https://dspace.org/

Guiding the development of the tuberculosis screening target product profile using singlescreen and multiscreen approaches a modelling study

Alexandra de Nooy; Cecily Miller; Tom Ockhuisen; Dennis Falzon; Nazir Ismail ...

ELSEVIER SCI LTD
https://hdl.handle.net/10539/49035
WOS

Downloaded from a DSpace repository. https://dspace.org/

Guiding the development of the tuberculosis screening target product profile using single-screen and multi-screen approaches: a modelling study

Alexandra de Nooy, Cecily Miller, Tom Ockhuisen, Dennis Falzon, Alexei Korobitsyn, Morten Ruhwald, Nazir Ismail, Mikashmi Kohli, Brooke E Nichols

Summary

Background Reducing tuberculosis incidence requires new, symptom-agnostic screening tools. Target product profiles guide their development by defining performance. In this study, we aimed to support the 2025 WHO tuberculosis screening target product profile update by modelling the minimum sensitivity and specificity thresholds needed across different prevalence levels and screening algorithms.

Methods In this modelling study, we used decision trees and parameter multiplication to simulate screening strategies across four levels of tuberculosis prevalence (0·10%, 0·25%, 0·50%, and 1·00%). Tests were considered acceptable if they achieved a post-screen prevalence of 5% or higher and diagnosed at least 60% of tuberculosis cases after confirmatory testing. We modelled high-sensitivity (defined as ≥90% sensitivity) and high-specificity tests (defined as ≥98% specificity) within one-screen algorithms (novel screening test followed by confirmatory testing) and two-screen algorithms (inclusive of a secondary screen approximated by chest x-ray performance).

Findings At 1·00% prevalence, a test with 90% sensitivity required specificities of at least 83% (one-screen) and at least 52% (two-screen). High-specificity tests (98%) required sensitivities of 68% (one-screen) or 80% (two-screen) across all prevalence levels, with the exception of 0·10% prevalence in a one-screen algorithm. At 0·10% prevalence, no acceptable one-screen tests were identified. At 0·10% prevalence, tests in two-screen algorithms required 90% sensitivity with at least 96% specificity or at least 80% sensitivity with 98% specificity.

Interpretation In higher tuberculosis prevalence settings, a test with moderate sensitivity and specificity is sufficient to achieve high case detection without overloading the confirmatory diagnostic cascade. Screening in very low prevalence settings remains challenging and requires a very high-sensitivity and high-specificity screening test even within two-screen algorithms.

Funding Bill and Melinda Gates Foundation.

Copyright © 2025 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Lancet Glob Health 2025; 13: e1750–60

Department of Global Health, Amsterdam University Medical Center, Amsterdam, Netherlands (A de Nooy MSc, T Ockhuisen MSc, B E Nichols PhD); **Amsterdam Institute of Global Health and Development, Amsterdam, Netherlands** (A de Nooy, T Ockhuisen); **Global Tuberculosis Programme, WHO, Geneva, Switzerland** (C Miller PhD, D Falzon MD, A Korobitsyn MD, Prof N Ismail FCPath); **FIND, Geneva, Switzerland** (M Ruhwald MD, M Kohli PhD); **Novo Nordisk Foundation Initiative for Vaccines and Immunity, Copenhagen, Denmark** (M Ruhwald); **Department of Clinical Microbiology and Infectious Diseases** (Prof N Ismail) **and Wits Diagnostic Innovation Hub, Faculty of Health Sciences** (B E Nichols), **University of the Witwatersrand, Johannesburg, South Africa; Department of Global Health, Boston University School of Public Health, Boston, MA, USA** (B E Nichols)

Correspondence to: Ms Alexandra de Nooy, Department of Global Health, Amsterdam University Medical Center, Amsterdam 1105 AZ, Netherlands **a.denooy1@amsterdamumc.nl**

Introduction

Tuberculosis remains a global health challenge, with the 2025 END TB strategy targets unlikely to be met. Relative to 2015, these targets aimed for a 50% reduction in tuberculosis incidence and a 75% reduction in mortality by 2025; however, by 2023, reductions of 8·3% and 23·0%, respectively, had been achieved. Additionally, the goal of zero tuberculosis-affected households facing catastrophic costs is unmet at around 49%.[1,2] A key barrier to progress is the case-detection gap. Although millions of people are notified with tuberculosis each year, estimates suggest that only 75% of new cases were identified in 2023.[2] In addition to under-reporting, many individuals remain undiagnosed, or are incorrectly diagnosed, because of inadequate access to accurate and appropriate diagnostics.[3] Furthermore, many individuals are either unaware of the need to seek care for symptoms or might have subclinical tuberculosis and are asymptomatic.[2–5] Considering that many tuberculosis programmes rely on passive case finding, in which only individuals presenting at health-care facilities with tuberculosis symptoms might be referred for testing, these individuals will likely remain undiagnosed.[6,7] Although resource intensive, active case finding can reduce the diagnostic gap by identifying individuals unlikely to present at health-care facilities. Although active case finding commonly increases case notifications,[8] an impact on prevalence is not always observed. The 4-year TREATs study in South Africa and Zambia aimed to quantify the effect on tuberculosis prevalence that universal test and treatment of HIV, in conjunction with universal tuberculosis symptom screening, might have in communities. Prevalence was found to be unaffected, likely due to the intervention's focus on symptomatic individuals and close contacts only.[9] To improve early detection and reduce missed diagnoses, there is need for effective, affordable, and symptom-agnostic tools suitable for both facility-based and community active case finding programmes.

Few tuberculosis screening tests and tools are currently fit for use in high-burden settings. One such tool is the WHO-recommended four-symptom screen.[10] Others

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

Research in context

Evidence before this study

Tuberculosis remains a major global health challenge, with efforts to reduce incidence hindered by the small range of available screening tools that are accurate, easy to implement, and affordable for large groups. Target product profiles guide the development of new screening tests by defining their minimum accuracy requirements and expected use cases. We searched PubMed and WHO for studies published before March 1, 2025, using the search terms "TB screening", "target product profile", "screening algorithms", and "low-prevalence TB". Although tuberculosis screening has been a topic of interest in previous tuberculosis target product profiles, it has always been incorporated as a component alongside diagnostic test requirements, as opposed to having a dedicated target product profile. Furthermore, in these past target product profiles there was limited quantitative consideration of the trade-offs between sensitivity (important for case detection) and specificity (which impacts subsequent confirmatory test numbers). Additionally, previous target product profiles for tuberculosis predominantly relied on expert opinion, as opposed to tools such as mathematical modelling, to select accuracy thresholds for novel tests. Previous work by WHO has typically set screening guidelines around prevalences of 0·5% and higher. Although important gains could be made in screening at these levels of prevalence, it is important to consider that eventual tuberculosis elimination would require community screening at even lower levels. This would require careful consideration since prevalence is a driver of screening programme performance. Additionally, previous research has considered a range of different screening approaches, whether this is one screening test followed by confirmatory testing, multiple screening tests used in series or parallel, or test pooling. Several studies have explored tuberculosis screening strategies and diagnostic performance, for example through symptom screening and chest x-ray; however, few have systematically quantified the relationship between prevalence, test accuracy, and screening algorithm performance.

Added value of this study

This study provides additional evidence to support the development of a tuberculosis target product profile. In this way, modelling is used to quantitatively support the selection of minimum sensitivity and specificity requirements of novel screening tests as opposed to expert opinion alone. The study highlights the potential impact of screening tests within different serial screening algorithms, and includes analysis of symptom-agnostic tools at very low prevalence levels. Different classes of screening tests (high sensitivity and high specificity) have been explored and the potential use cases of each have been defined. In conjunction with WHO, minimum thresholds of case detection and post-screen prevalence have been selected to balance the sensitivity and specificity requirements of novel tests. These resulted in a set of sensitivity and specificity targets per prevalence, test class, and algorithm. Our findings demonstrate that in higher prevalence settings (1·00%), moderate sensitivity and specificity are sufficient for effective screening. However, in very low-prevalence settings (0·10%), no screening test—regardless of accuracy—was effective in a one-screen algorithm, and very high accuracy is needed in a two-screen algorithm.

Implications of all the available evidence

Effective tuberculosis screening, especially using a method that does not rely on symptoms, is needed to minimise the case detection gap among individuals with tuberculosis. Considering that there are relatively few screening tools currently available, new research and development is required to expand the tuberculosis screening landscape. Previous target product profiles only briefly focused on tuberculosis screening and provided a small set of targets against which developers could compare new products. The 2025 target product profile update, which has been supported by this study, provides an expanded set of options for developers, contributing to the potential for new, impactful screening tests to be developed.

include abnormal chest x-ray identification by human reader or computer-aided detection software, or sputum-based molecular WHO-recommended rapid diagnostics (mWRDs).[10] Furthermore, for people with HIV, C-reactive protein tests (combined with WHO-recommended four-symptom screen) have been recommended.[10] However, each tool has trade-offs. Symptom screening is inexpensive and easy to implement, but has variable sensitivity and specificity parameters, is subjective, will miss asymptomatic (and some symptomatic) cases, and frequently has a low rate of follow-up diagnostic investigation.[11] Chest x-ray with computer-aided detection and mWRDs have higher accuracy but require more resources, infrastructure, and staff to implement.[12,13] mWRDs also require sputum, which asymptomatic individuals might not easily produce, and could have long turnaround times.[14] Additionally, considering the lower accuracy of mWRDS in a screening population than in those already referred or indicated for confirmatory testing and the need to prioritise the often constrained number of available tests for diagnosis, there is relatively little research into wide-scale screening with mWRDs.[10] These challenges highlight the gaps in current screening tests and tools and demonstrate the need to develop new, affordable, easily implemented. and rapid tuberculosis screening tests.

To drive research and development of new tests to meet this demand, it is important for manufacturers to have a reference of the desired test attributes against which their products can be compared. To guide this process, WHO developed target product profiles describing optimal and minimum requirements for each test. The defined target product profile attributes include, but are not limited to, accuracy (sensitivity and specificity), intended target

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

groups and implementation strategy, and cost.[15] As such, considering the global tuberculosis community's current interest in pursuing broader tuberculosis screening (particularly in lower prevalence settings) and the subsequent diagnostic and treatment programmes, WHO has worked to update several tuberculosis target product profiles. These have included those on tests for tuberculosis treatment monitoring and optimisation,[16] tests for tuberculosis diagnosis and detection of drug resistance[17] and, for the first time, a screening-specific target product profile (intended to replace previous triage target product profiles).[18] Generally, updating a target product profile is a complex, iterative process that requires input from literature, clinical trials, national health programmes, and subject matter experts. Developing a screening target product profile is additionally challenging because of the complexities surrounding screening programmes and their successful implementation. The efficacy and feasibility of use for a screening test typically depends on the context, the expected size of target groups being screened, the underlying tuberculosis prevalence (eg, individuals at high risk, tuberculosis contacts, and community), the screening methodology (eg, single-screen or multiple screening tests), and the overall cost per screen. Within each screening programme, a desirable screening test would have both a high enough specificity to avoid overloading the confirmatory diagnostic cascade and a high enough sensitivity to detect a large proportion of cases, and additionally would be low cost to ensure financial feasibility of implementation. The minimum desired test targets must, however, simultaneously be balanced with the abilities and ambition of product manufacturers. Targets that are too ambitious or stringent could prevent manufactures from investing in the development of new tools and hence delay potentially impactful screening tests from reaching the market.

To account for some of these complexities, the development of the screening target product profile, alongside several of the newly updated tuberculosis target product profiles, represents an initial instance in which mathematical modelling has been used to inform relevant tuberculosis target product profile parameters—in this case modelling different sensitivity and specificities of the novel screen to determine the minimum and optimal values. In this study, we aimed to use modelling as additional evidence to inform the WHO Target Product Profile Development Group's decisions on the minimum and optimal values for these parameters by exploring a novel screening test in varied prevalence and screen settings, and when used in different screening algorithms.

Methods

Target product profile screening algorithms

Two screening algorithms, in the form of decision trees (figure 1), were simulated in R studio (version 4.2.1). These represent one-screen and two-screen approaches. In the one-screen algorithm, a cohort is screened using a simulated novel screening test, with all screen-positive individuals undergoing confirmatory testing via a low-complexity automated nucleic acid amplification test (LC-aNAAT) with assumed 90% sensitivity and 96% specificity.[19,20]

In the two-screen algorithm, individuals who screen positive with the novel test undergo a second screening with an additional test (ie, serial screening). To limit the number of unconstrained variables in the target product profile analysis, this second screen was assumed to be a chest x-ray with 85% sensitivity and 70% specificity (appendix pp 2–3).[21–24] Individuals with abnormal chest x-ray results are then referred for confirmatory testing through LC-aNAAT. The final sensitivity and specificity requirements would be mathematically equivalent if the screening order were reversed (ie, chest x-ray followed by the novel screen), although this would affect implementation requirements. Additionally, this chest x-ray represents a placeholder technology for a secondary screen, and could effectively be replaced by any screening tool with similar accuracy. Additional analyses, external to the target product profile, assess the impact of uncertainty surrounding the chest x-ray accuracy and guide how results might change if a different screen was used (appendix p 4).

See Online for appendix

Due to the nature of this study, and use of publicly available data, no ethical approval was required for this work.

Modelled screening tests

Based on discussions with the expert Target Product Profile Development Group, two main classes of screening tests were considered. These are high-sensitivity (90% per previous target product profiles) and high-specificity (98% based on current diagnostics) tests. In simulating each test, one accuracy parameter (sensitivity or specificity) was held constant while the other was varied between 40% and 98%. An additional analysis simultaneously varied sensitivity and specificity and determined all acceptable combinations that met the prerequisites.

Minimum prerequisites for target product profile

When simulating the described screening tests and algorithms, it was necessary to define prerequisites for determining the minimum acceptable test sensitivity and specificity. In consultation with WHO and the target product profile Development Group, two prerequisites were established: (1) a case detection (percentage of individuals with tuberculosis correctly diagnosed) of 60% or higher to reflect the working group decision that a screening test should be able to identify at least half of all people screened and (2) a post-screen prevalence (ie, the proportion of individuals who screen positive that truly

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

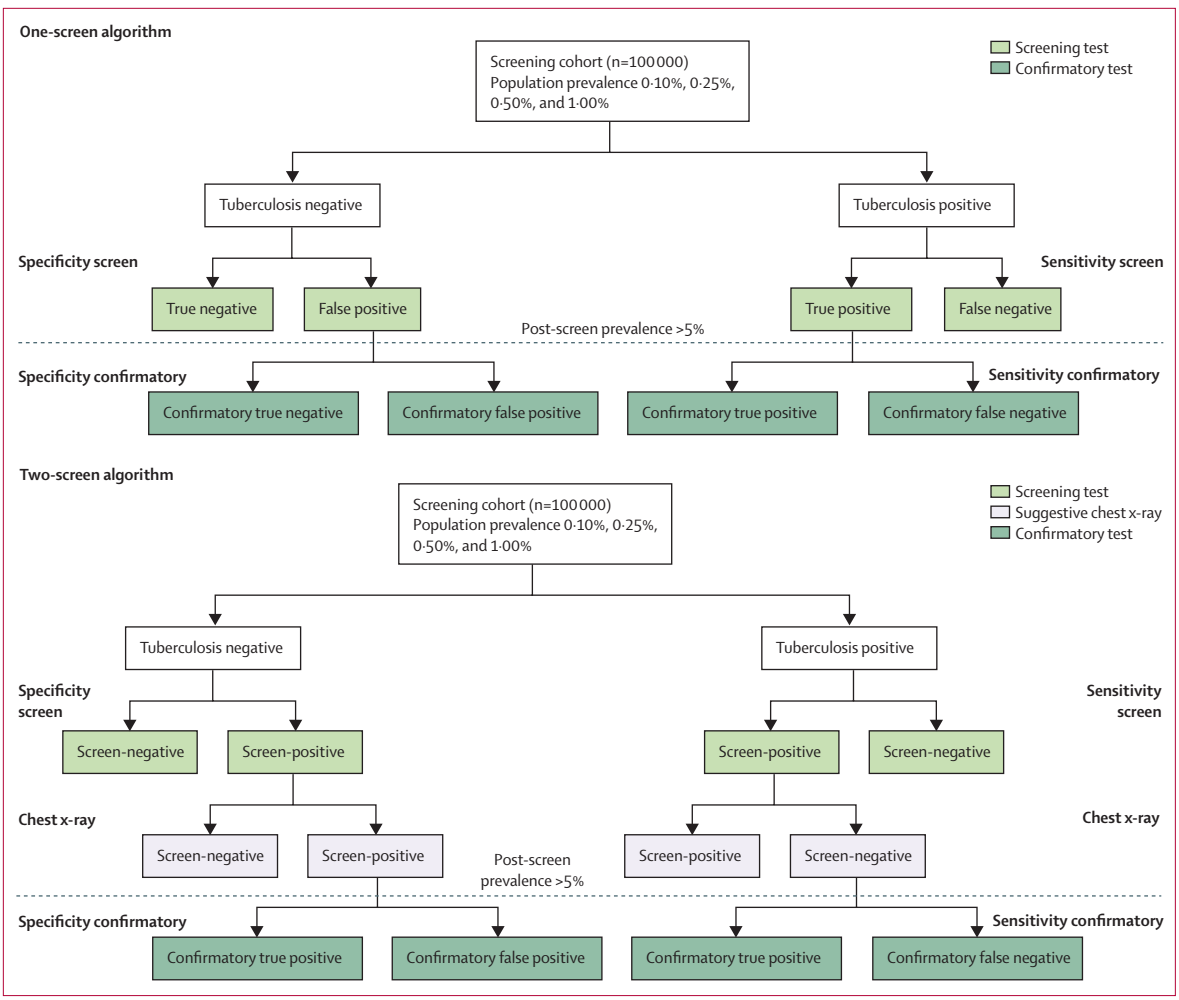


Figure 1: **Decision tree model and included tests for the one-screen and two-screen algorithms**
This figure highlights the model parameters used to calculate the number of individuals at each stage of the decision tree.

have tuberculosis) of 5% or higher to reflect positivity rates observed among individuals seeking care in high-burden countries. These values were calculated for each test and algorithm combination using the following equations:

$$\text{Case detection} = \frac{TP_{confirmed}}{TB_{total}}$$

$$\text{One-screen post-screen prevalence} = \frac{TP_{screen1}}{TP_{screen1} + FP_{screen1}}$$

$$\text{Two-screen post screen prevalence} = \frac{TP_{screen2}}{TP_{screen2} + FP_{screen2}}$$

where $TP_{confirmed}$ represents true positive cases detected after confirmatory testing with LC-aNAAT; TB_{total} represents the number of individuals with tuberculosis; $TP_{screen1}$ and $FP_{screen1}$ represent the true and false positive results after one screening test is conducted; and $TP_{screen2}$ and $FP_{screen2}$ represent the true and false positive results after two serial screening tests are conducted.

Model scenarios and implementation

For each combination of test and algorithm, we simulated a screening programme for an assumed cohort of 100 000 individuals (to have sufficient numbers to explore very low prevalence). Additionally, we simulated four prevalence levels selected by the target product profile development group: 0·10%, 0·25%, 0·50%, and 1·00%. Using the scenario conditions (ie, prevalence, algorithm, and novel screening test class), the population present at every step along the depicted decision tree was calculated through parameter multiplication. Using the results, the case detection and post-screen prevalence was determined. For the high-sensitivity and high-specificity tests, these metrics were plotted against the accuracy parameter being varied (either sensitivity or specificity). Additionally, the number of screening and confirmatory tests required per scenario is calculated and plotted, with these values

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

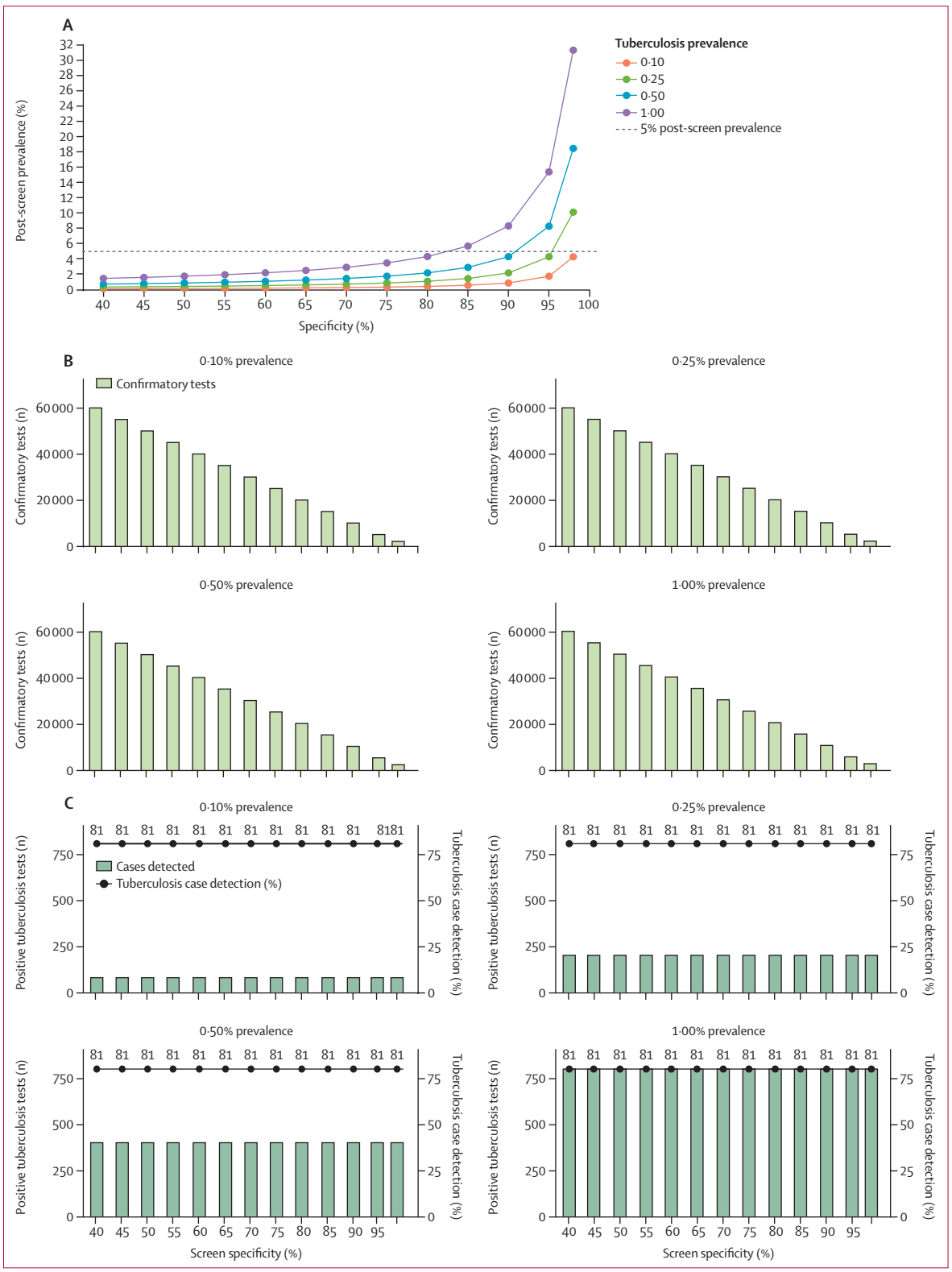


***Figure 2*: Post-screen prevalence and case detection results for a test with high sensitivity (90%) in a one-screen algorithm**
(A) Post-screen prevalence by tuberculosis prevalence and specificity of novel screen. (B) Number of required confirmatory tests by prevalence and specificity. (C) Number of positive tuberculosis tests and tuberculosis case detection (the percentage of all individuals with tuberculosis correctly diagnosed).

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

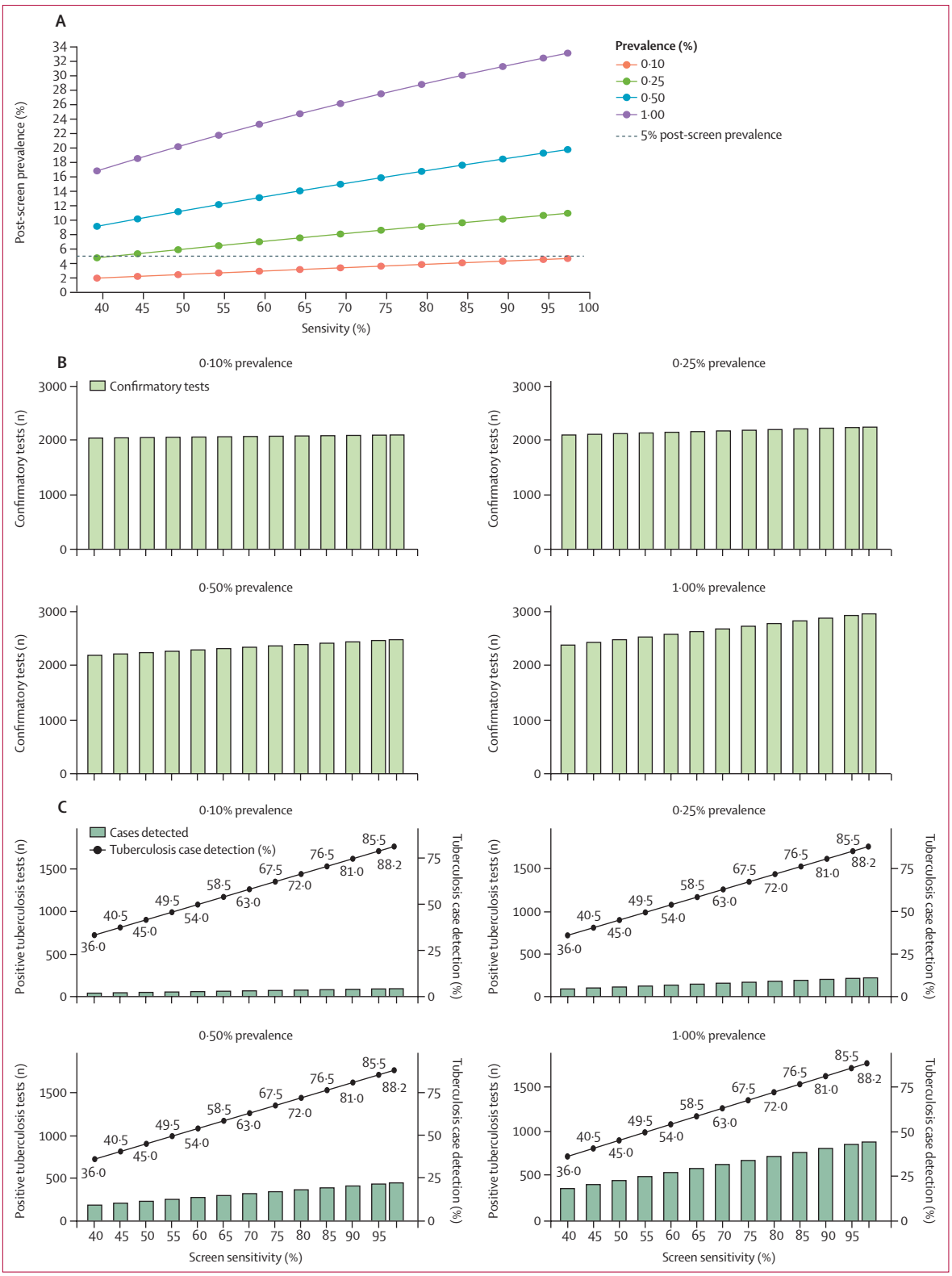


***Figure 3*: Post-screen prevalence, number of required tests, and case detection results for a high-specificity test (98%) in a one-screen algorithm**
(A) Post-screen prevalence results by tuberculosis prevalence and screen sensitivity. (B) Number of required confirmatory tests per prevalence and sensitivity. (C) Number of positive tuberculosis tests (green bars) and tuberculosis case detection (the percentage of all individuals with tuberculosis correctly diagnosed).

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

subsequently being used for a simple secondary costing analysis that was not conducted for the target product profile (appendix pp 7–8) to understand the screening costs alongside the related downstream diagnostic costs. In scenarios meeting the selected target product profile prerequisites, the minimum accuracy characteristics for each combination of prevalence, screening algorithm and screening test class were selected for target product profile consideration.

Role of the funding source

The funder had no role in study design, data collection and analysis, decision to publish, or writing of the manuscript.

Results

For a high-sensitivity test (90%) in a one-screen algorithm, the prerequisites were met for the 0·25%, 0·50%, and 1·00% prevalence levels with minimum specificity requirements of around 96%, 92%, and 83%, respectively (figure 2). The post-screen prevalence prerequisite drove these requirements given that the constant 90% sensitivity resulted in a case detection (the percentage of individuals with tuberculosis correctly diagnosed) of 83% regardless of specificity. At 0·10% prevalence, the 5% post-screen prevalence threshold could not be reached, which necessitated the requirement for a second screening step. The case detection and post-screen prevalence results for a high-sensitivity test in a two-screen algorithm are shown in the appendix (p 5). For the two-screen algorithm, case detection remained at 69% and the post-screen prevalence prerequisite was met at the following minimum specificity values for each prevalence: 96% (0·10% prevalence), 87% (0·25% prevalence), 75% (0·50% prevalence), and 52% (1·00% prevalence). In terms of confirmatory tests, across algorithms and prevalence, the number of diagnostic tests required decreases with increased specificity. This corresponds to lower costs per individual correctly diagnosed as specificity increased (appendix p 8). Prevalence also drove cost differences. For example, at 85% specificity in a one-screen algorithm, the cost per individual correctly diagnosed was US$526 and $5686 at 1·00% and 0·10% prevalence, respectively. Accuracy requirements varied with differing chest x-ray (or second screen) accuracy values (appendix p 4).

To meet the case detection and post-screen prevalence prerequisites for a high-specificity test (98%) in a one-screen algorithm, a minimum sensitivity value of 68% was required for prevalence levels greater than or equal to 0·25% (figure 3). For 0·10% prevalence, only a high-specificity test in a two-screen algorithm met both prerequisites (figure 3A; appendix p 6), where a minimum sensitivity of 80% was required across all four prevalence levels. Additionally, while the total number of confirmatory tests required increased with increasing screen sensitivity, the percentage of the screening cohort receiving these tests was low (<3000 diagnostic tests or <0·03% across both algorithms and all sensitivities; figure 3B; appendix p 6). Across algorithms, higher sensitivity resulted in lower costs per individual correctly diagnosed given the corresponding increase in case detection (appendix p 8). Additionally, as with the high-sensitivity test, higher prevalence resulted in lower costs per individual diagnosed.

The minimum accuracy values that met both the post-screen prevalence and case detection prerequisites for each combination of test class and algorithm are shown in the table. These represent the values presented to the WHO Target Product Profile Development Group for consideration.

The unconstrained sensitivity and specificity analysis presented all acceptable combinations of sensitivity and specificity when both accuracy parameters were simultaneously varied. At 0·10% prevalence, only a test with 95% sensitivity and specificity in a two-screen algorithm met the target product profile prerequisites (figure 4). At 0·25% prevalence, no one-screen combinations were acceptable, while a two-screen algorithm required at least 80% sensitivity and 90% specificity. At 1·00% prevalence multiple potential accuracy combinations were acceptable for both algorithms. In the one-screen algorithm, sensitivity could reduce to 70% for a test with 90% specificity, or specificity could be reduced to 85% if sensitivity was 80%. In a two-screen algorithm, the lowest specificity test (50%) would require a sensitivity of 95%. Conversely, the test with the lowest sensitivity (80%) would require a test specificity of 60%.

Discussion

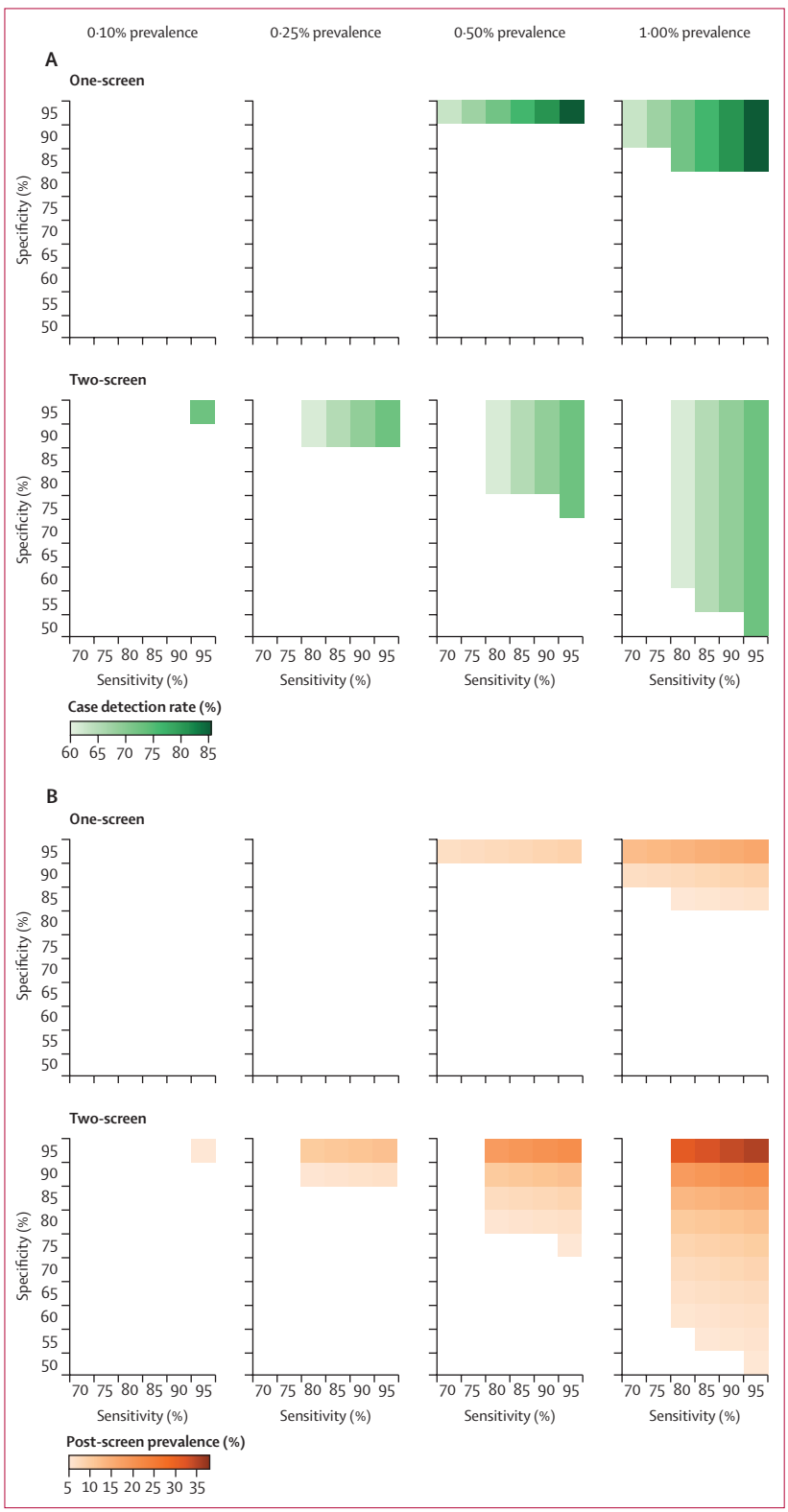
Currently, there are no screening tests available that meet the 2014 target product profile screening requirements of 90% sensitivity and 70% specificity at low tuberculosis prevalences.[10] However, since 2014, the technology landscape has expanded and is increasing rapidly. There is increasing interest in the early detection of all individuals with tuberculosis irrespective of

	One-screen algorithm		Two-screen algorithm	
	90% sensitivity*	98% specificity†	90% sensitivity*	98% specificity†
0·10%	Not possible	Not possible‡	96% specificity	80% sensitivity
0·25%	96% specificity§	68% sensitivity	87% specificity	80% sensitivity
0·50%	92% specificity	68% sensitivity	75% specificity	80% sensitivity
1·00%	83% specificity	68% sensitivity	52% specificity	80% sensitivity

*Requirements driven by post-screen prevalence prerequisite. †Requirements driven by case detection prerequisite. ‡Exception where post-screen prevalence prerequisite was not met. §Not included in the target product profile as during the first iterations a maximum of 95% was considered due to a consensus on the low feasibility of such a screen.

Table: **Minimum accuracy requirements for screening tests in the one-screen and two-screen algorithms at different tuberculosis prevalences**

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.



symptom presentation. This has led to greater interest in conducting large-scale screening in lower prevalence settings, and subsequently to an increased demand for new screening tools. To cater to this demand, a target product profile update was initiated, and we used a model-based approach to inform the included sensitivity and specificity parameters. We used modelling to assess acceptable parameters of a hypothetical novel screening test for various tuberculosis prevalences and in several screening algorithms. By identifying an array of potential minimum test characteristics, we aimed to provide multiple possible targets for developers.

At the highest modelled prevalence of 1·00%, a high-sensitivity test had requirements of 90% sensitivity and 83% specificity. Notably, these updated values are higher than those for screening in the previous target product profile.[25] The 2014 target product profile has minimum requirements for screening of 90% sensitivity and 70% specificity, which have themselves not yet been met by an easily implemented tool.[25] Additionally, existing tools might not meet both prerequisites defined in this study. For example, using the models developed in this work, chest x-ray and symptom screening (among people with HIV) within one-screen algorithms were simulated to have case detection rates of 76·5% and 74·7%, and post-screen prevalence values ranging between 0·28–2·78% and 0·13–1·3%, respectively. This prompted the consideration of two-screen algorithms (new to this target product profile) that might have more achievable accuracy values while simultaneously having a positive impact on the tuberculosis screening landscape. For example, at 1·00% prevalence, a test with 90% sensitivity and 52% specificity in a two-screen algorithm (in addition to a second screen similar to chest x-ray) would meet the defined prerequisites. In the near-term, a simple, affordable rapid test with these accuracy characteristics could assist in identifying individuals with tuberculosis who would otherwise be missed.

Generally, regardless of test class, as prevalence increases the required accuracy decreases. Further high-sensitivity tests within two-screen algorithms, have lower specificity requirements than in one-screen algorithms and can meet prerequisites down to low population prevalences of 0·10% (although 96% specificity and 90% sensitivity would be required). In terms of a high-specificity test, higher sensitivity is needed in two-screen algorithms to achieve the desired case detection, but can be used at low prevalence (0·10%). For the target product profile, these two classes of test and the choice of chest

***Figure 4*: Acceptable combinations of sensitivity and specificity for a novel screen that meets the 5% post-screen prevalence and 60% case detection prerequisites**

Coloured blocks represent acceptable combinations of sensitivity and specificity, whereas blank blocks represent unacceptable combinations. (A) Case detection rate. (B) Post-screen prevalence. The colour gradients in panel A and B represent the case detection rate and post-screen prevalence, respectively.

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

x-ray as a second screen were chosen intentionally to provide discrete targets and to limit the challenges to both manufacturers and regulators that target ranges might have introduced. However, the unconstrained accuracy analysis does highlight potential other combinations of sensitivity and specificity, which meet the prerequisites, with increased prevalence expanding the range of acceptable combinations. Additionally, the uncertainty analysis around the chest x-ray accuracy estimates provided insight into how requirements could shift if the accuracy of the second screen were to differ (either due to the characteristics of the screened population or type of screen).

With respect to implementation, a potentially key difference between high-sensitivity and high-specificity tests would be the number of follow-up screening and diagnostic tests. For example, at 1·00% prevalence and in a one-screen algorithm, a test with 90% sensitivity and 70% specificity could result in around 30 600 confirmatory tests being requested. Conversely, a test with 98% specificity would result in fewer than 3000 confirmatory tests regardless of sensitivity. This highlights specificity as an important component, particularly for programmes focused on low prevalence or community settings where the number of individual screened is likely to be high. The high-specificity test might reduce costs by limiting the number of subsequent confirmatory tests but might miss more people with tuberculosis if sensitivity is low. In a context where identification of tuberculosis is imperative, particularly among higher-risk individuals or in high-transmission settings, the high-sensitivity screen might be more appropriate. Thus, the type of test chosen must consider the objective of the screening programme and the target group being considered. Additionally, screening test cost and tuberculosis prevalence must be accounted for given the difference in cost per individual correctly diagnosed.

There are several limitations to this analysis. First, we did not explicitly account for various populations at risk of tuberculosis, although the simulated prevalence levels are likely to be representative of many of these groups. It would be useful, once novel screening tests are developed, to explore the impact that the differences in tuberculosis presentation and diagnosis among these groups would have on a screening programme's performance, particularly when accounting for heterogeneity in the screening population.[3,20,26] Furthermore, the overall implementation of screening programmes will differ substantially based on the context, current standard of care, algorithm, cohort size, and group being considered.[10] These complexities make it challenging to estimate an average global screening programme. Consequently, we did not model known screening challenges such as patients not returning for diagnosis or treatment, or the potential impact on diagnostic programmes, including a high patient influx and increased strain on limited diagnostics resources.[27–29] Considering that only broad characteristics of the test have been modelled, and that it is unknown what these tests might look like in terms of sample type, turnaround time, or infrastructure requirements when developed, aspects such as uptake and or acceptability of these screening tests could not be incorporated.[30] Once the characteristics and cost of newly developed screening tools and the intention of a specific screening programme are known, it would be essential to model these practical considerations to inform implementation.

Second, only two relatively simple serial screening algorithms were considered. Alternative approaches, such as parallel or pooled screening, would need to be further evaluated and compared with those presented in this study.[10,31,32] Additionally, this analysis assumes diagnostic confirmation through molecular testing when evaluating accuracy requirements. However, access to molecular testing might be limited in some settings. Consequently, if confirmation relies on less accurate technology, higher accuracy requirements may be required for the screening test. Notably, although smear microscopy is still widely used, its lower sensitivity (around 64%[30]) might prevent the case detection criteria from being reached in the modelled scenarios.[33] Third, only a simple costing analysis (in terms of expected numbers of each test) has been conducted. Full cost-effectiveness analyses would be highly dependent on context and the specific screening programme (size, target group, and objective). Additionally, costs would be driven by characteristics of the screening test that affect its implementation (eg, if it requires new infrastructure, equipment, or more personnel) and the diagnostic test that is used for confirmation.[34] Once these elements are known, additional context-driven costing analyses should be conducted.

Tuberculosis screening is complex to model and its performance is highly dependent on a range of factors, only one of which is the screening test accuracy. However, this work highlighted the potential trade-offs in performance directly related to the test accuracy characteristics, depending on the prevalence and algorithm followed, and used these results to recommend target product profile accuracy requirements. We showed the potential that high-specificity tests could have in screening lower prevalence populations and highlighted the impact of test specificity on resource requirements (in terms of confirmatory tests required). Additionally, this work presented the potential need for multiple screening stages as countries move towards very low prevalence screening. Overall, this work produced an initial framework that could provide results to support target product profile decision making, both in terms of content included, for example the choice to incorporate different screening algorithms, and the final accuracy values. Furthermore, although modelling has limitations, it proved to be a valuable addition to the target product profile development process by providing

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

quantitative evidence to support chosen accuracy values, rather than relying solely on expert opinion.

This modelling work provides an evidence base for the minimal screening test accuracy requirements needed to achieve the case detection and post-screen prevalence prerequisites. Through the identification of an array of possible minimum test characteristics, we have provided multiple acceptable possible targets for developers. Test requirements vary widely depending on the type of test, screening algorithm, and prevalence of the screening group. Notably, not all types of screening tests are suitable in settings with a very low prevalence. For example, only a highly sensitive and highly specific test in a two-screen algorithm is appropriate for settings in which tuberculosis prevalence is 0·10%.

Contributors

AdN, MK, and BEN conceived and designed the analysis. AdN performed the analysis. BEN and TO reviewed and verified the data. All authors participated in interpreting the results. AdN and BEN wrote the paper. All authors reviewed and edited the paper. All authors were involved in the decision to publish the manuscript.

Declaration of interests

DF, AK, and CM are employees of WHO. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of WHO. All other authors declare no competing interests.

Data sharing

Algorithms and values are summarised in the text. Simulation output can be made available on request and all relevant code is available online at https://github.com/adenooy2/Screening-TPP-Analysis.git.

Acknowledgments

This work was supported by Irwin Law who developed the supplemental table utilised for selecting key model parameters. During the preparation of this work the authors used ChatGPT to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

1 WHO. The end TB strategy. World Health Organization, 2015. https://www.who.int/publications/i/item/WHO-HTM-TB-2015.19 (accessed Aug 8, 2025).

2 WHO. Global tuberculosis report 2024. World Health Organization, 2024. https://iris.who.int/bitstream/handle/10665/379339/9789240101531-eng.pdf?sequence=1 (accessed Nov 17, 2024).

3 Subbaraman R, Jhaveri T, Nathavitharana RR. Closing gaps in the tuberculosis care cascade: an action-oriented research agenda. *J Clin Tuberc Other Mycobact Dis* 2020; **19:** 100144.

4 Helfinstein S, Engl E, Thomas BE, et al. Understanding why at-risk population segments do not seek care for tuberculosis: a precision public health approach in South India. *BMJ Glob Health* 2020; **5:** 2555.

5 South African National Department of Health, South African Medical Research Council, Human Sciences Research Council, National Institute for Communicable Diseases, WHO, National Institute of Communicable Diseases. The First National TB prevalence survey—short report. 2018 https://www.nicd.ac.za/wp-content/uploads/2021/02/TB-Prevalence-survey-report_A4_SA_TPS-Short_Feb-2021.pdf (accessed Aug 5, 2022).

6 Saunders MJ, Tovar MA, Collier D, et al. Active and passive case-finding in tuberculosis-affected households in Peru: a 10-year prospective cohort study. *Lancet Infect Dis* 2019; **19:** 519–28.

7 Shewade HD, Gupta V, Satyanarayana S, et al. Active versus passive case finding for tuberculosis in marginalised and vulnerable populations in India: comparison of treatment outcomes. *Glob Health Action* 2019; **12:** 1656451.

8 MacPherson P, Shanaube K, Phiri MD, et al. Community-based active-case finding for tuberculosis: navigating a complex minefield. *BMC Global and Public Health* 2024; **2:** 1–14.

9 Klinkenberg E, Floyd S, Shanaube K, et al, and the TREATS study team. Tuberculosis prevalence after 4 years of population-wide systematic TB symptom screening and universal testing and treatment for HIV in the HPTN 071 (PopART) community-randomised trial in Zambia and South Africa: a cross-sectional survey (TREATS). *PLoS Med* 2023; **20:** e1004278.

10 WHO. WHO consolidated guidelines on tuberculosis. Module 2: systematic screening for tuberculosis disease. World Health Organization, 2021. https://tbksp.org/en/node/1274 (accessed Oct 5, 2022).

11 Divala TH, Lewis J, Bulterys MA, et al. Missed opportunities for diagnosis and treatment in patients with TB symptoms: a systematic review. *Public Health Action* 2022; **12:** 10–17.

12 WHO. Chest radiography in tuberculosis detection: summary of current WHO recommendations and guidance on programmatic approaches. World health Organization, 2016. https://iris.who.int/handle/10665/252424 (accessed Oct 23, 2024).

13 Nguyen TBP, Nguyen TA, Luu BK, et al. A comparison of digital chest radiography and Xpert® MTB/RIF in active case finding for tuberculosis. *Int J Tuberc Lung Dis* 2020; **24:** 934–40.

14 Davids M, Dheda K, Pant Pai N, Cogill D, Pai M, Engel N. A survey on use of rapid tests and tuberculosis diagnostic practices by primary health care providers in South Africa: implications for the development of new point-of-care tests. *PLoS One* 2015; **10:** e0141453.

15 WHO. WHO target product profiles. https://www.who.int/observatories/global-observatory-on-health-research-and-development/analyses-and-syntheses/target-product-profile/who-target-product-profiles (accessed July 8, 2022).

16 WHO. Target product profiles for tests for tuberculosis treatment monitoring and optimization. World Health Organization, 2023. https://www.who.int/publications/i/item/9789240081178 (accessed Oct 23, 2024).

17 WHO. Target product profiles for tuberculosis diagnosis and detection of drug resistance. World Health Organization, 2024. https://iris.who.int/bitstream/handle/10665/378358/9789240097698-eng.pdf?sequence=1 (accessed Oct 23, 2024).

18 WHO. Target product profiles for tuberculosis screening tests. 2025. https://iris.who.int/bitstream/handle/10665/382179/9789240113572-eng.pdf (accessed Aug 15, 2025).

19 Cepheid. Xpert MTB/RIF Ultra. https://www.cepheid.com/en-GB/tests/tb-emerging-infectious-diseases/xpert-mtb-rif-ultra.html (accessed Sept 27, 2023).

20 WHO. WHO consolidated guidelines on tuberculosis. Module 3: diagnosis-rapid diagnostics for tuberculosis detection. World Health Organization 2021. https://www.who.int/publications/i/item/9789240089488 (accessed Aug 8, 2025).

21 WHO. National tuberculosis prevalence surveys: 2007–2016. World Health Organization, 2021. https://iris.who.int/bitstream/handle/10665/341072/9789240022430-eng.pdf (accessed Oct 23, 2024).

22 Moyo S, Ismail F, Van der Walt M, et al. Prevalence of bacteriologically confirmed pulmonary tuberculosis in South Africa, 2017-19: a multistage, cluster-based, cross-sectional survey. *Lancet Infect Dis* 2022; **22:** 1172–80.

23 Aung ST, Nyunt WW, Moe MM, Aung HL, Lwin T. The fourth national tuberculosis prevalence survey in Myanmar. *PLoS Glob Public Health* 2022; **2:** e0000588.

24 Matji R, Maama L, Roscigno G, et al. Policy and programmatic directions for the Lesotho tuberculosis programme: findings of the national tuberculosis prevalence survey, 2019. *PLoS One* 2023; **18:** e0273245.

25 WHO. High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. 2014. https://iris.who.int/bitstream/handle/10665/135617/WHO_HTM_TB_2014.18_eng.pdf?sequence=1 (accessed March 3, 2025).

26 Swaminathan S, Rekha B. Pediatric tuberculosis: global overview and challenges. *Clin Infect Dis* 2010; **50** (suppl 3)**:** S184–94.

27 Mohammed H, Oljira L, Roba KT, et al. Burden of tuberculosis and challenges related to screening and diagnosis in Ethiopia. *J Clin Tuberc Other Mycobact Dis* 2020; **19:** 100158.

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.

28 Telisinghe L, Ruperez M, Amofa-Sekyi M, et al. Does tuberculosis screening improve individual outcomes? A systematic review. *eClinicalMedicine* 2021; **40:** 101127.

29 Andom AT, Gilbert HN, Ndayizigiye M, et al. Understanding barriers to tuberculosis diagnosis and treatment completion in a low-resource setting: a mixed-methods study in the Kingdom of Lesotho. *PLoS One* 2023; **18:** e0285774.

30 Davis JL, Cattamanchi A, Cuevas LE, Hopewell PC, Steingart KR. Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis. *Lancet Infect Dis* 2013; **13:** 147–54.

31 Cuevas LE, Santos VS, Lima SVMA, et al. Systematic review of pooling sputum as an efficient method for Xpert MTB/RIF tuberculosis testing during the COVID-19 pandemic. *Emerg Infect Dis* 2021; **27:** 719–27.

32 WHO. Diagnosis of tuberculosis and detection of drug-resistance rapid communication. 2024. https://iris.who.int/bitstream/handle/10665/379015/B09111-eng.pdf?sequence=1 (accessed Nov 18, 2024).

33 Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev* 2014; **2014:** CD009593. https://doi.org/10.1002/14651858.CD009593.pub3.

34 Alsdurf H, Empringham B, Miller C, Zwerling A. Tuberculosis screening costs and cost-effectiveness in high-risk groups: a systematic review. *BMC Infect Dis* 2021; **21:** 935.

Downloaded for Anonymous User (n/a) at University of the Witwatersrand Johannesburg from ClinicalKey.com by Elsevier on October 13, 2025. For personal use only. No other uses without permission. Copyright ©2025. Elsevier Inc. All rights reserved.