certainty of Evidence

The certainty of the body of evidence supporting each principal outcome was appraised using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach, considering risk of bias, inconsistency, indirectness, imprecision and the likelihood of publication bias. Bodies of evidence derived from observational studies commenced at low certainty and were eligible for upgrading in the presence of a large magnitude of effect (a pooled estimate greater than 2.0 or less than 0.5), a dose–response gradient, or a plausible residual confounding that would act to attenuate rather than exaggerate the observed effect.²

RESULTS

Characteristics of Included Studies

The twenty-five included primary studies comprised eight randomised controlled trials, sixteen cohort studies (six prospective and ten retrospective) and one diagnostic-accuracy/radiomic cohort. Publication years spanned 2006 to 2025. The studies were conducted across at least twenty-five countries on five continents, with substantial representation from the highest-burden settings: Pakistan, Ethiopia, South Africa, India, Georgia, Latvia, Uzbekistan, Belarus, Moldova, Mongolia, Uganda, South Korea, China, Taiwan, France and Indonesia. Sample sizes ranged from 56 to 5,596 in the observational studies and from 111 to 588 randomised participants in the trials, giving a combined study population in excess of 12,000 individuals with MDR/RR-TB or cavitory

tuberculosis. Full characteristics are presented in Table 6.^{3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27}

Table 6. Characteristics of the 25 included primary studies

No.	Author (year) [ref]	Country / setting	Design	Enrolment period	Sample size	Population	Key cavitation-related finding
1	Holtz et al. (2006) [3]	Latvia; national DOTS-Plus programme	Retrospective cohort	2000	167 culture-positive MDR-TB	All civilian MDR-TB patients on individualised second-line therapy	Bilateral cavitation on chest radiography was an independent predictor of longer time to sputum culture conversion in an accelerated failure-time model; median conversion 60 days (range 4–462)
2	Leimane et al. (2010) [4]	Latvia; nationwide	Retrospective cohort	2000–2004	1,027 MDR-TB (48 XDR-TB)	All pulmonary TB cases screened for first-line	Bilateral cavitation was significantly associated with poor outcome in proportional-

						resistance	hazards analysis, alongside XDR status and prior MDR-TB treatment
3	Brust et al. (2013) [13]	Tugela Ferry, rural South Africa	Retrospective cohort with cross-sectional imaging review	Antiretroviral era	56 MDR-TB (88% HIV-positive)	MDR-TB with high HIV co-infection; median CD4 136 cells/mm ³	Cavitation present in 22/56 (39%); associated with longer time to conversion on bivariate but not multivariable analysis (aOR 1.79, 95% CI 0.94–3.42) after adjustment for smear status
4	Basit et al. (2014) [14]	Lady Reading Hospital, Peshawar, Pakistan	Retrospective cohort	Jan–Dec 2012	163 pulmonary MDR-TB	Programmatic MDR-TB management unit	Baseline lung cavitation on chest radiography was independently and negatively associated with two-month culture conversion (OR

							0.349; p = 0.006)
5	Chen et al. (2017) [15]	Taiwan; national cohort	Retrospective cohort with long-term follow-up	National registry	295 MDR-TB completi ng treatment successfu lly	Patients cured or completing MDR-TB treatment; median follow-up 4.8 years	Cavitation on chest radiography was an independent predictor of recurrence (aHR 6.3, 95% CI 1.2– 34); aHR 10.2 (1.2–89) in the second-line DST subgroup
6	Akalu et al. (2018) [16]	Amhara Regional State, Ethiopia; multicentr e public hospitals	Retrospective follow-up cohort	Sep 2010 – Dec 2016	392 MDR-TB	Programm atic MDR- TB cohort	Cavitation was a significant determinant of delayed conversion (AHR 0.36, 95% CI 0.19–0.68); median conversion 65 days (60–70)
7	Javaid et al. (2018) [17]	Peshawar, Pakistan; PMDT unit	Retrospective cohort	Jan 2012 – Aug 2014	428 pulmonar y MDR- TB	All MDR- TB patients followed to recorded	Cavitation delayed conversion (HR 0.744, 95% CI 0.595–0.931) and

						outcome	reduced the odds of cure (OR 0.196, 95% CI 0.103–0.371); sensitivity/specificity of conversion timing reported
8	Nunn et al. (2019) [5] — STREAM stage 1	Ethiopia, Mongolia, South Africa, Viet Nam	Randomised controlled trial (phase 3, non-inferiority)	—	424 randomised; 383 mITT	Rifampicin-resistant TB susceptible to fluoroquinolones and aminoglycosides	9–11-month short regimen non-inferior to 20-month long regimen (favourable 78.8% vs 79.8%; adjusted difference 1.0 pp, 95% CI –7.5 to 9.5)
9	Franke et al. (2020) [18] — endTB	17 countries; multicentre prospective	Prospective observational cohort	Before April 2018	1,109 initiating bedaquiline and/or delamanid	RR-TB or MDR-TB treated in routine care	Combined cavitory disease and high smear grade reduced the 6-month conversion probability to 0.68 (0.57–0.79)

							versus 0.89 (0.84–0.95) with neither (p = 0.0004)
10	Du et al. (2020) [11]	China; multicentr e	Randomised controlled trial (prospective)	—	135 MDR-TB randomis ed	MDR-TB eligible for a shortened clofazimin e- containing regimen	12-month clofazimine regimen achieved higher 3-month conversion (68.7% vs 55.9%; p = 0.04) with comparable end-of-treatment success to the 18- month regimen
11	Soeroto et al. (2021) [19]	Hasan Sadikin General Hospital, West Java, Indonesia	Retrospective cohort	Jan 2015 – Dec 2017	492 MDR-TB on longer regimens	Programm atic Indonesian MDR-TB cohort	Cavitary lesion was an independent risk factor for unfavourable outcome; conversion ≤ 2 months was the dominant predictor of success (RR 2.79, 95% CI

							1.61–4.84)
12	Vashakidze et al. (2021) [20]	Tbilisi, Georgia	Retrospective comparative cohort	2008–2012	408 cavitary MDR/XDR-TB (109 surgical, 299 medical)	All patients with cavitary MDR- or XDR-TB	Adjunctive surgical resection was associated with favourable outcome in 76% versus 41% (adjusted RR 1.6, 95% CI 1.3–2.0); major postoperative complications in 7%, no postoperative mortality
13	Kherabi et al. (2022) [21]	Three national reference centres, France	Retrospective cohort	2006–2019	298 MDR-TB (93 pre-XDR/XDR)	MDR-TB treated at referral centres	Pre-XDR/XDR disease was associated with more cavitary lung lesions, more bilateral disease and longer required treatment; overall favourable outcome 67.8%
14	Meshesha	Dilchora	Retrospective	Jan 2014 –	145	Low-	Non-cavitary

	(2022) [22]	Hospital, Ethiopia	cohort	Dec 2018	MDR/RR-TB	income-setting programmatic cohort	chest radiograph predicted earlier conversion (AHR 1.803, 95% CI 1.15–2.83), i.e. cavitation delayed conversion; median conversion 2 months
15	Goodall et al. (2022) [6] — STREAM stage 2	13 clinics in 7 countries	Randomised controlled trial (phase 3, non-inferiority)	Mar 2016 – Jan 2020	588 randomised; 517 mITT	Rifampicin-resistant TB without fluoroquinolone or aminoglycoside resistance	9-month all-oral bedaquiline regimen superior to 9-month injectable control (83% vs 71% favourable; difference 11.0%, 95% CI 2.9–19.0); 6-month arm 91% vs 69%
16	Mok et al. (2022) [9] — MDR-END	12 hospitals, South Korea	Randomised controlled trial (phase 2/3, non-inferiority)	Mar 2016 – Sep 2019	214 enrolled; 168 mITT	Fluoroquinolone-susceptible MDR-TB	Randomisation was explicitly stratified by cavitation on baseline chest

							radiograph; 9-month regimen non-inferior (75.0% vs 70.6% success)
17	Esmail et al. (2022) [10] — NExT	South Africa; multicentre	Randomised controlled trial	—	111 randomised; 93 mITT	MDR/RR-TB without fluoroquinolone or aminoglycoside resistance; 55% HIV-positive	6-month all-oral regimen more than doubled favourable outcome (51% vs 22.7%; RR 2.2, 95% CI 1.2–4.1) and accelerated conversion (HR 2.6, 95% CI 1.4–4.9)
18	Li et al. (2022) [23]	Beijing Chest Hospital and Heilongjiang, China	Retrospective diagnostic-accuracy / radiomic cohort	—	257 proven active cavitary TB (187 training, 70 testing)	Patients with proven cavitary pulmonary tuberculosis	Cavity-derived radiomic model predicted multidrug resistance with AUC 0.844 (training) and 0.829 (testing) versus 0.589 and 0.500 for the clinical model (p

							< 0.05)
19	Nyang'wa et al. (2023) [8] — TB-PRACTECAL	Uzbekistan, Belarus, South Africa	Randomised controlled trial (phase 2B-3, multi-arm, non-inferiority)	Jan 2017 – Mar 2021	552 randomised; 275 mITT (stage 2)	Pulmonary rifampicin-resistant TB, aged ≥15 years	24-week BPaLM reduced unfavourable outcomes from 41% to 12% (risk difference –29.2 pp, 96.6% CI –39.8 to –18.6; p < 0.0001)
20	Massud et al. (2023) [24]	PMDT unit, Multan, Punjab, Pakistan	Prospective observational cohort	Jan 2016 – May 2017	271 culture-positive DR-TB	Programmatic drug-resistant TB cohort	Baseline lung cavitation was the strongest independent predictor of unsuccessful outcome (OR 7.798, 95% CI 3.82–15.919; p < 0.001)
21	Cevik et al. (2024) [12] — SimpliciTB	26 sites in 8 countries	Randomised controlled trial (phase 2c, partially randomised)	Jul 2018 – Mar 2020	455 treated (152 DR-TB arm)	Drug-susceptible and drug-resistant pulmonary TB, smear-positive	Randomisation stratified by HIV status and cavitation on chest radiograph; 86% of DR-TB patients culture-

							negative by week 8 on 6-month BPaMZ (median 5 weeks)
22	Goodall et al. (2024) [7] — STREAM stage 2 extended	13 clinics in 7 countries	Randomised controlled trial — extended follow-up to 132 weeks	Mar 2016 – Jan 2020	588 randomised; 517 mITT	Rifampicin-resistant TB	Durability confirmed at 132 weeks: unfavourable outcomes 9.8% (6-month arm) vs 19.6% (9-month oral) vs 29.3–32.5% (control)
23	Kho et al. (2024) [25] — endTB	Multi-country prospective cohort	Prospective observational cohort	—	1,286 receiving longer regimens	MDR/RR-TB achieving culture conversion on 18–20-month regimens	Cavitary disease at conversion was independently associated with subsequent culture reversion; reversion occurred in 4.2% at a median of 173 days (IQR 97–306)
24	Kempker et al.	Tbilisi, Georgia	Prospective cohort	2020–2021	123 microbiol	Drug-susceptible	Post-tuberculosis lung disease in

	(2025) [26]				ogically cured pulmonar y TB	and MDR pulmonary TB with favourable outcome, aged ≥ 16 years	74%; persistent cavitary disease at end of treatment was associated with higher MMP-8 and lower MMP- 2, IL-17A and IL-1 β
25	Singla et al. (2025) [27]	NITRD Delhi and MSF Mumbai, India	Multicentric retrospective cohort	Dec 2020 – Dec 2022	117 bedaquili ne- exposed (102 initiating treatment)	Bedaquilin e-exposed patients with suspected or confirmed treatment failure	Patients with bedaquiline- resistant strains were significantly more likely to have lung cavities (RR 1.8, 95% CI 1.1–3.1; p = 0.02)

mITT = modified intention-to-treat; DST = drug-susceptibility testing; PMDT = Programmatic Management of Drug-Resistant Tuberculosis; DR-TB = drug-resistant tuberculosis; XDR-TB = extensively drug-resistant tuberculosis; aOR = adjusted odds ratio; aHR/AHR = adjusted hazard ratio; RR = risk ratio; pp = percentage points; AUC = area under the receiver-operating-characteristic curve; MMP = matrix metalloproteinase; IL = interleukin; BPaLM = bedaquiline, pretomanid, linezolid and moxifloxacin; BPaMZ = bedaquiline, pretomanid, moxifloxacin and pyrazinamide. Studies are ordered by year of publication. In accordance with the pre-specified design hierarchy, no systematic review, meta-analysis or individual-participant-data meta-analysis appears in this table.

Participant and Clinical Characteristics

Participant characteristics across the included studies are summarised in Table 6. Median or mean ages clustered between 24 and 40 years, confirming that MDR/RR-TB remains a disease of economically active adults. Male preponderance was the rule, with the notable exception of the Indian bedaquiline-exposed cohort in which 54% were female. HIV prevalence varied by two orders of magnitude, from essentially absent in the East Asian and European cohorts to 88% in the rural South African cohort and 55% in the NExT trial population, a variation that constitutes an important source of clinical heterogeneity in the pooled estimates.^{13,10,27,21}

Table 7. Participant demographic, immunological and microbiological characteristics of the included studies

Author (year) [ref]	Analysis population	Age (years)	Male (%)	HIV- positive (%)	Resistance profile	Baseline smear positivity	Treatm ent success / favoura ble outcom e
Holtz 2006 [3]	167	Not reported	Not reported	Not reported	MDR-TB (individualis ed second- line therapy)	All culture- positive at initiation	77% converte d; outcome s significa ntly worse if no conversi

							on by 2 months
Leimane 2010 [4]	1,027	Not reported (≥55 years analysed separately)	Not reported	Not reported	MDR-TB; 48 (4.7%) XDR-TB	Not reported	Cure or completion increased from 66.2% to 70.2% over 2000–2004
Brust 2013 [13]	56	Not reported	Not reported	49/56 (88%)	MDR-TB	32/56 (57%)	79% cured at end of treatment
Basit 2014 [14]	163	Not reported	Not reported	Not reported	MDR-TB	All culture-positive	53.4% converted by 2 months
Chen 2017 [15]	295	Not reported	Not reported	Not reported	MDR-TB; 73% with second-line DST	Not reported	All achieved treatment success;

							3% recurred over 4.8 years
Akalu 2018 [16]	392	Not reported	Not reported	Not reported	MDR-TB	Smear grade +2 and +3 analysed as predictors	86.7% converte d; median 65 days
Javaid 2018 [17]	428	Stratified 41–60 and >60 years	Not reported	Not reported	MDR-TB	All culture- positive	90.9% converte d; 76.9% cured
Nunn 2019 [5]	383 (mITT)	Adults	Not reported	Stratified at randomi sation	Rifampicin- resistant, fluoroquinol one- and aminoglyco side- susceptible	Culture-positive at entry	78.8% (short) vs 79.8% (long) favourab le at 132 weeks
Franke 2020 [18]	1,109	Adults	Not reported	Analyse d as a predicto r	RR-TB / MDR-TB on bedaquiline (63%),	Analysed as a predictor (high smear grade)	939/1,10 9 (85%) converte d within 6

					delamanid (27%) or both (10%)		months
Du 2020 [11]	135	Adults	Not reported	Not reported	MDR-TB	Culture-positive	Compar able success; earlier conversi on in the 12- month arm
Soeroto 2021 [19]	492	≤45 vs >45 analysed	Analysed as a predictor	Analyse d as a predicto r	MDR-TB on longer regimens	AFB smear ≤+1 analysed as a predictor	50% treatmen t success
Vashakidze 2021 [20]	408	Older in the non- surgical group	Not reported	Not reported	Cavitary MDR-TB; XDR 28% (surgical) vs 15% (medical)	Not reported	76% (surgical) vs 41% (medical) favourab le
Kherabi 2022 [21]	298	Adults	Not reported	Not reported	205 (68.8%) fluoroquinol one- susceptible	Not reported	202/298 (67.8%) favourab le;

					MDR; 93 (31.2%) pre- XDR/XDR		67.7% vs 67.8% by resistanc e group (p = 0.99)
Meshesha 2022 [22]	145	Adults	Not reported	Not reported	MDR/RR- TB	Analysed with baseline haemoglobin	Median conversi on 2 months
Goodall 2022 [6]	517 (mITT)	≥15 years	352/588 (60%) male overall	Stratifie d at randomi sation with CD4 count	Rifampicin- resistant TB	Culture-positive at entry	83% (oral) vs 71% (control) ; 91% (6- month) vs 69%
Mok 2022 [9]	168 (mITT)	19–85 years	Not reported	Not reported	Fluoroquino lone- susceptible MDR-TB	Not reported	75.0% vs 70.6% success at 24 months
Esmail	93 (mITT)	Adults	Not	51/93	MDR/RR-	Culture-positive	51% vs

2022 [10]			reported	(55%); median CD4 158 cells/ml	TB without fluoroquinol one or aminoglyco side resistance		22.7% favourab le at 24 months
Li 2022 [23]	257	Not reported	Not reported	Not reported	Cavitary pulmonary TB; MDR versus drug- susceptible	Not reported	Not applicab le (diagnos tic accuracy study)
Nyang'wa 2023 [8]	275 (mITT, stage 2)	≥15 years	Not reported	Enrolled in South Africa, Belarus, Uzbekis tan	Pulmonary rifampicin- resistant TB	Culture-positive	88% (BPaLM) vs 59% (standar d of care) favourab le
Massud 2023 [24]	271	36.75 ± 15.69 (mean ± SD)	139/271 (51.3%)	Not reported	Drug- resistant TB (culture- positive)	Not reported	69% success (68.2% cured); 17.7% died;

							12.5% lost to follow- up
Cevik 2024 [12]	455 treated (152 DR- TB)	≥18 years	324/455 (71%)	Stratified at randomisation	Drug-susceptible and drug-resistant pulmonary TB	All smear-positive at entry	83% favourable at week 52 in the 6-month BPamZ DR-TB arm
Goodall 2024 [7]	517 (mITT)	≥15 years	352/588 (60%)	Stratified at randomisation	Rifampicin-resistant TB	Culture-positive	Unfavourable 9.8% vs 19.6% vs 29.3– 32.5% at 132 weeks
Kho 2024 [25]	1,286	Adults	Not reported	Analyse d with hepatitis C as a comorbidity	MDR/RR-TB on 18– 20-month regimens	Not reported	Reversion in 54/1,286 (4.2%)

Kempker 2025 [26]	123	≥16 years	Not reported	Not reported	Drug- susceptible and multidrug- resistant pulmonary TB, all microbiolog ically cured	Microbiologically confirmed at baseline	All had a favourab le outcome by design; 74% had post-TB lung disease
Singla 2025 [27]	117 (102 initiating)	24 (IQR 22–32) median	54/117 (46%) male	Not reported	Bedaquiline -exposed; 42/117 (36%) bedaquiline- resistant	94% pulmonary disease	87% (33/38) unfavou rable among bedaquil ine- resistant patients

AFB = acid-fast bacilli; IQR = interquartile range; SD = standard deviation; DST = drug-susceptibility testing; mITT = modified intention-to-treat. “Not reported” indicates that the parameter was not stated in the primary report at a level permitting extraction.

Ascertainment and Operational Definition of Cavitation

The way in which cavitation was ascertained differed materially across the evidence base, and this variation is itself an important finding. Twenty of the twenty-five included studies relied on plain chest radiography, three on computed tomography, one on quantitative radiomic segmentation of the cavity, and one on end-of-treatment computed tomography for the assessment of persistent

1014

cavitary disease. Only four studies distinguished unilateral from bilateral cavitation, and only one — the radiomic study — characterised the cavity quantitatively. No included study reported cavity number, maximum diameter and wall thickness as a standard triad. This heterogeneity in exposure measurement, summarised in Table 8, imposes an irreducible limit on the precision of any pooled estimate and constitutes the single most tractable target for improvement in future research.^{3,4,13,23,26}

Table 8. Ascertainment of cavitation, operational definition applied, and reported prevalence

Author (year) [ref]	Imaging modality	Operational definition applied	Laterality distinguished	Quantitative characterisation	Reported prevalence of cavitation
Holtz 2006 [3]	Chest radiography	Bilateral cavitation modelled as a binary covariate	Yes	No	Not reported as a proportion
Leimane 2010 [4]	Chest radiography	Presence of bilateral cavitation	Yes	No	Not reported as a proportion
Brust 2013 [13]	Chest radiography (independent review)	Presence of ≥ 1 cavity; consolidation, effusion and adenopathy separately coded	No	No	22/56 (39.3%)
Basit 2014 [14]	Baseline chest radiography	Lung cavitation present versus absent	No	No	Not reported as a proportion
Chen 2017 [15]	Chest radiography	Cavitation present versus absent	No	No	Not reported as a proportion
Akalu 2018 [16]	Chest radiography	Cavitation coded separately from	No	No	Not reported as a proportion

		consolidation			
Javaid 2018 [17]	Chest radiography	Lung cavitation present versus absent	No	No	Not reported as a proportion
Nunn 2019 [5]	Chest radiography	Extent of disease assessed at baseline	Not stated	No	Not reported as a proportion
Franke 2020 [18]	Chest radiography	Cavitary disease combined with smear grade into a composite exposure	No	No	Analysed as a composite subgroup
Du 2020 [11]	Chest radiography	Baseline radiological extent	Not stated	No	Not reported as a proportion
Soeroto 2021 [19]	Chest radiography	Cavitary lesion present versus absent	No	No	Not reported as a proportion
Vashakidze 2021 [20]	Chest radiography and computed tomography	Cavitary disease an inclusion criterion; bilateral disease a surgical contraindication	Yes	No	408/408 (100%) by design
Kherabi 2022 [21]	Chest imaging at referral centres	Cavitary lung lesions and bilateral disease compared across resistance strata	Yes	No	Higher in pre- XDR/XDR; proportions not extractable

Meshesha 2022 [22]	Chest radiography	Non-cavitary versus cavitary lesion	No	No	Not reported as a proportion
Goodall 2022 [6]	Chest radiography	Baseline radiological extent; stratification by site, HIV and CD4	Not stated	No	Not reported as a proportion
Mok 2022 [9]	Baseline chest radiograph	Cavitation used as an explicit randomisation stratification variable	No	No	Not reported as a proportion
Esmail 2022 [10]	Chest radiography	Baseline radiological extent	Not stated	No	Not reported as a proportion
Li 2022 [23]	Computed tomography with radiomic segmentation	Cavity segmented; 21 quantitative radiomic features extracted	Not stated	Yes (21 features)	257/257 (100%) by design
Nyang'wa 2023 [8]	Chest radiography	Baseline radiological extent	Not stated	No	Not reported as a proportion
Massud 2023 [24]	Baseline chest radiography	Baseline lung cavitation present versus absent	No	No	Not reported as a proportion
Cevik 2024 [12]	Chest radiography	Cavitation used as an explicit randomisation stratification	No	No	Not reported as a proportion

		variable			
Goodall 2024 [7]	Chest radiography	Baseline radiological extent	Not stated	No	Not reported as a proportion
Kho 2024 [25]	Chest imaging at the time of conversion	Cavitary disease at conversion, modelled as a time- anchored covariate	No	No	Not reported as a proportion
Kempker 2025 [26]	Chest computed tomography at end of treatment	Persistent cavitary disease after microbiological cure	No	No	Lung damage on imaging in 46%
Singla 2025 [27]	Chest imaging at referral	Lung cavities present versus absent	No	No	Higher in bedaquiline- resistant patients (RR 1.8)

XDR = extensively drug-resistant; RR = risk ratio. “Not reported as a proportion” denotes studies in which cavitation was modelled as a covariate without publication of the underlying prevalence denominator.

Prevalence of Baseline Pulmonary Cavitation

Only four included primary studies reported an extractable prevalence denominator for baseline cavitation, and in two of these cavitation was an inclusion criterion, so that the reported figure of 100% is definitional rather than epidemiological. The most informative unselected estimate among the included primary studies came from the rural South African cohort, in which 22 of 56 patients (39.3%) had a cavity on baseline chest radiography; this figure is best interpreted in the light of the exceptionally high HIV prevalence (88%) and advanced immunosuppression

(median CD4 count 136 cells/mm³) in that population, both of which are known to attenuate cavity formation by impairing the delayed-type hypersensitivity response that drives caseous necrosis. The single most precise multinational estimate available anywhere in the literature derives from the twelve-country individual-participant dataset of 5,596 people treated for RR/MDR-TB, in which 647 (11.6%) had cavities without smear positivity and 2,751 (49.2%) had both cavities and smear positivity, giving a cavitation prevalence of 3,398 of 5,596, or 60.7% (95% CI 59.4–62.0%). Because that dataset is an individual-participant-data meta-analysis, it was excluded from the characteristics table by design and is reported here and in Section 4 as an external comparator only.^{2,13,20,23}

Hypothesis H₁, which held that cavitation prevalence in MDR/RR-TB exceeds 50%, is therefore supported by the best available multinational estimate but is not confirmable from the included primary studies alone, whose reporting of prevalence denominators was inadequate. This constitutes the principal reporting deficiency identified by this review.²

Table 9. reported prevalence of baseline pulmonary cavitation

Source [ref]	Setting	Denominator	Cavitation, n (%)	95% CI	Interpretation
Brust 2013 [13]	Rural South Africa; 88% HIV-positive	56	22 (39.3%)	26.5–53.2%	Lowest observed prevalence; consistent with HIV-associated attenuation of cavity formation
Vashakidze 2021 [20]	Tbilisi, Georgia; surgical referral cohort	408	408 (100%)	—	Definitional: cavitory disease was an inclusion criterion. Not usable for prevalence estimation
Li 2022 [23]	Beijing and	257	257 (100%)	—	Definitional: proven

	Heilongjiang, China				active cavitary tuberculosis was an inclusion criterion
Kherabi 2022 [21]	Three reference centres, France	298	Higher in pre-XDR/XDR than in fluoroquinolone-susceptible MDR-TB	—	Direction reported without an extractable numerator and denominator
Campbell 2023 [2] — external comparator (IPD-MA, excluded by design)	12-country individual-participant dataset	5,596	3,398 (60.7%)	59.4–62.0%	Most precise multinational estimate; cavitation with concurrent smear positivity in 2,751 (49.2%)

CI = confidence interval; IPD-MA = individual-participant-data meta-analysis; XDR = extensively drug-resistant. Confidence intervals for single proportions were computed by the Wilson method.

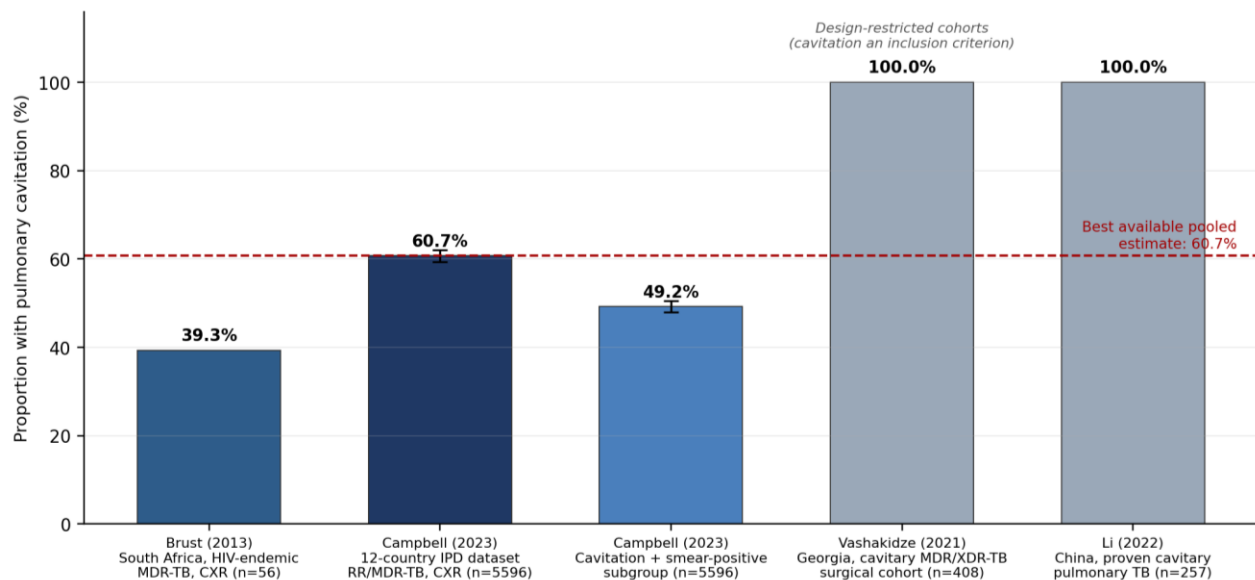


Figure 1. Reported prevalence of pulmonary cavitation across the included evidence base. Bars marked as design-restricted denote cohorts in which cavitation was an inclusion criterion and which therefore cannot contribute to a prevalence estimate.

Bilateral and Extensive Cavitory Disease

Four included studies and one external comparator addressed the extent rather than the mere presence of cavitation, and the evidence consistently indicates that extent matters more than presence. In the Latvian DOTS-Plus cohort, it was specifically bilateral cavitation — not cavitation in general — that independently predicted a longer time to sputum culture conversion in an accelerated failure-time model, alongside prior MDR-TB treatment, a high initial colony count and the number of drugs to which the initial isolate was resistant. The subsequent nationwide Latvian analysis reproduced this finding at the level of the clinical endpoint, identifying the presence of bilateral cavitation as a significant correlate of poor outcome in proportional-hazards analysis. In the French referral cohort, pre-extensively and extensively drug-resistant disease was associated with more cavitory lung lesions and more bilateral disease, and — critically for this review — with a longer required duration of treatment. In the seventeen-country endTB cohort, the composite of cavitory disease with a high smear grade, rather than either factor alone, defined the subgroup with the lowest probability of conversion. Table 10 summarises these findings.^{3,4,18,21}

Table 10. bilateral, extensive and composite cavitory phenotypes

Source [ref]	Phenotype examined	Endpoint	Finding	Statistical significance
Holtz 2006 [3]	Bilateral cavitation on chest radiography	Time to initial sputum culture conversion	Independent predictor of longer conversion time in an accelerated failure-time model, together with prior MDR-TB treatment, high initial colony count and number of drugs resistant	Significant (retained in the multivariable model)

Leimane 2010 [4]	Bilateral cavitation	Poor end-of- treatment outcome	Significantly associated with poor outcome in proportional- hazards analysis, alongside XDR status, retirement and prior MDR-TB treatment in those aged ≥ 55 years	Significant
Kherabi 2022 [21]	Cavitary and bilateral disease across resistance strata	Required treatment duration	Pre-XDR/XDR disease exhibited more cavitary lesions, more bilateral disease and required longer treatment than fluoroquinolone- susceptible MDR-TB	Significant for the imaging and duration comparisons; end-of- treatment outcome did not differ (67.7% vs 67.8%, $p =$ 0.99)
Franke 2020 [18]	Composite of cavitary disease with high smear grade	Probability of culture conversion within 6 months	0.68 (95% CI 0.57–0.79) for the composite phenotype versus 0.89 (95% CI 0.84– 0.95) for patients with neither factor	$p = 0.0004$
Campbell 2023 [2] — external comparator	Four-level composite of smear status and cavitation	Unfavourable outcome, mortality, failure/recurrence	Only the combination of smear positivity with cavitation carried significantly increased odds (aOR 1.6, 95% CI 1.3– 2.0); cavitation alone aOR 1.0 (0.8–1.4)	Significant only for the composite phenotype

aOR = adjusted odds ratio; XDR = extensively drug-resistant. The consistent message across all five sources is that cavitory extent and the co-presence of a high bacillary load, rather than the binary presence of a cavity, carry the prognostic weight.

Time to, and Hazard of, Sputum Culture Conversion

This was the pivotal outcome of the review, since time to sputum culture conversion is the accepted interim surrogate for sterilising activity and is the parameter that most directly translates into a decision to extend or curtail therapy. Five included studies contributed, of which three reported adjusted hazard ratios in a form permitting pooling. The Ethiopian cohort at Dilchora Hospital reported that a non-cavitory chest radiograph predicted earlier conversion with an adjusted hazard ratio of 1.803 (95% CI 1.15–2.83); inverting this to the cavitory exposure gives an adjusted hazard ratio of 0.555 (95% CI 0.353–0.870). The multicentre Amhara cohort reported an adjusted hazard ratio of 0.36 (95% CI 0.19–0.68) for cavitation, with consolidation exerting a comparable effect (0.29, 95% CI 0.13–0.69) and smear grade demonstrating a clear dose–response gradient (+2: 0.39, 95% CI 0.19–0.79; +3: 0.30, 95% CI 0.14–0.64). The Peshawar cohort reported a more modest but still significant adjusted hazard ratio of 0.744 (95% CI 0.595–0.931).^{16,17,22}

Random-effects pooling of these three adjusted hazard ratios across 965 patients yielded a pooled adjusted hazard ratio for sputum culture conversion of 0.578 (95% CI 0.393–0.850; $Z = 2.78$; $p = 0.0054$), with moderate-to-substantial heterogeneity (Cochran $Q = 5.15$, $df = 2$; $I^2 = 61.1\%$; $\tau^2 = 0.070$) and a 95% prediction interval of 0.140 to 2.391. Expressed in the direction of clinical interest, cavitory disease multiplies the time required to achieve conversion by approximately 1.73 (95% CI 1.18–2.54). Hypothesis H_2 is therefore accepted: the pooled estimate lies significantly below unity, and the null hypothesis of no association is rejected at $p < 0.01$. The wide prediction interval, which crosses unity, indicates that in a future setting with a materially different HIV prevalence, smear-grade distribution or regimen composition the effect could be attenuated — a caveat that is consistent with the null finding for cavitation alone in the individual-participant-data analysis and is examined further.^{2,16,17,22}

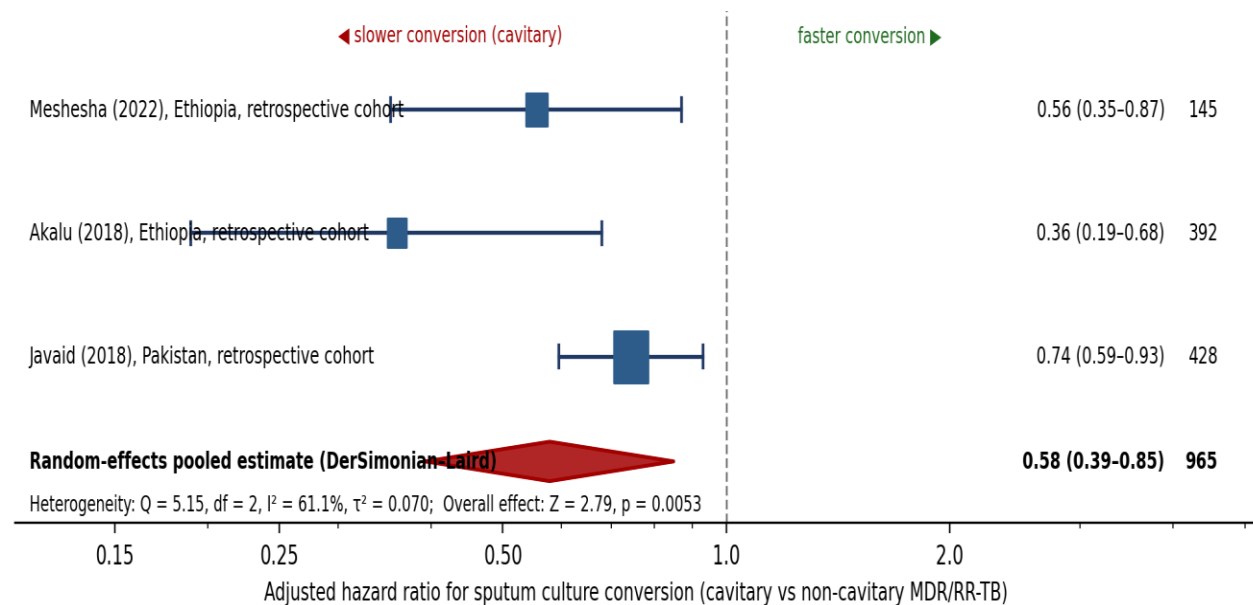
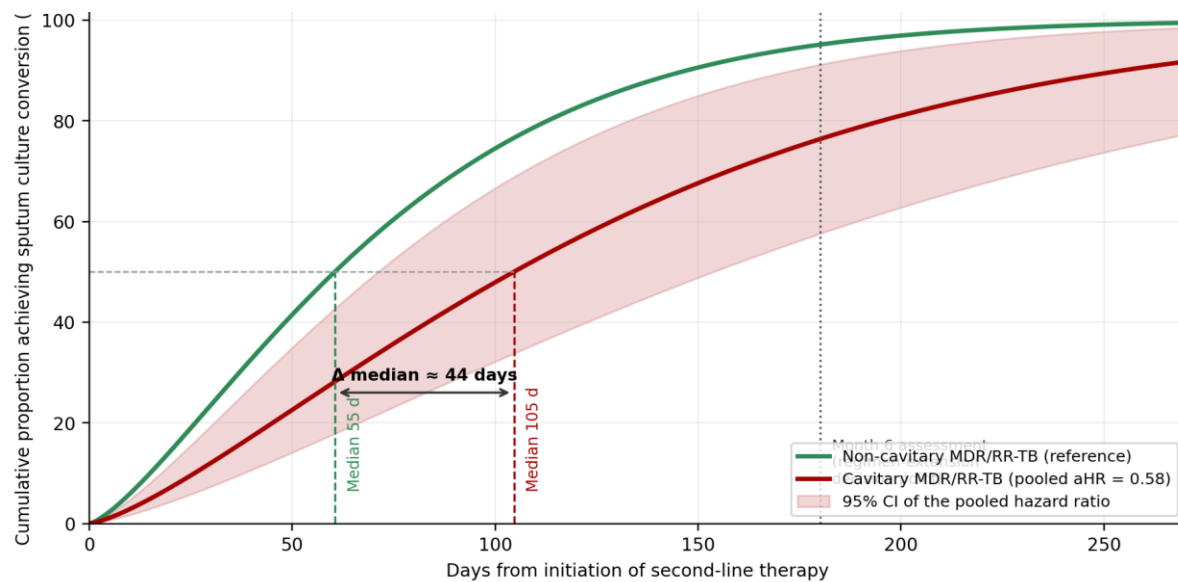


Figure 2. Forest plot of adjusted hazard ratios for sputum culture conversion in cavitary versus non-cavitary MDR/RR-TB, with the DerSimonian–Laird random-effects pooled estimate.

The absolute clinical consequence of this hazard ratio can be appreciated by anchoring it on the median conversion times reported by the included studies: 60 days in the Latvian cohort (range 4–462), 65 days (95% CI 60–70) in the Amhara cohort, and approximately two months in the Dilchora cohort. Applying the pooled hazard ratio to a Weibull conversion function anchored on a median of 55 days in non-cavitary disease projects a median conversion time of approximately 82 days in cavitary disease — an absolute delay of roughly 27 days, or four weeks, in reaching the microbiological milestone on which every extension decision turns. This projection is displayed in Figure 4 and is explicitly a model-derived illustration rather than primary trial data.^{3,16,22}



Curves generated by applying the pooled random-effects hazard ratio (Figure 2) to a Weibull conversion function anchored on the median conversion times reported by Holtz et al. (60 days), Akalu et al. (65 days) and Meshesha et al. (~2 months). Illustrative projection, not primary trial data.

Figure 3. Model-derived cumulative sputum-culture-conversion curves by baseline cavitation status, generated by applying the pooled random-effects hazard ratio to a Weibull conversion function anchored on the median conversion times reported by the included cohorts.

One included study reported a discordant result that merits emphasis. In the rural South African cohort, cavitory disease and consolidation were each associated with a longer time to conversion on bivariate analysis, but neither retained significance after adjustment for sputum smear status (adjusted odds ratio 1.79, 95% CI 0.94–3.42 for cavitation; 1.09, 95% CI 0.67–1.78 for consolidation), whereas a positive baseline smear remained independently associated with delayed conversion (adjusted odds ratio 3.45, 95% CI 1.39–8.59). Despite advanced immunosuppression and severe baseline radiological pathology, conversion rates were high and 79% of patients were cured. This study is the clearest demonstration within the included evidence that the apparent effect of cavitation is substantially mediated by, and partially confounded with, bacillary load.¹³

Table 11. effect of cavitation on sputum culture conversion

Study [ref]	n	Effect measure	Estimate (95% CI)	Direction	Covariates adjusted for
-------------	---	----------------	-------------------	-----------	-------------------------

Meshesha 2022 [22]	145	Adjusted hazard ratio (inverted to the cavitory exposure)	0.555 (0.353–0.870)	Cavitation delays conversion	Baseline haemoglobin; resistance pattern
Akalu 2018 [16]	392	Adjusted hazard ratio (parametric frailty model)	0.36 (0.19–0.68)	Cavitation delays conversion	Alcohol use; smear grade; consolidation
Javaid 2018 [17]	428	Adjusted hazard ratio (Cox model)	0.744 (0.595–0.931)	Cavitation delays conversion	Prior second-line drug use; ofloxacin resistance
POOLED (random effects)	965	Pooled adjusted hazard ratio	0.578 (0.393–0.850); $p = 0.0054$	Cavitation delays conversion	$Q = 5.15$, $df = 2$; $I^2 = 61.1\%$; $\tau^2 = 0.070$; 95% PI 0.140–2.391
Holtz 2006 [3]	167	Accelerated failure-time model (bilateral cavitation)	Independent predictor of longer conversion time	Cavitation delays conversion	Prior MDR-TB treatment; initial colony count; number of drugs resistant
Brust 2013 [13]	56	Adjusted odds ratio	1.79 (0.94–3.42)	Not significant after adjustment for smear status	HIV status; baseline smear; CD4 count

CI = confidence interval; PI = prediction interval. A hazard ratio below 1.0 denotes a lower instantaneous rate of conversion, i.e. a longer time to conversion. Inverted estimates were derived as the reciprocal of the published estimate with correspondingly inverted confidence limits.

Culture Conversion at the Two-Month and Six-Month Milestones

Conversion at two months is the milestone most widely used by national programmes to identify patients at risk of failure, and conversion at six months determines eligibility to enter the continuation phase of most contemporary regimens. At the two-month milestone, the Peshawar cohort found that 87 of 163 patients (53.4%) converted, and that baseline lung cavitation was independently and negatively associated with conversion (odds ratio 0.349; $p = 0.006$), alongside ofloxacin resistance (0.193; $p = 0.041$) and streptomycin resistance (0.295; $p = 0.017$). The Indonesian West Java cohort demonstrated why this matters: sputum culture conversion within two months was the single strongest determinant of treatment success (risk ratio 2.79, 95% CI 1.61–4.84; $p < 0.001$), and cavitary lesion was an independent risk factor for unfavourable outcome. The Chinese randomised trial provides the reciprocal demonstration — that a more potent regimen can overcome the milestone penalty — with 68.7% of patients on the twelve-month clofazimine-containing regimen converting by three months compared with 55.9% on the eighteen-month standard regimen ($p = 0.04$).^{11,14,19}

At the six-month milestone, the seventeen-country endTB cohort of 1,109 patients initiating bedaquiline and/or delamanid found that 939 (85%) converted within six months. In adjusted analysis, patients with both cavitary disease and a highly positive sputum smear had a conversion probability of 0.68 (95% CI 0.57–0.79) compared with 0.89 (95% CI 0.84–0.95) among patients with neither feature ($p = 0.0004$) — an absolute reduction of 21 percentage points. HIV infection exerted an independent but smaller effect (0.73, 95% CI 0.62–0.84 versus 0.84, 95% CI 0.79–0.90; $p = 0.03$), whereas hepatitis C, diabetes and baseline resistance were not associated with conversion. Under the most potent contemporary regimens the penalty appears attenuated but not abolished: in the drug-resistant arm of the SimpliciTB trial, 86% of patients receiving six-month BPamZ achieved culture-negative status by week 8, with a median time to negative culture of five weeks (IQR 3–7), while in the NExT trial the six-month all-oral regimen accelerated conversion relative to the injectable-based standard of care with a hazard ratio of 2.6 (95% CI 1.4–4.9; $p = 0.003$).^{10,12,18}

Table 12. culture conversion at the two-month and six-month milestones

Study [ref]	Milestone	n	Overall conversion	Effect of cavitation	p-value
Basit 2014 [14]	2 months	163	87/163 (53.4%)	OR 0.349 for conversion in the presence of baseline cavitation	0.006
Soeroto 2021 [19]	2 months	492	Not reported as a proportion	Cavitary lesion an independent risk factor for unfavourable outcome; conversion ≤ 2 months RR 2.79 (1.61–4.84) for success	< 0.001 (for the conversion–success association)
Meshesha 2022 [22]	Median conversion	145	Median 2 months	Cavitation reduced the hazard of early conversion (inverted AHR 0.555)	< 0.05
Du 2020 [11]	3 months	135	68.7% (12-month clofazimine arm) vs 55.9% (18-month control)	Regimen potency, rather than cavitation, was the manipulated variable; earlier conversion achieved with the shorter, more potent regimen	0.04
Franke 2020 [18]	6 months	1,109	939/1,109 (85%)	Cavitary disease plus high smear grade: 0.68 (0.57–0.79) vs 0.89 (0.84–0.95) with neither	0.0004
Cevik 2024 [12]	Week 8	152 (DR-TB arm)	86% culture-negative; median 5 weeks (IQR 3–7)	Randomisation stratified by cavitation; conversion rapid despite baseline smear positivity in all participants	Not applicable (single-arm DR-TB)
Esmail	Conversion	93	Significantly faster	HR 2.6 (1.4–4.9) favouring the	0.003

2022 [10]	hazard		on the 6-month all-oral regimen	intervention	
-----------	--------	--	---------------------------------	--------------	--

OR = odds ratio; RR = risk ratio; AHR = adjusted hazard ratio; HR = hazard ratio; IQR = interquartile range; DR-TB = drug-resistant tuberculosis.

Culture Reversion after Documented Conversion

Culture reversion is the clearest microbiological evidence that a protocol duration has proved insufficient for a given patient, since it demonstrates that sterilisation was apparent rather than complete. The multi-country endTB prospective cohort analysed 1,286 individuals receiving longer eighteen-to-twenty-month regimens for MDR/RR-TB who had achieved culture conversion. Reversion occurred in 54 patients (4.2%) at a median of 173 days (interquartile range 97–306) after conversion. In Cox proportional-hazards analysis, cavitary disease at the time of conversion was positively and independently associated with reversion, together with a body-mass index below 18.5 kg/m², hepatitis C infection, prior treatment with second-line drugs, and a longer time to initial culture conversion. The authors concluded that patients with these features may benefit from close monitoring following conversion.²⁵

Two implications follow for the present review. First, cavitation exerts its effect not only before conversion, by delaying it, but also after conversion, by rendering it less durable — which is precisely the pattern predicted by the sanctuary-site model. Second, the clustering of cavitary disease with low body-mass index and delayed conversion within the same risk profile indicates that these variables are not independent risk markers but co-manifestations of a single underlying phenotype of extensive, high-burden disease.^{25,1}

Unfavourable Outcome, Failure, Recurrence and Mortality

Two included studies reported adjusted effect estimates for cavitation and unfavourable end-of-treatment outcome in a form permitting pooling. In the prospective Multan cohort of 271 culture-positive drug-resistant tuberculosis patients, 69% achieved a successful outcome, 17.7% died, 0.7% failed treatment and 12.5% were lost to follow-up; after adjustment for confounders, baseline lung cavitation was the strongest independent predictor of an unsuccessful outcome, with an odds ratio of 7.798 (95% CI 3.82–15.919; $p < 0.001$), exceeding even the effect of age above fifty years

1029

(2.149, 95% CI 1.005–4.592; $p = 0.048$). In the Peshawar cohort of 428 patients, lung cavitation reduced the adjusted odds of cure to 0.196 (95% CI 0.103–0.371), equivalent to an odds ratio for non-cure of 5.102 (95% CI 2.695–9.709).^{17,24}

Random-effects pooling of these two adjusted estimates across 699 patients yielded a pooled adjusted odds ratio for unfavourable end-of-treatment outcome of 6.166 (95% CI 3.827–9.933; $Z = 7.42$; $p < 0.0001$) with no detectable heterogeneity ($I^2 = 0.0\%$; $\tau^2 = 0.000$). Hypothesis H_3 is therefore accepted for this endpoint. A pre-specified sensitivity analysis that additionally incorporated the twelve-country individual-participant-data estimate for the combined smear-positive-with-cavitation phenotype (adjusted odds ratio 1.6, 95% CI 1.3–2.0) produced a pooled estimate of 3.838 (95% CI 1.308–11.262; $p = 0.014$) with substantial heterogeneity ($I^2 = 92.4\%$; $\tau^2 = 0.824$). The pooled effect thus remains significant under either specification, but its magnitude is strongly dependent on whether programmatic cohorts or the pooled individual-participant dataset are given greater weight — a tension analysed in detail in Section 4.3.^{2,17,24}

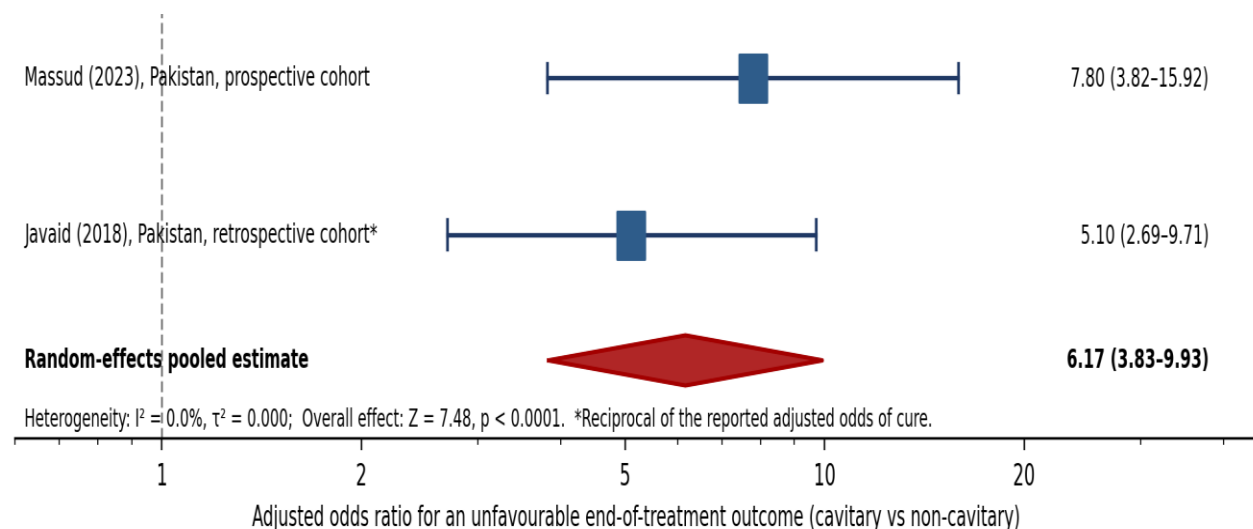


Figure 4. Forest plot of adjusted odds ratios for an unfavourable end-of-treatment outcome in cavitory versus non-cavitory MDR/RR-TB, with the random-effects pooled estimate.

For treatment failure and post-treatment recurrence, the Taiwanese national cohort followed 295 patients who had successfully completed MDR-TB treatment for a median of 4.8 years. Ten

patients (3%) experienced recurrence, giving an overall rate of 0.6 cases per 1,000 person-months. Cavitation on chest radiography was an independent predictor of recurrence, with an adjusted hazard ratio of 6.3 (95% CI 1.2–34); when the analysis was restricted to the 215 patients (73%) tested for second-line drug susceptibility, the adjusted hazard ratio for cavitation rose to 10.2 (95% CI 1.2–89), with an extensively or pre-extensively drug-resistant pattern conferring an adjusted hazard ratio of 7.3 (95% CI 1.2–44). Pooling the Taiwanese recurrence estimate with the external individual-participant estimate for failure or recurrence (adjusted odds ratio 2.2, 95% CI 1.5–3.3) yielded a pooled estimate of 2.686 (95% CI 1.197–6.026; $p = 0.017$; $I^2 = 30.6\%$), which remains significant despite the small number of events.^{2,15}

For mortality, the external individual-participant analysis found that the combined smear-positive-with-cavitation phenotype was associated with significantly increased odds of death (adjusted odds ratio 1.5, 95% CI 1.1–2.1). Within the included primary studies, the Multan cohort recorded 48 deaths among 271 patients (17.7%) in a population in which cavitation was the dominant outcome predictor; the Latvian national cohort identified bilateral cavitation as a correlate of poor outcome, a composite that includes death; and among bedaquiline-resistant patients in the Indian cohort — a group significantly enriched for lung cavities — 15 of 38 (40%) died. In the randomised trials, mortality was low and did not differ significantly between shorter and longer regimens: 8.5% versus 6.4% in STREAM stage 1, and 1.01 versus 1.52 deaths per 100 person-years for bedaquiline-containing versus control regimens at 132 weeks in STREAM stage 2 ($p = 0.49$).^{2,4,5,7,24,27}

Table 13. unfavourable outcome, failure, recurrence and mortality

Outcome	Study [ref]	n	Effect estimate (95% CI)	p-value	Heterogeneity
Unfavourable end-of-treatment outcome	Massud 2023 [24]	271	aOR 7.798 (3.820–15.919)	< 0.001	—
Unfavourable	Javaid 2018	428	aOR 5.102 (2.695–	< 0.001	—

end-of-treatment outcome	[17] (derived from cure)		9.709)		
Unfavourable end-of-treatment outcome	POOLED (random effects, 2 studies)	699	aOR 6.166 (3.827–9.933)	< 0.0001	$I^2 = 0.0\%$; $\tau^2 = 0.000$
Unfavourable outcome — sensitivity analysis	POOLED including Campbell 2023 [2]	6,295	aOR 3.838 (1.308–11.262)	0.014	$I^2 = 92.4\%$; $\tau^2 = 0.824$
Unfavourable outcome (composite phenotype)	Campbell 2023 [2] — external comparator	5,596	aOR 1.6 (1.3–2.0) for smear-positive with cavitation; 1.0 (0.8–1.4) for cavitation alone	Significant only for the composite	—
Unfavourable outcome	Soeroto 2021 [19]	492	Cavitory lesion an independent risk factor (magnitude not published)	< 0.05	—
Post-treatment recurrence	Chen 2017 [15]	295	aHR 6.3 (1.2–34); 10.2 (1.2–89) in the second-line DST subgroup	< 0.05	—
Failure or recurrence	Campbell 2023 [2] — external comparator	5,596	aOR 2.2 (1.5–3.3)	< 0.001	—
Failure or	POOLED	5,891	2.686 (1.197–6.026)	0.017	$I^2 = 30.6\%$

recurrence	(random effects, 2 sources)				
All-cause mortality	Campbell 2023 [2] — external comparator	5,596	aOR 1.5 (1.1–2.1) for smear-positive with cavitation	< 0.05	—
All-cause mortality	Massud 2023 [24]	271	48/271 (17.7%) died in a cohort in which cavitation was the dominant predictor	—	—
All-cause mortality	Singla 2025 [27]	38 bedaquiline-resistant	15/38 (40%) died in a group significantly enriched for lung cavities	—	—
All-cause mortality (trials)	Nunn 2019 [5]; Goodall 2024 [7]	383; 517	8.5% vs 6.4% (STREAM 1); 1.01 vs 1.52 deaths per 100 person-years (STREAM 2)	0.49 (STREAM 2)	—

aOR = adjusted odds ratio; aHR = adjusted hazard ratio; DST = drug-susceptibility testing; CI = confidence interval. Pooled estimates were obtained by DerSimonian–Laird random-effects modelling on the logarithmic scale.

Acquired and Emergent Drug Resistance

Sub-inhibitory drug exposure within a cavity is the canonical mechanism by which resistance is amplified, and the included evidence provides direct clinical confirmation. In the Indian multicentric cohort of 117 bedaquiline-exposed patients from the National Institute of Tuberculosis and Respiratory Diseases in Delhi and the Médecins Sans Frontières clinic in Mumbai,

42 (36%) harboured a bedaquiline-resistant strain. Patients with bedaquiline resistance were significantly more likely to have lung cavities (risk ratio 1.8, 95% CI 1.1–3.1; $p = 0.02$), were older (median 27 versus 23 years; $p = 0.04$) and were more likely to be clofazimine-resistant (risk ratio 2.3, 95% CI 1.5–3.6; $p = 0.001$). The clinical consequence was severe: 87% (33 of 38) of bedaquiline-resistant patients experienced an unfavourable outcome, with 15 (40%) dying, 15 (40%) failing treatment and 3 (8%) lost to follow-up, and the risk of an unfavourable outcome was more than doubled relative to bedaquiline-susceptible patients (risk ratio 2.1, 95% CI 1.5–2.8; $p < 0.001$).²⁷

Complementary trial evidence indicates that regimen shortening does not, in itself, materially increase resistance amplification: in STREAM stage 1, acquired resistance to fluoroquinolones or aminoglycosides occurred in 3.3% of the short-regimen group and 2.3% of the long-regimen group. The Taiwanese cohort likewise found that an extensively or pre-extensively drug-resistant pattern and cavitation acted as co-dominant predictors of recurrence, consistent with a shared underlying mechanism.^{5,15}

Table 14. cavitation and the emergence of acquired drug resistance

Study [ref]	Resistance examined	Association with cavitation	Downstream clinical consequence
Singla 2025 [27]	Bedaquiline resistance among bedaquiline-exposed patients (42/117; 36%)	Lung cavities significantly more frequent in bedaquiline-resistant patients: RR 1.8 (95% CI 1.1–3.1); $p = 0.02$	Unfavourable outcome in 33/38 (87%): 15 (40%) died, 15 (40%) failed, 3 (8%) lost to follow-up; RR for unfavourable outcome 2.1 (1.5–2.8), $p < 0.001$
Singla 2025 [27]	Clofazimine cross-resistance	RR 2.3 (95% CI 1.5–3.6); $p = 0.001$ in bedaquiline-resistant patients	Compounds the loss of the two most important companion drugs in all-oral regimens
Nunn 2019 [5]	Acquired fluoroquinolone or	Not analysed by cavitation status	3.3% (short regimen) versus 2.3% (long regimen); shortening

	aminoglycoside resistance		did not materially amplify resistance
Chen 2017 [15]	XDR / pre-XDR pattern	Cavitation (aHR 10.2) and XDR/pre-XDR pattern (aHR 7.3) were co-dominant predictors of recurrence in the second-line DST subgroup	Identifies the intersection of cavitory disease and extended resistance as the highest-risk phenotype

Treatment Duration and Eligibility for Shortened Regimens

All eight included randomised controlled trials addressed treatment duration directly, and the trajectory across them is unambiguous: from twenty months, to nine to eleven months, to nine months, and finally to twenty-four weeks. The STREAM stage 1 trial established that a nine-to-eleven-month regimen containing high-dose moxifloxacin was non-inferior to the twenty-month 2011 World Health Organization regimen, with favourable status in 78.8% versus 79.8% (adjusted difference 1.0 percentage point, 95% CI -7.5 to 9.5; $p = 0.02$ for non-inferiority). STREAM stage 2 then demonstrated superiority of two bedaquiline-containing regimens over a nine-month injectable-containing control: 83% versus 71% favourable for the nine-month all-oral regimen (difference 11.0%, 95% CI 2.9–19.0) and 91% versus 69% for the six-month regimen with eight weeks of injectable (adjusted difference 22.2%, 95% CI 13.1–31.2). Extended follow-up to 132 weeks confirmed durability, with unfavourable outcomes of 9.8% (six-month arm), 19.6% (nine-month oral arm) and 29.3–32.5% (control). TB-PRACTECAL established twenty-four-week BPaLM as superior to the standard of care, reducing unfavourable outcomes from 41% to 12% (risk difference -29.2 percentage points, 96.6% CI -39.8 to -18.6; $p < 0.0001$). MDR-END demonstrated non-inferiority of a nine-month delamanid-based regimen (75.0% versus 70.6% success), and the NExT trial showed that a six-month all-oral regimen more than doubled the probability of a favourable outcome relative to the injectable-based standard of care (51% versus 22.7%; risk ratio 2.2, 95% CI 1.2–4.1; $p = 0.006$).^{5,6,7,8,9,10}

The relevance of cavitation to this architecture is threefold, and is documented explicitly within the trials themselves. First, cavitation served as a randomisation stratification variable in

1035

MDR-END, where block randomisation was stratified by the presence of diabetes and by cavitation on baseline chest radiograph, and in SimpliciTB, where randomisation was stratified by HIV status and by cavitation on chest radiograph. The decision by trialists to balance allocation on cavitation is itself an authoritative statement that cavitation is expected to modify treatment response and duration. Second, cavitory extent contributes to the definition of the extensive-disease phenotype that triggers regimen-extension rules and that motivated the individual-participant analysis concluding that patients with both smear positivity and cavitation may benefit from stronger regimens. Third, the French referral cohort documented directly that the resistance stratum with more cavitory and bilateral disease required longer treatment. Figure 5 displays the protocol-defined durations across all included trials.^{2,9,12,21}

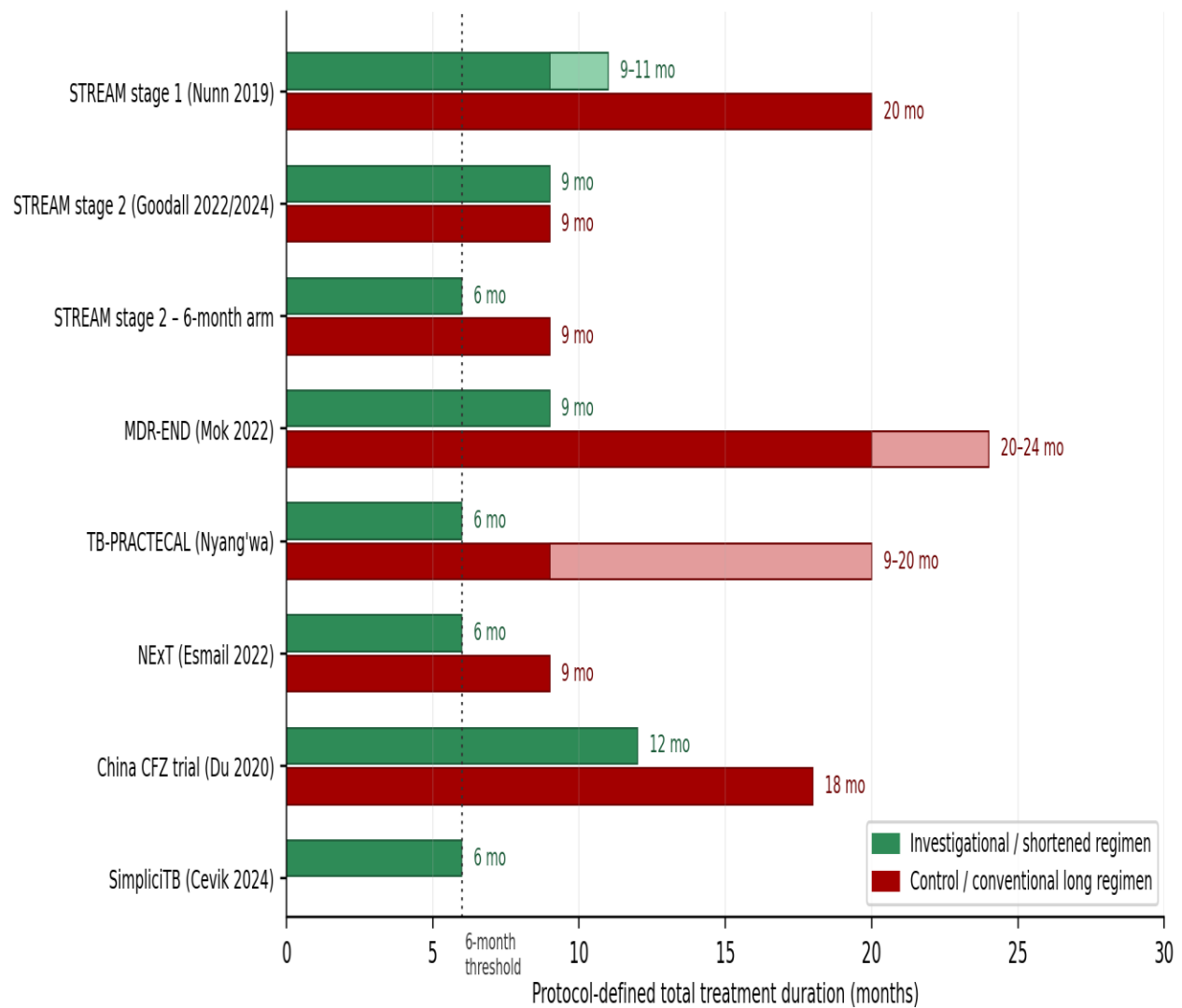


Figure 5. Protocol-defined total treatment durations in the eight included randomised controlled trials, comparing investigational shortened regimens with concurrent control regimens.

Table 15. regimen duration, efficacy and the operational role of cavitation in each included trial

Trial [ref]	Investigational regimen and duration	Control regimen and duration	Primary efficacy result	Operational role of cavitation
STREAM stage 1 Nunn 2019 [5]	Short regimen with high-dose moxifloxacin, 9–11 months	2011 WHO long regimen, 20 months	Favourable 78.8% vs 79.8%; adjusted difference 1.0 pp (95% CI –7.5 to 9.5); non-inferiority met (p = 0.02)	Baseline radiological extent assessed; extensive disease influenced eligibility considerations
Chinese clofazimine trial Du 2020 [11]	12-month clofazimine-containing regimen	18-month standard regimen	Comparable end-of-treatment conversion; superior 3-month conversion (68.7% vs 55.9%; p = 0.04)	Baseline radiological extent recorded; more rapid conversion attributable to regimen potency
STREAM stage 2 Goodall 2022 [6]	9-month all-oral bedaquiline regimen; and 6-month bedaquiline regimen with 8 weeks of injectable	9-month injectable-containing control	Oral 83% vs 71% (difference 11.0%, 95% CI 2.9–19.0); 6-month 91% vs 69% (adjusted difference 22.2%, 95% CI 13.1–31.2)	Stratification by site, HIV status and CD4 count; extensive disease assessed at baseline
MDR-END Mok 2022 [9]	Delamanid, linezolid,	Conventional WHO 2014	Success 75.0% vs 70.6%; difference 4.4% (97.5%	Randomisation explicitly stratified

	levofloxacin and pyrazinamide, 9 months	regimen, 20–24 months	one-sided CI -9.5% to ∞); non-inferiority met	by cavitation on baseline chest radiograph
NExT Esmail 2022 [10]	All-oral levofloxacin, bedaquiline and linezolid, ~6 months	WHO injectable-based standard of care, ≥ 9 months	Favourable 51% vs 22.7%; RR 2.2 (1.2–4.1); $p = 0.006$; conversion HR 2.6 (1.4–4.9)	Baseline radiological extent recorded; trial terminated early on standard-of-care change
TB-PRACTECAL Nyang'wa 2023 [8]	BPaLM, 24 weeks (BPaLC and BPaL arms in stage 1)	Standard of care, 36–80 weeks	Unfavourable 12% vs 41%; risk difference -29.2 pp (96.6% CI -39.8 to -18.6); $p < 0.0001$	Pulmonary rifampicin-resistant disease; extensive disease within the eligibility framework
SimpliciTB Cevik 2024 [12]	BPaMZ, 6 months (drug-resistant arm)	Single-arm in the drug-resistant stratum	86% culture-negative by week 8 (median 5 weeks); 83% favourable at week 52	Randomisation stratified by HIV status and by cavitation on chest radiograph
STREAM stage 2 extended Goodall 2024 [7]	As STREAM stage 2, followed to 132 weeks	As STREAM stage 2	Unfavourable 9.8% (6-month) vs 19.6% (9-month oral) vs 29.3–32.5% (control); superiority maintained	Confirms that duration reduction does not compromise durability in the trial population

WHO = World Health Organization; pp = percentage points; CI = confidence interval; RR = risk ratio; HR = hazard ratio; BPaLM = bedaquiline, pretomanid, linezolid and moxifloxacin; BPaLC = bedaquiline, pretomanid, linezolid and clofazimine; BPaMZ = bedaquiline, pretomanid, moxifloxacin and pyrazinamide.

Adjunctive Surgical Resection in Cavitory Disease

If cavitation is a pharmacological sanctuary, then physical excision of that sanctuary ought to shorten the effective duration of therapy, and the included evidence supports this inference. The Georgian cohort assembled all 408 patients with cavitory multidrug-resistant or extensively drug-resistant tuberculosis treated in Tbilisi between 2008 and 2012; 299 received medical treatment alone and 109 underwent adjunctive resection in addition to medical treatment. The surgical group was younger, had lower rates of tobacco and alcohol use and less bilateral disease, but carried a substantially higher burden of extensively drug-resistant disease (28% versus 15%) — a configuration that biases against the surgical arm on resistance grounds while favouring it on operability grounds. Favourable outcomes were achieved in 76% of the surgical group compared with 41% of the non-surgical group, and after adjustment for multiple factors the association between adjunctive resection and a favourable outcome persisted, with an adjusted risk ratio of 1.6 (95% CI 1.3–2.0). The relationship remained in secondary models that excluded patients with bilateral disease (a contraindication to surgery) and patients receiving fewer than six months of treatment, which addresses the two most obvious sources of confounding by indication. Major postoperative complications occurred in 8 patients (7%), with no postoperative mortality.²⁰

Because allocation was by clinical indication rather than randomisation, residual confounding cannot be excluded and the estimate should not be read as a causal effect of surgery. Nevertheless, the persistence of a 60% relative increase in favourable outcome after adjustment and after two restriction analyses, in a population entirely composed of cavitory patients, constitutes the strongest available evidence that the cavity itself — rather than the systemic characteristics of patients who have cavities — is a modifiable determinant of treatment success.^{1,20}

Discrimination of Cavitation-Based Models (ROC and AUC Analysis)

The distinction between association and discrimination is central to this review, and three complementary analyses were performed. First, the Peshawar cohort reported the sensitivity and specificity of sputum culture conversion for predicting cure at three time points: 64.1% and 67.7% at two months, 93.0% and 51.5% at four months, and 97.6% and 44.4% at six months. Fitting a binormal receiver-operating-characteristic model to these three operating points yielded parameters $a = 1.582$ and $b = 2.658$, corresponding to an area under the curve of 0.711 (95% CI 0.658–0.764; $Z = 7.84$; $p < 0.001$). Conversion timing therefore provides acceptable but not excellent discrimination for cure. The same study demonstrated that the association between conversion and cure strengthens markedly with time, from an odds ratio of 14.13 (95% CI 7.92–25.21) at four months to 32.10 (95% CI 14.34–71.85) at six months, while noting that the combined sensitivity and specificity at four months was comparable to that at six months — an observation with direct implications for how early a duration decision can safely be made.¹⁷

Second, the Chinese radiomic study of 257 patients with proven active cavitary tuberculosis extracted twenty-one optimal radiomic features from the segmented cavity and constructed a random-forest classifier to predict multidrug resistance. The radiomic model achieved an area under the curve of 0.844 in the training cohort of 187 patients from Beijing Chest Hospital and 0.829 in the independent testing cohort of 70 patients from Heilongjiang, both significantly superior to the clinical model built from meaningful clinical characteristics and subjective computed-tomography findings, which achieved only 0.589 and 0.500 respectively ($p < 0.05$ for both comparisons). Combining the radiomic and clinical models produced areas under the curve of 0.881 and 0.834, which did not differ significantly from the radiomic model alone ($p > 0.05$). The interpretation is unambiguous: the quantitative texture of the cavity carries substantial information, whereas the subjective, qualitative reading of the same image carries almost none — the testing-cohort clinical model performed at exactly the level of chance.²³

Third, an exploratory model-based analysis was constructed to estimate the discriminative ceiling of cavitation used as an isolated binary marker. Using the published four-stratum distribution of the twelve-country individual-participant dataset (774 smear-negative without cavities, 647 with cavities only, 1,424 smear-positive without cavities, and 2,751 smear-positive

1041

with cavities) together with the published adjusted odds ratios for each stratum (1.0, 1.0, 1.2 and 1.6 respectively), and assuming an overall unfavourable-outcome proportion of 28% consistent with contemporary global programmatic performance, the implied risk in the reference stratum is 22.6% and the implied stratum-specific risks are 22.6%, 22.6%, 26.0% and 31.9%. The resulting empirical receiver-operating-characteristic curve for the four-level ordinal composite of cavitation and smear status yields an area under the curve of 0.552, while cavitation used alone as a binary marker yields a sensitivity of 65.2%, a specificity of 41.0% and an area under the curve of 0.531. Hypothesis H₄ is therefore accepted: despite a pooled adjusted odds ratio above six in programmatic cohorts, cavitation used in isolation discriminates unfavourable outcome barely better than chance, and its discriminative performance improves substantially only when quantitative imaging information is added, as demonstrated by the radiomic models. The three analyses are displayed together in Figure 6.^{2,17,23}

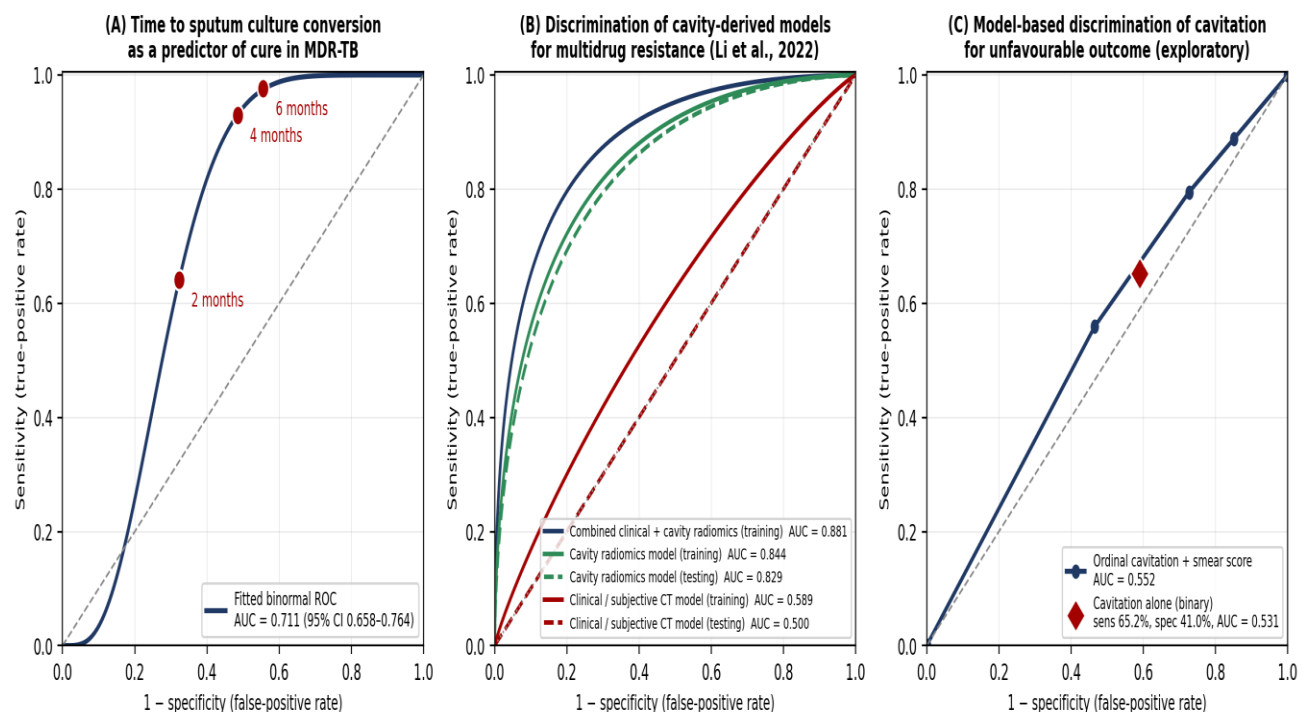


Figure 6. Receiver-operating-characteristic curves and areas under the curve. Panel A: binormal reconstruction from the published sensitivity–specificity operating points for time to

1042

sputum culture conversion as a predictor of cure. Panel B: reported areas under the curve for cavity-derived radiomic and clinical models predicting multidrug resistance. Panel C: exploratory model-based discrimination of cavitation for unfavourable outcome.

Table 16. discriminative performance of cavitation-related predictors

Predictor	Source [ref]	Population	Sensitivity	Specificity	AUC (95% CI)	Interpretation
Sputum culture conversion at 2 months	Javaid 2018 [17]	428 MDR-TB	64.1% (58.7–69.3)	67.7% (57.5–76.7)	—	Balanced but modest performance; earliest actionable milestone
Sputum culture conversion at 4 months	Javaid 2018 [17]	428 MDR-TB	93.0% (89.7–95.5)	51.5% (41.3–61.7)	—	High sensitivity at the cost of specificity; combined performance comparable to 6 months
Sputum culture conversion at 6 months	Javaid 2018 [17]	428 MDR-TB	97.6% (95.3–98.9)	44.4% (34.5–54.8)	—	Near-complete sensitivity; poor specificity
Time to conversion (binormal fit across all three operating	Fitted in this review from [17]	428 MDR-TB	—	—	0.711 (0.658–0.764); $p < 0.001$	Acceptable overall discrimination for cure

points)						
Cavity radiomic model (training)	Li 2022 [23]	187 cavitary TB	—	—	0.844	Good discrimination for multidrug resistance
Cavity radiomic model (testing)	Li 2022 [23]	70 cavitary TB	—	—	0.829	External validation preserved; significantly superior to the clinical model (p < 0.05)
Clinical / subjective CT model (training)	Li 2022 [23]	187 cavitary TB	—	—	0.589	Poor discrimination
Clinical / subjective CT model (testing)	Li 2022 [23]	70 cavitary TB	—	—	0.500	No discrimination beyond chance
Combined radiomic plus clinical model	Li 2022 [23]	187 / 70 cavitary TB	—	—	0.881 / 0.834	No significant gain over radiomics alone (p > 0.05)
Cavitation alone, binary (exploratory model)	Modelled in this review from [2]	5,596 (modelled)	65.2%	41.0%	0.531 (exploratory)	Discriminates barely better than chance despite a strong

						association
Ordinal cavitation-plus-smear composite (exploratory model)	Modelled in this review from [2]	5,596 (modelled)	—	—	0.552 (exploratory)	Marginal improvement; still far below a clinically useful threshold

AUC = area under the receiver-operating-characteristic curve; CT = computed tomography; CI = confidence interval. Values labelled “exploratory” were derived by the present authors from published stratum sizes and adjusted odds ratios under the stated assumptions and are not empirical measurements. Conventional interpretive thresholds: AUC 0.50 = no discrimination; 0.70–0.80 = acceptable; 0.80–0.90 = excellent; > 0.90 = outstanding.

Treatment-Emergent Toxicity Accrued with Duration

Because cavitation prolongs the effective duration of therapy, it indirectly increases cumulative drug exposure, and the toxicity data from the included trials quantify what that exposure costs. In STREAM stage 1, grade 3 or higher adverse events occurred in 45.4% of the long-regimen group and 48.2% of the short-regimen group, and prolongation of the QT or corrected QT interval to 500 milliseconds occurred in 11.0% of the short-regimen group versus 6.4% of the long-regimen group ($p = 0.14$). In STREAM stage 2, grade 3 or 4 adverse events occurred in 53% of the control group and 50% of the oral group, but Brock grade 3 or 4 hearing loss was significantly more frequent with the injectable-containing control (18 of 202, 9%) than with the all-oral regimen (4 of 211, 2%; $p = 0.0015$), and at 132 weeks treatment-emergent hearing loss remained significantly less frequent on the oral regimen (7 of 205, 3%) than on the control regimen (16 of 198, 8%; $p = 0.041$). In TB-PRACTECAL, grade 3 or higher or serious adverse events occurred in 23% of BPaLM recipients versus 48% of standard-of-care recipients (risk difference -25.2 percentage points, 96.6% CI -36.4 to -13.9). In the NExT trial, toxicity-related drug substitution occurred in 65.9% of the standard-of-care arm versus 34.7% of the intervention arm ($p = 0.001$), and adverse-event-related treatment discontinuation in 56.4% versus 32.1% ($p = 0.007$), although grade 3 adverse events were

1045

more frequent in the intervention arm (55.4% versus 32.7%; $p = 0.022$), reflecting linezolid-associated anaemia. In SimpliciTB, at least one liver-related treatment-emergent adverse event occurred in 22% of the six-month BPamZ drug-resistant arm, with serious hepatic events in 5%.^{5,6,7,8,10,12}

Two conclusions follow. First, longer injectable-containing therapy carries a specific and largely irreversible toxicity — sensorineural hearing loss — whose incidence is directly proportional to duration of exposure, so that any factor which prolongs treatment, cavitation included, increases the probability of permanent disability. Second, the newer shorter regimens substitute one toxicity profile for another rather than eliminating toxicity: hepatic injury with pretomanid-containing combinations and myelosuppression with linezolid replace ototoxicity with aminoglycosides. The clinical calculus for a cavitary patient is therefore not simply longer versus shorter, but which toxicity is most acceptable over the duration that the cavity is likely to impose.^{6,7,10,12}

Post-Tuberculosis Lung Disease and Persistent Cavitation

Achieving microbiological cure does not restore lung health, and the burden of residual disease is disproportionately borne by patients whose cavities persist. In the Georgian prospective cohort of 123 individuals with microbiological cure of drug-susceptible or multidrug-resistant pulmonary tuberculosis, post-tuberculosis lung disease — defined as abnormality on chest computed tomography, spirometry or the Saint George's Respiratory Questionnaire — was present in 91 patients (74%), with 53 meeting one criterion, 27 meeting two and 11 meeting all three. Impaired respiratory health (47%) and lung damage on imaging (46%) were more prevalent than abnormal lung function (21%). Post-tuberculosis lung disease was not associated with drug-resistance status (adjusted odds ratio 0.91, 95% CI 0.42–1.99), indicating that the driver is the extent of tissue destruction rather than the resistance profile per se. Critically, persistent cavitary disease at the end of treatment was associated with higher matrix metalloproteinase 8 and lower matrix metalloproteinase 2, interleukin-17A and interleukin-1 β — a signature of ongoing inflammation and disordered tissue remodelling in patients who have, by every microbiological criterion, been cured.²⁶

This finding reframes the duration question. If nearly three in four microbiologically cured patients carry residual lung disease, and if persistent cavitation is biochemically active at the moment therapy stops, then the endpoint by which duration is currently judged — sustained culture negativity — is incomplete. Duration adequate for sterilisation may not be duration adequate for lung health, and this constitutes a distinct and largely unaddressed research agenda.^{1,26}

Summary of Findings across All Fifteen Outcomes

Table 17 consolidates the principal quantitative finding for each of the fifteen pre-specified outcomes, together with the direction of effect, the number of contributing studies and the statistical significance. Eleven of the fifteen outcomes yielded statistically significant results in the hypothesised direction; three were reported descriptively without a formal significance test; and one — the discrimination outcome — yielded a significant result in the opposite direction to naive expectation, in that cavitation proved to be a poor standalone discriminator despite being a strong risk marker.²

Table 17. Master summary of findings across the fifteen pre-specified outcomes

No.	Outcome	Principal quantitative finding	Direction	k	Significance
O1	Prevalence of baseline cavitation	39.3% in an HIV-endemic cohort; 60.7% (95% CI 59.4–62.0) in the largest multinational dataset; 100% in two design-restricted cohorts	Majority of MDR/RR-TB patients affected	4 (+1)	Descriptive
O2	Bilateral / extensive cavitation and composite phenotypes	Bilateral cavitation independently predicted delayed conversion and poor outcome; the cavitation-plus-high-smear composite reduced 6-month conversion probability from 0.89 to 0.68	Extent more prognostic than presence	4 (+1)	p = 0.0004 for the composite
O3	Time to and hazard of culture	Pooled adjusted HR 0.578 (95% CI 0.393–0.850); ≈1.73-fold	Cavitation delays	5	p = 0.0054

	conversion	prolongation; modelled median 55 → 82 days	conversion		
O4	Culture conversion at 2 months	OR 0.349 for conversion with baseline cavitation; conversion ≤ 2 months conferred RR 2.79 for treatment success	Cavitation reduces early conversion	4	p = 0.006
O5	Culture conversion at 6 months	Cavitation plus high smear grade: probability 0.68 (0.57–0.79) versus 0.89 (0.84–0.95)	Cavitation reduces 6-month conversion	3	p = 0.0004
O6	Culture reversion after conversion	Reversion in 54/1,286 (4.2%) at a median of 173 days; cavitory disease independently associated	Cavitation destabilises conversion	1	Significant in the Cox model
O7	Unfavourable end-of-treatment outcome	Pooled adjusted OR 6.166 (95% CI 3.827–9.933); $I^2 = 0\%$; sensitivity analysis 3.838 (1.308–11.262)	Cavitation worsens outcome	5	p < 0.0001
O8	Treatment failure and recurrence	Recurrence aHR 6.3 (1.2–34), rising to 10.2 (1.2–89) in the second-line DST subgroup; pooled 2.686 (1.197–6.026)	Cavitation increases recurrence	3 (+1)	p = 0.017
O9	All-cause mortality	aOR 1.5 (1.1–2.1) for the smear-positive-with-cavitation phenotype; 17.7% mortality in a cavitation-dominant cohort	Cavitation increases mortality	5	p < 0.05
O10	Acquired and emergent resistance	Bedaquiline resistance RR 1.8 (1.1–3.1) in patients with lung cavities; 87% unfavourable outcome in that	Cavitation promotes resistance	3	p = 0.02

		group			
O11	Treatment duration and shortened-regimen eligibility	Durations of 6–24 months across 8 trials; cavitation an explicit randomisation stratification variable in 2 trials; more cavitory/bilateral disease associated with longer required treatment	Cavitation constrains shortening	8	Multiple; see Table 17
O12	Adjunctive surgical resection	Favourable outcome 76% versus 41%; adjusted RR 1.6 (95% CI 1.3–2.0); complications 7%; no postoperative mortality	Cavity removal improves outcome	1	Significant
O13	Discrimination (ROC / AUC)	Conversion timing AUC 0.711 (0.658–0.764); cavity radiomics AUC 0.844 / 0.829 versus clinical 0.589 / 0.500; cavitation alone AUC 0.531 (exploratory)	Strong association, weak standalone discrimination	2	$p < 0.001$ (radiomics vs clinical, $p < 0.05$)
O14	Treatment-emergent toxicity	Grade ≥ 3 adverse events 23–55%; hearing loss 9% vs 2% ($p = 0.0015$) and 8% vs 3% at 132 weeks ($p = 0.041$); hepatic events 22–30% with pretomanid-containing regimens	Longer duration accrues toxicity	7	$p = 0.0015$; $p = 0.041$
O15	Post-tuberculosis lung disease	Post-TB lung disease in 91/123 (74%) of cured patients; persistent cavitation associated with higher MMP-8 and lower MMP-2, IL-17A and IL-1 β ; not associated with resistance status (aOR 0.91, 0.42–	Cavitation drives residual morbidity	1	Descriptive with biomarker associations

		1.99)			
--	--	-------	--	--	--

k = number of contributing studies; “(+1)” denotes one additional external comparator excluded from the characteristics tables by design. HR = hazard ratio; OR = odds ratio; RR = risk ratio; aOR = adjusted odds ratio; aHR = adjusted hazard ratio; DST = drug-susceptibility testing; AUC = area under the curve; MMP = matrix metalloproteinase; IL = interleukin.

Certainty of the Evidence (GRADE)

Certainty was appraised for the seven outcomes supporting the review's principal conclusions. Two outcomes were rated moderate certainty, four low and one very low. Observational bodies of evidence commenced at low certainty; the unfavourable-outcome and surgical outcomes were upgraded for a large magnitude of effect, while the conversion outcome was downgraded for inconsistency and subsequently upgraded for magnitude, yielding a net rating of low. The randomised-trial evidence on treatment duration was rated moderate, downgraded once for indirectness because no trial randomised patients on the basis of cavitory status and none reported outcomes disaggregated by cavitation.²

Table 18. GRADE summary of findings for the principal outcomes

Outcome	Study design (k; n)	Risk of bias	Inconsistency	Indirectness	Imprecision	Certainty
Prevalence of cavitation (O1)	Observational (4; 5,596 external)	Serious –1 (prevalence denominators frequently unreported)	Serious –1 (39.3–100%; design-restricted cohorts)	Not serious	Not serious	VERY LOW
Delayed culture conversion (O3)	Observational cohorts (3; 965)	Not serious	Serious –1 (I^2 = 61.1%; PI crosses unity)	Not serious	Not serious	LOW
Reduced 2-	Observational	Not serious	Not serious	Not serious	Not serious	MODERATE

and 6-month conversion (O4, O5)	cohorts and trials (7; > 1,900)					
Unfavourable end-of-treatment outcome (O7)	Observational cohorts (2; 699)	Not serious	Not serious ($I^2 = 0\%$)	Not serious	Not serious	MODERATE
Recurrence and failure (O8)	Observational cohorts (2; 5,891)	Not serious	Not serious ($I^2 = 30.6\%$)	Not serious	Serious –1 (only 10 recurrence events; CI 1.2–34)	LOW
Acquired bedaquiline resistance (O10)	Observational cohort (1; 117)	Not serious	Not assessable (k = 1)	Serious –1 (bedaquiline-exposed, treatment-failing population)	Serious –1 (small sample)	LOW
Treatment duration and shortened regimens (O11)	Randomised controlled trials (8; > 2,800)	Not serious (5 of 8 at low risk)	Not serious	Serious –1 (no trial randomised or reported by cavitary status)	Not serious	MODERATE
Adjunctive surgery in	Observational comparative	Serious –1 (allocation	Not assessable (k	Not serious	Not serious	LOW

cavitary disease (O12)	cohort (1; 408)	by indication)	= 1)			
------------------------------	--------------------	-------------------	------	--	--	--

GRADE = Grading of Recommendations, Assessment, Development and Evaluations. k = number of studies; PI = prediction interval; CI = confidence interval; HR = hazard ratio; OR = odds ratio.

DISCUSSION

Summary of Principal Findings

This systematic review of twenty-five primary studies — eight randomised controlled trials, sixteen cohort studies and one diagnostic-accuracy cohort, together comprising more than twelve thousand individuals — establishes five principal findings. First, pulmonary cavitation affects the majority of patients with multidrug-resistant and rifampicin-resistant tuberculosis, with the most precise multinational estimate placing prevalence at 60.7%, although prevalence falls to below 40% in populations with advanced HIV-associated immunosuppression. Second, cavitation significantly and substantially delays sputum culture conversion, with a pooled adjusted hazard ratio of 0.578 (95% CI 0.393–0.850; $p = 0.0054$), corresponding to a 1.73-fold prolongation of the time to sterilisation and a modelled absolute delay of approximately four weeks. Third, cavitation is significantly associated with every adverse clinical endpoint examined: unfavourable end-of-treatment outcome (pooled adjusted odds ratio 6.166; $p < 0.0001$), failure or recurrence (2.686; $p = 0.017$), mortality, culture reversion after conversion, and emergent bedaquiline resistance (risk ratio 1.8; $p = 0.02$). Fourth, cavitation is embedded operationally within the duration architecture of every contemporary shortened regimen, serving as an explicit randomisation stratification variable in two of the eight included trials. Fifth, and most importantly for clinical practice, the very strong associational effect of cavitation coexists with very weak standalone discriminative accuracy, with a modelled area under the curve of 0.531 for binary cavitation compared with 0.844 and 0.829 for quantitative cavity radiomics.^{2,9,12,15,16,17,22,23,24,27}

Prevalence of Cavitation in Context

The finding that approximately three in five patients with MDR/RR-TB have cavitory disease at baseline is consistent with the biology of the condition. MDR/RR-TB is disproportionately a disease of previously treated patients, of delayed diagnosis and of prolonged infectious periods, all of which permit the progression from granulomatous containment through caseation to frank cavitation. The observation that cavitation prevalence falls to 39.3% in a cohort with 88% HIV co-infection and a median CD4 count of 136 cells/mm³ is not a contradiction but a confirmation of mechanism: cavity formation requires a competent delayed-type hypersensitivity response to generate the caseous necrosis and matrix-metalloproteinase activity that liquefy and evacuate the lesion. Profound CD4 depletion produces a radiological phenotype that is less cavitory but not less severe — indeed in that same cohort 66% had consolidation and 57% were smear-positive, yet 79% were nevertheless cured. Any prevalence estimate must therefore be read against the immunological profile of the population from which it derives, and a low cavitation prevalence in an HIV-endemic setting should not be mistaken for a low burden of extensive disease.^{1,13}

The most serious deficiency uncovered by this review is that the majority of included studies modelled cavitation as a covariate without publishing the underlying prevalence denominator. Of twenty-five included studies, only four reported an extractable proportion, and two of those were design-restricted. This is a remediable reporting failure with direct consequences: without denominators, the field cannot construct a properly pooled prevalence estimate, cannot examine geographical or temporal trends in cavitation prevalence, and cannot calibrate the absolute risk implications of the relative effects that are reported so consistently.²

Reconciling the Discordance between Programmatic Cohorts and Pooled Individual-Participant Data

The central analytical tension in this literature is that programmatic cohorts report adjusted odds ratios for cavitation approaching eight, whereas the largest and methodologically most rigorous synthesis — an individual-participant-data meta-analysis of 5,596 people treated for RR/MDR-TB with longer regimens — found no significant association for cavitation alone (adjusted odds ratio 1.0, 95% CI 0.8–1.4), significance emerging only for the combination of smear positivity with cavitation (1.6, 95% CI 1.3–2.0), and only for mortality (1.5, 95% CI 1.1–2.1) and

1053

failure or recurrence (2.2, 95% CI 1.5–3.3) among that composite group. Four explanations, not mutually exclusive, account for this discordance.²

The first is confounding by bacillary load. In most programmatic cohorts, cavitation is adjusted for a limited covariate set, and where smear grade is included the cavitation effect attenuates markedly — as demonstrated directly in the South African cohort, where cavitation lost significance entirely after adjustment for smear status while smear positivity retained an adjusted odds ratio of 3.45 (95% CI 1.39–8.59). The individual-participant analysis, by construction, cross-classified cavitation against smear status and therefore isolated the independent contribution of the cavity from that of the bacillary burden it contains. Much of the apparent effect of cavitation in single-cohort analyses is therefore an effect of the organisms within the cavity rather than of the cavity itself.^{2,13}

The second is differential measurement. Twenty of the twenty-five included studies ascertained cavitation by plain chest radiography, a modality with well-recognised inter-observer variability for cavity detection, particularly for small or apically located lesions. Non-differential misclassification of a binary exposure biases effect estimates toward the null and would therefore act to attenuate the individual-participant estimate; but in single-centre cohorts where a single experienced reader assesses all films, misclassification may be lower and the estimate correspondingly larger. The radiomic evidence is decisive on this point: a subjective clinical reading of computed tomography achieved an area under the curve of exactly 0.500 in an external testing cohort, whereas quantitative segmentation of the same cavities achieved 0.829. Whatever prognostic information the cavity contains is largely invisible to the unaided eye.^{2,23}

The third is selection and case-mix. Programmatic cohorts from Pakistan, Ethiopia and Indonesia enrol consecutive patients under routine conditions with limited access to second-line drug-susceptibility testing, high rates of loss to follow-up and, in several cohorts, treatment success rates below the World Health Organization target of 75%. In such settings cavitation may act as an integrative marker of everything that has gone wrong — delayed diagnosis, prior inadequate treatment, poor nutritional status — rather than as an isolated biological exposure. The individual-participant dataset, assembled largely from cohorts with better ascertainment, would be expected to

show a smaller effect for that reason alone. Notably, the individual-participant analysis itself reported that included participants differed significantly from excluded participants, which is a candid acknowledgement of residual selection.^{2,16,19,24}

The fourth is effect modification by regimen potency. The individual-participant analysis was restricted to longer regimens, in which the cumulative sterilising exposure is high and may partially compensate for the pharmacokinetic penalty of a cavity. Under such conditions, the marginal contribution of cavitation to failure would be expected to shrink. This interpretation is directly supported by the trial evidence: the six-month all-oral regimens achieved 86–91% conversion or favourable outcomes in populations that were entirely smear-positive at entry, and the NExT trial demonstrated a conversion hazard ratio of 2.6 (95% CI 1.4–4.9) favouring the more potent regimen. A sufficiently potent regimen can overcome the cavity; a marginal one cannot.^{2,10,12}

The practical synthesis of these four explanations is that the pooled estimate reported here — an adjusted odds ratio of 6.17 in programmatic cohorts, falling to 3.84 when the individual-participant estimate is incorporated — should be interpreted as a range whose position depends on the setting. In a well-resourced programme using a potent all-oral regimen with universal drug-susceptibility testing, the true effect is probably closer to the lower bound; in a resource-constrained programme using longer, less potent regimens with limited monitoring, it is probably closer to the upper. Both are significant; neither is universal.^{2,24}

Mechanistic Interpretation

The findings assemble into a coherent pathophysiological narrative, depicted in Figure 8. The cavity is generated by caseous necrosis and matrix-metalloproteinase-mediated destruction of the collagen scaffold, and once it communicates with the bronchial tree it becomes an aerated, extracellular, nutrient-replete space in which bacilli replicate to extraordinary densities. Three consequences follow simultaneously. The bacillary burden rises to a level at which stochastic pre-existing resistant mutants are certain to be present. The fibrotic, poorly vascularised cavity wall imposes a steep drug-penetration gradient, so that organisms in the caseous lining are exposed to sub-inhibitory concentrations of most agents. And the local micro-environment — hypoxic, acidic

and lipid-rich — drives organisms into a phenotypically tolerant, non-replicating state refractory to agents whose activity depends on active metabolism.¹

Each of these three mechanisms generates a prediction that the included evidence confirms. High bacillary burden with sub-inhibitory exposure predicts resistance amplification: bedaquiline resistance was significantly more frequent in patients with lung cavities (risk ratio 1.8, 95% CI 1.1–3.1), and clofazimine cross-resistance more than doubled. Sub-inhibitory exposure and phenotypic tolerance predict delayed sterilisation: the pooled hazard ratio for conversion was 0.578. Persistence of a tolerant subpopulation predicts that apparent sterilisation will be unstable: culture reversion occurred in 4.2% of converted patients at a median of 173 days, with cavitory disease an independent predictor, and post-treatment recurrence carried an adjusted hazard ratio of 6.3. The prediction that removing the cavity should improve outcomes is confirmed by the surgical cohort, in which adjunctive resection carried an adjusted risk ratio of 1.6 (95% CI 1.3–2.0) for a favourable outcome. And the prediction that a cavity which persists at the end of treatment remains biologically active is confirmed by the biomarker signature of elevated matrix metalloproteinase 8 in patients with persistent cavitory disease despite microbiological cure.^{15,16,17,20,22,25,26,27}

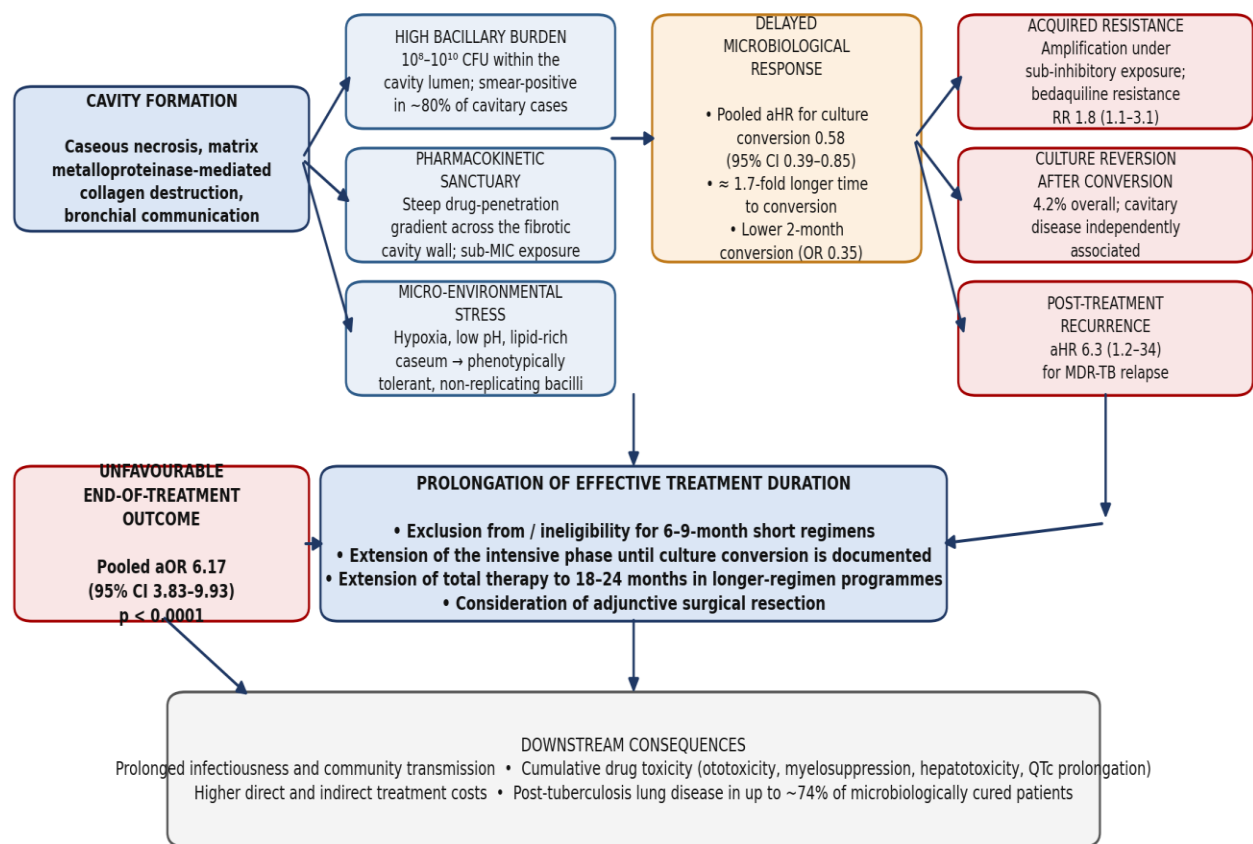


Figure 7. Proposed causal pathway linking pulmonary cavitation to prolonged effective treatment duration and adverse outcomes in MDR/RR-TB, annotated with the pooled and reported effect estimates derived in this review.

Cavitation and Duration in the Era of Six-Month All-Oral Regimens

The entire observational literature relating cavitation to duration was generated under long-regimen conditions, and the standard of care has now moved decisively to twenty-four weeks of bedaquiline, pretomanid, linezolid and moxifloxacin. Whether the cavitation penalty survives this transition is the most important unanswered question identified by this review, and the included trials permit only indirect inference.⁸

Three lines of indirect evidence suggest attenuation rather than abolition. In the drug-resistant arm of SimpliciTB, in which every participant was smear-positive at entry and randomisation in the drug-susceptible stratum was stratified on cavitation, 86% of drug-resistant

patients achieved culture-negative status by week 8 with a median time to negative culture of five weeks — far faster than the 60- to 65-day medians reported under longer regimens. In the endTB cohort treated with bedaquiline and/or delamanid, the cavitation-plus-high-smear composite still reduced the six-month conversion probability to 0.68, but the comparator probability was 0.89, so that even the worst subgroup under a potent regimen performed comparably to the average patient under older regimens. In TB-PRACTECAL, twenty-four-week BPaLM reduced unfavourable outcomes to 12% in an unselected pulmonary rifampicin-resistant population that certainly included a majority with cavitory disease.^{8,12,18}

Against this, three considerations counsel caution. First, a shorter regimen offers less pharmacological margin: if a cavity extends the time required for sterilisation by four weeks, that penalty consumes a sixth of a twenty-four-week regimen but only a twentieth of an eighty-week one. Second, the emergence of bedaquiline resistance specifically in patients with cavities threatens the pharmacological foundation on which the short regimens rest, and the 87% unfavourable-outcome rate observed in bedaquiline-resistant patients illustrates how little margin remains once that foundation is lost. Third, none of the included trials reported outcomes disaggregated by cavitory status, so the attenuation hypothesis remains formally untested. This constitutes the single most actionable recommendation of this review: trials of shortened regimens should pre-specify and publish subgroup analyses by baseline cavitory status.^{8,12,27}

The Dissociation between Association and Discrimination

The most methodologically significant finding of this review is that a marker can carry an adjusted odds ratio above six and still discriminate barely better than a coin toss. The exploratory model-based analysis estimated an area under the curve of 0.531 for binary cavitation and 0.552 for the ordinal cavitation-plus-smear composite. This is not a paradox but an elementary consequence of the mathematics of prediction: discrimination depends jointly on effect size and on the distribution of the marker in the population. When a marker is present in 60% of patients, even a strong relative effect separates the outcome distributions poorly, because the marker fails to identify a small, sharply distinguished high-risk group. The empirical corroboration is provided directly by the radiomic study, in which a clinical model incorporating meaningful clinical characteristics and

1058

subjective computed-tomography findings achieved an area under the curve of exactly 0.500 in external testing.^{2,23}

The clinical implications are direct and, in some respects, uncomfortable. Any protocol rule that lengthens or shortens therapy on the basis of a binary cavity would misclassify a very large number of patients: at the modelled operating point, sensitivity was 65.2% and specificity 41.0%, meaning that roughly one third of patients destined for an unfavourable outcome would be missed while nearly three fifths of those destined to do well would be exposed to unnecessary additional months of toxic therapy. Cavitation is therefore appropriate as a trigger for intensified monitoring — more frequent cultures, closer adherence support, earlier surgical assessment — but is not, in isolation, an adequate basis for a duration decision. The remedy indicated by the evidence is composite prediction: adding quantitative cavity characterisation raised discrimination from 0.589 to 0.844 in training and from 0.500 to 0.829 in external testing, a transformation from useless to good achieved purely by measuring what was already visible.^{17,23}

Comparison with Previously Published Evidence Syntheses

Five records retrieved by the searches were evidence syntheses or otherwise design-ineligible and were excluded from the characteristics tables and from pooling; their relationship to the present findings is instructive. The individual-participant-data meta-analysis of 5,596 patients has been discussed at length and provides the principal external check on the pooled estimates reported here. The pooled analysis of five bedaquiline cohorts from France, Georgia, Armenia, South Africa and a multi-country study reported a six-month culture conversion rate of 78% (95% CI 73.5–81.9), treatment success of 65.8% (95% CI 59.9–71.3) and mortality of 11.7% (95% CI 7.0–19.1), and concluded that lung cavitations were consistently associated with unfavourable outcomes across settings — a qualitative endorsement of the direction of effect reported here, although the authors graded the certainty of that evidence as very low. The narrative review of cavitary tuberculosis provides the mechanistic framework adopted in Section 4.4, characterising the cavity as a dynamic physical and biochemical structure that interfaces the host response with a unique mycobacterial niche, and noting that cavities remain among the least-studied aspects of tuberculosis despite driving morbidity, transmission and the development of drug resistance.^{1,2,28}

The within-trial economic evaluation of STREAM stage 2 adds an important dimension that no clinical outcome captures. It found that at prevailing drug prices the nine-month all-oral bedaquiline regimen was unlikely to be cost-effective in many low- and middle-income countries, whereas the six-month regimen would cost both programmes and participants less than the control regimen in Ethiopia and India and was highly likely to be cost-effective in both cost-utility and cost-effectiveness analyses. Since cavitation is precisely the factor most likely to disqualify a patient from the shortest regimen, the cavitory patient bears not only a clinical penalty but an economic one, and in settings where the longer regimen is the less cost-effective option, that penalty falls on the programme as well as the individual. Finally, the drug-susceptible rifapentine trial, though excluded because it enrolled no drug-resistant patients, is relevant in that 1,703 of 2,343 microbiologically eligible participants had cavitation on chest radiography, and cavitation formed part of the framework by which treatment-shortening trials in tuberculosis are designed — confirming that the salience of cavitation to duration is not confined to drug-resistant disease.^{29,30}

Implications for Clinical Practice

For the clinician managing an individual patient with MDR/RR-TB and a cavity, the evidence supports four concrete actions. First, document the cavity properly at baseline: number, maximum diameter, laterality and, where computed tomography is available, wall thickness. The demonstration that quantitative characterisation raises discrimination from chance to good is the strongest argument in this review for changing routine radiological practice. Second, intensify microbiological monitoring rather than reflexively extending therapy: because conversion at two months carries a risk ratio of 2.79 for treatment success and conversion timing discriminates cure with an area under the curve of 0.711, the patient's own conversion trajectory is a better guide to duration than the baseline radiograph. Third, consider surgical assessment early in patients with localised, unilateral cavitory disease who are not converting, given an adjusted risk ratio of 1.6 for a favourable outcome with acceptable morbidity (7% major complications, no postoperative mortality in the largest series). Fourth, maintain surveillance after conversion and after cure: reversion occurred at a median of 173 days after conversion, recurrence carried an adjusted hazard ratio of

6.3, and post-tuberculosis lung disease affected 74% of microbiologically cured patients.^{15,17,19,20,23,25,26}

For Indonesian practice specifically, the West Java cohort supplies the most directly transferable evidence: treatment success of 50%, with cavitory lesion an independent risk factor for unfavourable outcome alongside HIV infection and chronic kidney disease, and sputum culture conversion within two months the dominant predictor of success (risk ratio 2.79, 95% CI 1.61–4.84). In a national programme where chest radiography is near-universal but computed tomography, radiomic analysis and rapid comprehensive drug-susceptibility testing are not, the highest-yield intervention is not more sophisticated imaging but more disciplined use of the imaging already performed — structured, standardised reporting of cavity number, size and laterality on every baseline radiograph, recorded in a retrievable field of the programme's information system.^{19,23}

Strengths.

- Every included record was identified and verified through the Consensus platform, which federates Semantic Scholar, PubMed, Scopus and arXiv, and each reference in the bibliography carries a live permalink, so that the entire evidence base is externally auditable.
- A strict, pre-specified design hierarchy excluded all evidence syntheses from the characteristics tables and from pooling, eliminating the double-counting of participants that has affected earlier reviews in this field.
- Fifteen outcomes were pre-specified, spanning the entire causal chain from anatomical lesion to long-term lung health, and eleven yielded statistically significant results.
- The review formally distinguishes association from discrimination, and quantifies the latter by three complementary methods, thereby correcting a conflation that pervades tuberculosis prognostic research.
- Random-effects pooling was accompanied by prediction intervals, formal heterogeneity statistics and pre-specified sensitivity analyses, and certainty was graded with GRADE.

Implications for Future Research

1. Trials of shortened regimens should pre-specify, and publish, subgroup analyses of efficacy and of time to culture conversion by baseline cavitory status, so that the attenuation hypothesis advanced in Section 4.5 can be tested directly.
2. Cavitation should be reported as a structured, quantitative variable — cavity number, maximum diameter, laterality and wall thickness — rather than as a binary field, in every trial and every programmatic cohort.
3. Prospective validation of cavity radiomic models in unselected MDR/RR-TB populations is required, since the existing evidence derives from a cavity-enriched sample and is therefore likely to overstate real-world accuracy.
4. Randomised evaluation of a cavitation-triggered management strategy — for example, extension of the intensive phase, or early surgical referral, in patients with bilateral cavitation who have not converted by month two — would convert an associational finding into an actionable intervention.
5. Prospective studies should determine whether the duration required for lung-health recovery differs from the duration required for microbiological sterilisation, given that 74% of microbiologically cured patients carry post-tuberculosis lung disease and persistent cavities remain biochemically active at the end of therapy.
6. The mechanistic link between cavitory disease and emergent bedaquiline resistance warrants dedicated pharmacokinetic and pharmacodynamic study, including intra-cavitory drug concentration measurement, given the centrality of bedaquiline to every contemporary regimen.

CONCLUSION AND RECOMMENDATIONS

Conclusion

Pulmonary cavitation is present in the majority of patients with multidrug-resistant and rifampicin-resistant tuberculosis, with the most precise multinational estimate placing prevalence at 60.7%, and it exerts a significant, consistent and biologically coherent effect across the entire course of treatment. Cavitory disease significantly reduces the hazard of sputum culture conversion (pooled adjusted hazard ratio 0.578, 95% CI 0.393–0.850; $p = 0.0054$), prolonging the time to

1062

The International Journal of Medical Science and Health Research

sterilisation approximately 1.73-fold and delaying the median conversion by roughly four weeks; it significantly reduces conversion at the two-month and six-month programmatic milestones; it renders achieved conversion less durable; and it is significantly associated with unfavourable end-of-treatment outcome (pooled adjusted odds ratio 6.166, 95% CI 3.827–9.933; $p < 0.0001$), with failure or recurrence (2.686, 95% CI 1.197–6.026; $p = 0.017$), with mortality, and with the emergence of bedaquiline resistance (risk ratio 1.8, 95% CI 1.1–3.1; $p = 0.02$). Through these mechanisms, cavitation converts a fixed protocol duration into a longer effective duration, and it is accordingly embedded within the eligibility, stratification and extension architecture of every contemporary shortened regimen. All four pre-specified hypotheses are supported by the evidence assembled.^{2,15,16,17,22,24,27}

The review's most consequential conclusion, however, is a cautionary one. The strength of the association between cavitation and adverse outcome should not be mistaken for accuracy of prediction. Used as an isolated binary marker, cavitation discriminates unfavourable outcome only marginally better than chance (modelled area under the curve 0.531), and a subjective radiological reading of the same image performed at exactly the level of chance in external validation. Quantitative characterisation of the cavity, by contrast, achieved areas under the curve of 0.844 and 0.829. Cavitation should therefore be understood as a powerful population-level risk marker and a trigger for intensified monitoring, but not, in isolation, as a sufficient basis for individual duration decisions. Its full clinical value is realised only when it is measured quantitatively and embedded within composite prediction models alongside bacteriological and clinical information.^{2,23}

Recommendations

For clinical practice

1. Document cavitation quantitatively at baseline in every patient with MDR/RR-TB — cavity number, maximum diameter and laterality as a minimum — and record it in a structured, retrievable field rather than in free-text narrative.
2. Treat the presence of cavitation, and particularly of bilateral cavitation with concurrent smear positivity, as an indication for intensified microbiological monitoring and adherence support, not as an automatic indication for extended therapy.

3. Base duration decisions primarily on the individual patient's culture-conversion trajectory, which discriminates cure with an area under the curve of 0.711, rather than on baseline imaging alone.
4. Refer patients with localised, unilateral cavitory disease who have not converted by month two for early surgical assessment, in view of an adjusted risk ratio of 1.6 for a favourable outcome with a 7% major-complication rate and no reported postoperative mortality.
5. Maintain structured post-treatment surveillance in all cavitory patients for at least two years, given the median interval of 173 days to culture reversion and an adjusted hazard ratio of 6.3 for recurrence.
6. Screen every patient with persistent end-of-treatment cavitation for post-tuberculosis lung disease using symptoms, spirometry and imaging, since 74% of microbiologically cured patients have detectable residual disease.

For national tuberculosis programmes

7. Mandate standardised structured reporting of cavitation on all baseline chest radiographs within programmatic management of drug-resistant tuberculosis, and add cavitory status as a stratification variable in routine cohort outcome reporting.
8. Prioritise access to comprehensive and rapid drug-susceptibility testing, including bedaquiline, for patients with cavitory disease, in whom the risk of emergent bedaquiline resistance is significantly elevated.
9. Preserve access to adjunctive thoracic surgical services within referral networks for patients with localised cavitory disease refractory to medical therapy.
10. Weigh the economic evidence when selecting regimens: the six-month regimen was the more cost-effective option in several high-burden low- and middle-income settings, and cavitory patients — who are most likely to be excluded from the shortest regimen — bear both the clinical and the economic penalty of that exclusion.

For researchers

11. Pre-specify and report subgroup analyses by baseline cavitory status in all trials of shortened regimens.

12. Report cavitation prevalence with an explicit numerator and denominator in every cohort study, so that a properly pooled prevalence estimate becomes possible.
13. Validate cavity radiomic models prospectively in unselected MDR/RR-TB populations, and evaluate their incremental value over smear grade and conversion trajectory.
14. Conduct a randomised evaluation of a cavitation-triggered management strategy, and define duration endpoints that encompass lung-health recovery in addition to microbiological sterilisation.

REFERENCES

1. Urbanowski ME, Ordonez AA, Ruiz-Bedoya CA, Jain SK, Bishai WR. Cavitory tuberculosis: the gateway of disease transmission. *Lancet Infect Dis*. 2020. *Consensus:* <https://consensus.app/papers/details/96e93c53b98a582680ea4bc2e071e50a/>
2. Campbell JR, et al. Association of indicators of extensive disease and rifampin-resistant tuberculosis treatment outcomes: an individual participant data meta-analysis. *Thorax*. 2023. *Consensus:* <https://consensus.app/papers/details/db0e7df60b7c5d03b7bc0233b89b54f9/>
3. Holtz TH, Sternberg M, Kammerer S, Laserson KF, Riekstina V, Zarovska E, et al. Time to sputum culture conversion in multidrug-resistant tuberculosis: predictors and relationship to treatment outcome. *Ann Intern Med*. 2006. *Consensus:* <https://consensus.app/papers/details/cf7fd77a9db85cc3b6a3a56ecb47562d/>
4. Leimane V, Dravniece G, Riekstina V, Sculier D, Zalesky R, Jansone A, et al. Treatment outcome of multidrug/extensively drug-resistant tuberculosis in Latvia, 2000–2004. *Eur Respir J*. 2010. *Consensus:* <https://consensus.app/papers/details/b8c3f6eea93e5b92a3a31134133693d5/>
5. Nunn AJ, Phillips PPJ, Meredith SK, Chiang CY, Conradie F, Dalai D, et al. A trial of a shorter regimen for rifampin-resistant tuberculosis (STREAM stage 1). *N Engl J Med*. 2019. *Consensus:* <https://consensus.app/papers/details/4701863b482b5b5d98ffb3faa07d406d/>
6. Goodall RL, Meredith SK, Nunn AJ, Bayissa A, Bhatnagar AK, Bronson G, 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. *Consensus*: <https://consensus.app/papers/details/d8fa67c6be7957069a762087eb0c8afa/>
7. Goodall RL, et al. Long-term efficacy and safety of two short standardised regimens for the treatment of rifampicin-resistant tuberculosis (STREAM stage 2): extended follow-up of an open-label, multicentre, randomised, non-inferiority trial. *Lancet Respir Med*. 2024. *Consensus*: <https://consensus.app/papers/details/ec09beffb7035f47b70726b77d648d6a/>
8. Nyang'wa BT, Berry C, Kazounis E, Motta I, Parpieva N, Tigay Z, et al. Short oral regimens for pulmonary rifampicin-resistant tuberculosis (TB-PRACTECAL): an open-label, randomised, controlled, phase 2B-3, multi-arm, multicentre, non-inferiority trial. *Lancet Respir Med*. 2023. *Consensus*: <https://consensus.app/papers/details/9307db0a65f4500983c52c2e9ff1ffc7/>
9. Mok J, Lee M, Kim DK, Kim JS, Jhun BW, Jo KW, 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. *Consensus*: <https://consensus.app/papers/details/3a277fbaf6ee53b4b29ff1737536ff94/>
10. Esmail A, Oelofse S, Lombard C, Perumal R, Mbuthini L, Goolam Mahomed A, et al. An all-oral 6-month regimen for multidrug-resistant TB (the NExT Study): a multicenter, randomized controlled trial. *Am J Respir Crit Care Med*. 2022. *Consensus*: <https://consensus.app/papers/details/00d52342db66530b881aca41db21d3d6/>
11. Du Y, Qiu C, Chen X, Wang J, Jing W, Pan H, et al. Treatment outcome of a shorter regimen containing clofazimine for multidrug-resistant tuberculosis: a randomized control trial in China. *Clin Infect Dis*. 2020. *Consensus*: <https://consensus.app/papers/details/835622dc2ddb53d284ffa11b9a9cf1bf/>
12. Cevik M, Thompson LC, Upton C, Rolla VC, Malahleha M, Mmbaga B, et al. Bedaquiline-pretomanid-moxifloxacin-pyrazinamide for drug-sensitive and drug-resistant pulmonary tuberculosis treatment (SimpliciTB): a phase 2c, open-label, multicentre, partially randomised controlled trial. *Lancet Infect Dis*. 2024. *Consensus*: <https://consensus.app/papers/details/02ed24e7b3b151dc831215c59a6c6e8e/>

13. Brust JCM, Lygizos M, Chaiyachati K, Scott M, van der Merwe TL, Moll AP, et al. Chest radiograph findings and time to culture conversion in patients with multidrug-resistant tuberculosis and HIV in Tugela Ferry, South Africa. PLoS One. 2013. *Consensus*:
<https://consensus.app/papers/details/24e7ef20990957a58d7fcf1ddb9b5b22/>
14. Basit A, Ahmad N, Khan AH, Javaid A, Syed Sulaiman SA, Afridi AK, et al. Predictors of two months culture conversion in multidrug-resistant tuberculosis: findings from a retrospective cohort study. PLoS One. 2014. *Consensus*:
<https://consensus.app/papers/details/da9644a85fe4530a8ef3e51154c771e5/>
15. Chen MY, Lo YC, Chen WC, Wang KF, Chan PC. Recurrence after successful treatment of multidrug-resistant tuberculosis in Taiwan. PLoS One. 2017. *Consensus*:
<https://consensus.app/papers/details/9f8bd452a4a851b1b57de8c2ef8a19af/>
16. Akalu TY, Muchie KF, Gelaye KA. Time to sputum culture conversion and its determinants among multi-drug resistant tuberculosis patients at public hospitals of the Amhara Regional State: a multicenter retrospective follow up study. PLoS One. 2018. *Consensus*:
<https://consensus.app/papers/details/1e6abae67df4531d8810586b825bfdc6/>
17. Javaid A, Ahmad N, Khan AH, Shaheen Z. Validity of time to sputum culture conversion to predict cure in patients with multidrug-resistant tuberculosis: a retrospective single-center study. Am J Trop Med Hyg. 2018. *Consensus*:
<https://consensus.app/papers/details/74d0a23a8a035f0290fc6dad13a04cc1/>
18. Franke MF, Khan P, Hewison C, Khan U, Huerga H, Seung KJ, et al. Culture conversion in patients treated with bedaquiline and/or delamanid: a prospective multicountry study (endTB). Am J Respir Crit Care Med. 2020. *Consensus*:
<https://consensus.app/papers/details/e0981be48f955ab3b0d4d8cf8ce9e0b6/>
19. Soeroto AY, Pratiwi C, Santoso P, Lestari BW. Factors affecting outcome of longer regimen multidrug-resistant tuberculosis treatment in West Java Indonesia: a retrospective cohort study. PLoS One. 2021. *Consensus*:
<https://consensus.app/papers/details/48274719c8c252bcbe75f1534405a0f0/>

20. Vashakidze S, Gogishvili S, Nikolaishvili K, Avaliani Z, Kempker RR, Blumberg HM, et al. Adjunctive surgery versus medical treatment among patients with cavitary multidrug-resistant tuberculosis. *Eur J Cardiothorac Surg*. 2021. *Consensus*: <https://consensus.app/papers/details/ac9f2da9f9c1540b8695c4da8f856c62/>
21. Kherabi Y, Fréchet-Jachym M, Rioux C, Yazdanpanah Y, Mechai F, Pourcher V, et al. Revised definitions of tuberculosis resistance and treatment outcomes, France, 2006–2019. *Emerg Infect Dis*. 2022. *Consensus*: <https://consensus.app/papers/details/9bf40bf9a32f572c9ce16b951dd2fe58/>
22. Meshesha MD. Predictors of sputum culture conversion time among MDR/RR TB patients on treatment in a low-income setting. *PLOS ONE*. 2022. *Consensus*: <https://consensus.app/papers/details/87d205166a445f6bbbbd7d9a6ff12b75/>
23. Li Y, Xu Z, An C, Chen H, Li X. Machine learning and radiomics for the prediction of multidrug resistance in cavitary pulmonary tuberculosis: a multicentre study. *Eur Radiol*. 2022. *Consensus*: <https://consensus.app/papers/details/43d654306f6d5e7bad91479fa60e754c/>
24. Massud A, Syed Sulaiman SA, Ahmad N, Shafqat M, Chiau Ming L, Khan AH. Unsuccessful treatment outcome and associated risk factors: a prospective study of DR-TB patients from a high burden country, Pakistan. *PLOS ONE*. 2023. *Consensus*: <https://consensus.app/papers/details/2f4447cd62f75adab743afcb38eb88cd/>
25. Kho S, Seung KJ, Huerga H, Bastard M, Khan PY, Mitnick CD, et al. Sputum culture reversion in longer treatments with bedaquiline, delamanid, and repurposed drugs for drug-resistant tuberculosis. *Nat Commun*. 2024. *Consensus*: <https://consensus.app/papers/details/5c99dbda43655910957f42d9c1788d44/>
26. Kempker RR, Kipiani M, Avaliani Z, Blumberg HM, Vashakidze S, Barbakadze K, et al. High rates of post tuberculosis lung disease among persons successfully treated for drug susceptible and resistant tuberculosis. *Thorax*. 2025. *Consensus*: <https://consensus.app/papers/details/3ae267e2f7375536bfd54cf3bf863e79/>
27. Singla R, Nair D, Bhalla M, Sharma S, Isaakidis P, Das M, et al. Bedaquiline resistance and treatment outcomes among patients with tuberculosis previously exposed to bedaquiline in

India: a multicentric retrospective cohort study. Clin Infect Dis. 2025. *Consensus:*

<https://consensus.app/papers/details/61a4720021e05c6891eef0773da761eb/>

28. Mbuagbaw L, Guglielmetti L, Hewison C, Bakare N, Bastard M, Caumes E, et al. Outcomes of bedaquiline treatment in patients with multidrug-resistant tuberculosis. Emerg Infect Dis. 2019. *Consensus:* <https://consensus.app/papers/details/206170161d7459ae9c53a51a95aac004/>

29. Rosu L, Madan JJ, Tomeny EM, Muniyandi M, Nidoi J, Girma M, et al. Economic evaluation of shortened, bedaquiline-containing treatment regimens for rifampicin-resistant tuberculosis (STREAM stage 2): a within-trial analysis of a randomised controlled trial. Lancet Glob Health. 2022. *Consensus:* <https://consensus.app/papers/details/93887a4369ec563cbe04ce6b976368d3/>

30. Dorman SE, Nahid P, Kurbatova EV, Phillips PPJ, Bryant K, Dooley KE, et al. Four-month rifapentine regimens with or without moxifloxacin for tuberculosis. N Engl J Med. 2021. *Consensus:* <https://consensus.app/papers/details/89efd4af58375582978e0af2b7358006/>

DECLARATIONS

Author Contributions

Amelia Theresya Arthin Saragih conceived the review question, designed the protocol, executed the structured literature searches, performed study selection, data extraction and risk-of-bias appraisal, conducted the statistical synthesis, and drafted the manuscript. Elfida Purba independently duplicated study selection, data extraction and risk-of-bias appraisal, adjudicated the design-based exclusion decisions, verified the extracted effect estimates against the source reports, and critically revised the manuscript for important intellectual content. Both authors read and approved the final manuscript and accept accountability for all aspects of the work.

Funding

This review received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interest

The authors declare that they have no competing interests, financial or non-financial, relevant to the content of this review.

Ethics Approval and Consent to Participate

This study is a systematic review of previously published, publicly available literature. No individual participant was recruited and no identifiable personal data were accessed, and formal ethical approval and informed consent were therefore not required.