INT J TUBERC LUNG DIS 21(4):381–388

http://dx.doi.org/10.5588/ijtld.16.0432

Operational modelling: the mechanisms influencing TB diagnostic yield in an Xpert® MTB/RIF-based algorithm

R. Dunbar,* P. Naidoo,* N. Beyers,* I. Langley†

***Desmond Tutu TB Centre, Department of Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, South Africa; †Centre for Applied Health Research and Delivery, Liverpool School of Tropical Medicine, Liverpool, UK**

SUMMARY

SETTING: Cape Town, South Africa.

OBJECTIVE: To compare the diagnostic yield for smear/culture and Xpert® MTB/RIF algorithms and to investigate the mechanisms influencing tuberculosis (TB) yield.

METHOD: We developed and validated an operational model of the TB diagnostic process, first with the smear/culture algorithm and then with the Xpert algorithm. We modelled scenarios by varying TB prevalence, adherence to diagnostic algorithms and human immunodeficiency virus (HIV) status. This enabled direct comparisons of diagnostic yield in the two algorithms to be made.

RESULTS: Routine data showed that diagnostic yield had decreased over the period of the Xpert algorithm roll-out compared to the yield when the smear/culture algorithm was in place. However, modelling yield under identical conditions indicated a 13.3% increase in diagnostic yield from the Xpert algorithm compared to smear/culture. The model demonstrated that the extensive use of culture in the smear/culture algorithm and the decline in TB prevalence are the main factors contributing to not finding an increase in diagnostic yield in the routine data.

CONCLUSION: We demonstrate the benefits of an operational model to determine the effect of scale-up of a new diagnostic algorithm, and recommend that policy makers use operational modelling to make appropriate decisions before new diagnostic algorithms are scaled up.

KEY WORDS: TB diagnostic yield; modelling; simulation; TB diagnosis

TUBERCULOSIS (TB) remains a major cause of morbidity and mortality worldwide. Of the global estimated 10.4 million incident TB cases in 2015, 1.2 million were infected with the human immunodeficiency virus (HIV).[1] The Africa region accounted for 26% of global TB cases, 31% of whom are estimated to be HIV–co-infected. The main contributing factor driving the TB epidemic is ongoing transmission due to undiagnosed TB cases, diagnosed cases not initiating treatment[2–4] and diagnostic and treatment initiation delays.[5,6]

Increased investment in recent years has resulted in a number of new, more sensitive and rapid diagnostic tests for TB, with the expectation that this would lead to an increase in the number of cases diagnosed and earlier diagnosis and initiation of treatment, thus reducing transmission and, ultimately, the burden of disease. One of these tests, the Xpert® MTB/RIF assay (Cepheid, Sunnyvale, CA, USA), was endorsed by the World Health Organization (WHO) and recommended as the initial diagnostic test for those with suspected multidrug-resistant TB (MDR-TB) or HIV-associated pulmonary TB (PTB).[7] South Africa replaced smear microscopy with Xpert as the first test in the diagnostic algorithm for all presumptive PTB cases in 2011.[8]

The decision by policy makers about which new test to implement in a diagnostic algorithm can be complicated, and factors to be considered include the best combination of diagnostic tests, the resources required, who should be tested and the TB epidemiology in the setting (prevalence of TB, HIV co-infection and drug resistance). Often many of these factors are not known, and expensive and time-consuming clinical trials are required to make informed decisions,[9] or decisions are made without all the necessary information.

The variable results reported from studies evaluating the implementation of Xpert highlight the complexity in deciding if and how a new diagnostic test should be implemented within a diagnostic algorithm. For example, a population-level decision model estimated that full Xpert coverage would identify

Correspondence to: Rory Dunbar, Desmond Tutu TB Centre Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, P O Box 241, Francie van Zyl Avenue, Tygerberg 7505, South Africa. e-mail: rdun@sun.ac.za

Article submitted 9 June 2016. Final version accepted 1 December 2016.

30% more TB cases (with yield increasing from 15% to 19%) in South Africa in 2013 compared to smear and culture.[10] A prospective cluster-randomised trial of Xpert compared to smear microscopy and culture conducted in a primary care clinic in Cape Town, South Africa, showed an increase in TB yield from 17% with smear and culture to 26% with Xpert.[11] A study conducted in North Ethiopia among household contacts of TB index cases showed an increase of 64.3% in TB detection between smear microscopy (12.8%) and Xpert (35.9%).[12] However, other studies in South Africa and Zimbabwe have not found increases in TB yield.[13,14] Xpert is expensive to implement and use, and the health system and patient impacts and benefits under routine operational conditions are still uncertain.

Modelling as a framework to help with decision making is an attractive and viable option to guide policy makers in implementing new diagnostic tests and algorithms. Operational modelling could identify gaps within a health system and options for addressing these.[15] Projections of the impact of interventions on patient access and outcomes and health system costs and infrastructure could help guide policy makers on which new diagnostic tests and algorithms should be implemented.

As part of an evaluation of new TB diagnostics in South Africa (Policy Relevant Outcomes from Validating Evidence on ImpacT, PROVE IT), we developed an operational model using a discrete event simulation approach for the previous smear/culture-based TB diagnostic algorithm and the newly introduced Xpert-based algorithm in Cape Town and validated the model outputs by comparing these with routine TB programme data.[16]

We used the operational model to investigate the mechanisms influencing TB yield in our setting and to better understand why we did not find the expected increase in TB diagnostic yield in our own empirical study.[16] We used simulated model scenarios, including a decrease in TB prevalence, varying adherence to protocol in diagnostic algorithms and knowledge of HIV status to make direct comparisons of the proportion of presumptive cases diagnosed as TB (TB yield), missed cases (false-negatives) and unnecessarily treated cases (false-positives) in the smear/culture and Xpert-based algorithms.

METHODS

Setting

The model was developed (Appendix Tables A.1 and A.2* and Figures 1 and 2) and validated (Appendix Table A.3) using routine National Health Laboratory Service (NHLS) data collected for the period from 2010 to 2013 over seven time points (T1 to T7) in Cape Town,[16] one of the larger cities in South Africa, with a population of 3.7 million in 2011 (national census 2011) and 28 658 TB cases reported; 47% of TB cases tested were co-infected with HIV (source: routine TB programme data, Cape Town Health Directorate).

Municipal and provincial health authorities provided TB diagnostic services at 142 primary health care (PHC) facilities. Sputum samples collected for TB testing at PHC facilities were couriered to the central NHLS on a daily basis for testing, and results were returned via courier and fax.

Two diagnostic algorithms (Appendix Figure A.1) were used in the study period. A smear/culture-based TB algorithm was used before August 2011, with all presumptive cases required to submit two spot sputum samples taken at least 1 h apart. Both sputum samples were examined using fluorescence microscopy after being chemically treated, centrifuged and stained. Among previously treated presumptive cases, the second sample was cultured using BACTEC™ MGIT™ 960 (BD, Sparks, MD, USA) and tested for drug susceptibility using the GenoType® MTBDR*plus* (Hain LifeScience, Nehren, Germany) line-probe assay (LPA). For new presumptive cases who were smear-negative and HIV-infected, a third sample was required for culture.

An Xpert-based algorithm was phased in from August 2011 to February 2013, with Xpert replacing smear microscopy for all presumptive cases. The first of two sputum samples submitted was tested using Xpert. If TB was detected, smear microscopy was performed on the second sample. In HIV-infected cases with negative Xpert results, the second sample underwent culture and LPA. All definitions for terms used throughout the article are given in Table 1.

Model development

The Witness package, a discrete event and continuous process simulator,[17] was used to develop a comprehensive model to represent the diagnosis of PTB in Cape Town. The model incorporated TB diagnostic algorithms (Appendix Figure A.1) as well as patient pathways and sample flow (Appendix Figure A.2) from specimen collection through laboratory test procedures to a result being provided to the patient and treatment initiation.

Table 2 summarises the model validation, and a detailed account of model development (Appendix Tables A.1 and A.2), more details about validation (Appendix Table A.3) as well as model sensitivity analysis, are available in the online Appendix.

Simulated scenarios: comparing the smear/culture and Xpert-based algorithms

To determine why the expected increase in TB yield was not observed in our setting with the roll-out of the

* The appendix is available in the online version of this article, at http://www.ingentaconnect.com/content/iuatld/ijtld/2017/00000021/00000004/art00006

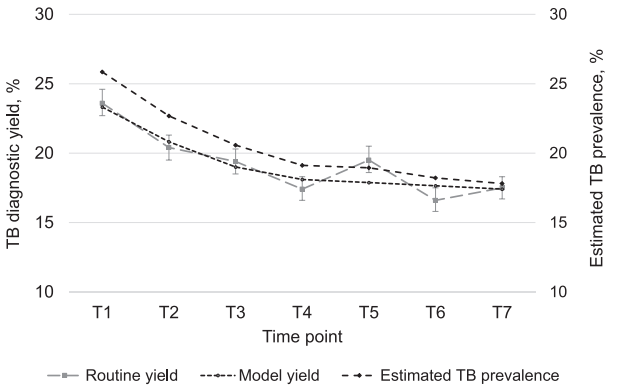

Figure 1 Model output comparing observed yield and model yield for all presumptive cases.

Xpert-based algorithm, we modelled both algorithms with identical input parameters to eliminate any differences in population characteristics (TB prevalence, HIV status, history of previous anti-tuberculosis treatment) during the time the smear/culture-based algorithm was in use and during the Xpert roll-out. We used estimated prevalence data from the most recent time point (T7) and an average of population parameters (proportion of previously treated cases and HIV status) over the seven time points. We ran the model for a period of 3 years for each simulation.

In the base-case scenario (Scenario A), we set levels of adherence to testing protocols in both algorithms at 85%, and assumed that 50% of presumptive cases knew their HIV status. Various other scenarios were modelled (Table 3) and, unless otherwise specified, the baseline parameters for Scenario A were maintained for all the following scenarios. In Scenario B, the estimated TB prevalence among presumptive TB cases was increased by 10%. In Scenario C, we increased the number of cultures for smear-negative and Xpert-negative presumptive TB cases to that found in routine practice: smear/culture-based algorithm: new HIV-negative 30% and HIV-positive 92%, previously treated HIV-negative 10% and HIV-positive 95%; Xpert-based algorithm: new HIV-negative 5% and HIV-positive 92%, previously treated HIV-negative 10% and HIV-positive 95%.

In Scenario D, we assessed the effect of an increase in known HIV status (from 50% to 85%) on outputs. In Scenario E, we tested the effect of 100% adherence to testing protocols in each algorithm, with 50% of presumptive TB cases' HIV status known. With Scenario F, we increased both HIV status known and adherence to testing protocols to 100%. With Scenario G, the use of culture was removed from both algorithms, as in most settings culture is not used as extensively as in Cape Town or is not used at all as part of the diagnostic algorithm. In Scenario H, we lowered the test sensitivity of smear microscopy by 10%. Scenario input parameters are summarised in Appendix Tables A.4 and A.5. For each scenario, we compared TB yield between the algorithms for all TB cases, and also assessed the proportion of missed cases (false-negative or TB not detected but TB present) and cases treated unnecessarily (false-positive or TB detected but no TB present). We undertook

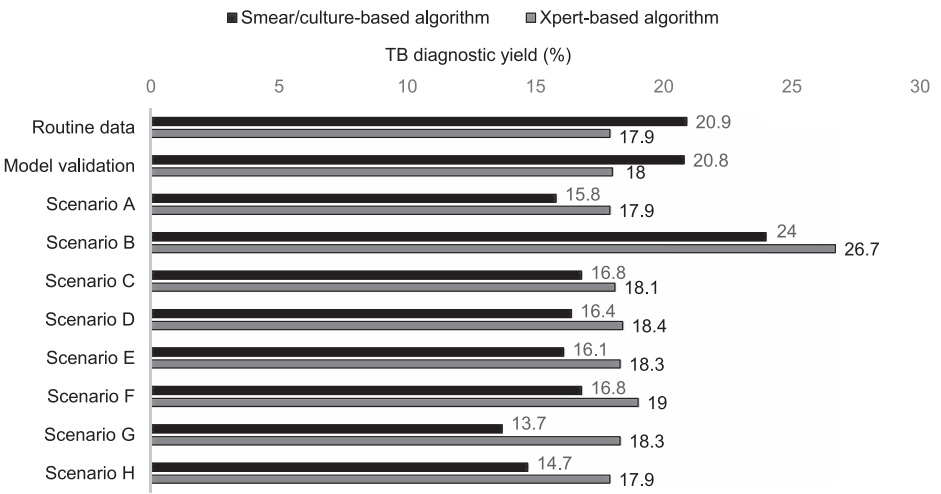

Figure 2 Model outputs from the scenarios comparing TB diagnostic yield (%) between algorithms and sensitivity of input parameters. Diagnostic yield from routine data. Model with routine data as input parameters. Scenario A: 85% adherence to algorithm and 50% of presumptive cases know their HIV status; Scenario B: increase estimated TB prevalence among presumptive cases by 10%; Scenario C: increase additional culture for smear or Xpert-negative presumptive cases per routine practice; Scenario D: increased proportion (85%) of presumptive cases know their HIV status; Scenario E: adherence (100%) to algorithms; Scenario F: increased proportion (100%) of presumptive cases know their HIV status and adherence (100%) to algorithms; Scenario G: remove culture as part of the sequence of tests required in each diagnostic algorithm; Scenario H: lower the sensitivity of smear microscopy by 10%. TB = tuberculosis.

Table 1 Definitions used throughout the study

Presumptive case	Defined as an individual with pre-treatment sputum samples submitted for diagnostic purposes
TB prevalence among presumptive cases	The proportion of true TB cases among presumptive cases. For model purposes, this is defined as culture-positive cases
TB case	An individual with one or more smears positive and/or culture positive for *M. tuberculosis* and/or *M. tuberculosis* detected on Xpert (includes true-positive cases and false-positive cases)
New presumptive cases	An individual with no previous anti-tuberculosis treatment or <4 weeks of previous anti-tuberculosis treatment
Previously treated presumptive cases	An individual with >4 weeks of previous anti-tuberculosis treatment
TB testing protocol	The sequence of tests required in each diagnostic algorithm
TB diagnostic yield	The number of TB cases diagnosed (based on the full TB testing protocol performed) expressed as a proportion of presumptive cases tested TB diagnostic yield = (True positive + false positive)/Presumptive cases
False-positive	The proportion of individuals with culture-negative TB who are incorrectly diagnosed with TB
False-negative	The proportion of individuals with culture-positive TB in whom a TB diagnosis is missed
True-positive	The proportion of individuals with culture-positive TB who are diagnosed with TB

TB = tuberculosis.

the analysis for new and previously treated TB cases, and for HIV-positive and HIV-negative TB cases (Appendix Table A.6–A.13).

Analysis

All patient and laboratory test information was written to a Microsoft SQL Server database (Microsoft, Redmond, WA, USA). Model outputs were aggregated by month over a 3-year period to produce means and 95% confidence intervals (CIs) for diagnostic yield and proportions that were false-negative and false-positive. We used the *t*-test to determine differences in means between observed and modelled outputs for the model validation and between algorithms for simulated scenarios. Differences are expressed as absolute values. All analyses were undertaken using STATA version 14 (StataCorp, College Station, TX, USA).

Ethics statement

The study was approved by the Health Research Ethics Committee at Stellenbosch University, Tygerberg, South Africa (IRB0005239) (N10/09/308) and the Ethics Advisory Group at the International Union

Table 2 Model validation comparing TB yield from routine data and model outputs (%)

	T1	T2	T3	T4	T5	T6	T7	Weighted mean ± SD	Mean difference % (95%CI)	*P* value
All presumptive cases										
Smear/culture-based algorithm										
Routine data	23.6	20.4	18.0	18.8	20.6	NA*	NA*	20.9 ± 2.1	0.1 (−3.0 to 3.3)	
Model outputs	23.4	20.8	18.7	17.3	17.4	NA*	NA*	20.8 ± 2.2		0.928
Xpert-based algorithm										
Routine data	NA*	NA*	21.2	16.9	19.3	16.6	17.5	17.9 ± 1.5	−0.1 (−1.8 to 1.6)	
Model outputs	NA*	NA*	19.5	18.5	18.1	17.7	17.4	18.0 ± 0.6		0.874
Overall										
Routine data	23.6	20.4	19.4	17.4	19.5	16.6	17.5	19.2 ± 2.2	0.0 (−2.4 to 2.5)	
Model outputs	23.4	20.8	19.1	18.1	17.9	17.7	17.4	19.2 ± 2.0		0.988
New presumptive cases										
Smear/culture-based algorithm										
Routine data	23.2	20.9	17.0	16.8	19.7	NA*	NA*	20.4 ± 2.5	0.8 (−2.6 to 4.3)	
Model outputs	22.1	19.8	17.6	16.3	16.0	NA*	NA*	19.6 ± 2.1		0.581
Xpert-based algorithm										
Routine data	NA*	NA*	21.7	15.7	18.7	15.8	16.2	17.1 ± 1.9	0.0 (−2.1 to 2.0)	
Model outputs	NA*	NA*	18.3	17.5	17.0	16.7	16.8	17.1 ± 0.5		0.975
Overall										
Routine data	23.2	20.9	18.8	16.1	18.9	15.8	16.2	18.5 ± 2.5	0.4 (−2.3 to 3.0)	
Model outputs	22.1	19.8	17.9	17.1	16.9	16.7	16.8	18.1 ± 1.9		0.764
Previously treated presumptive cases										
Smear/culture-based algorithm										
Routine data	26.6	23.1	21.6	25.3	23.5	NA*	NA*	24.4 ± 1.9	0.2 (−2.7 to 3.1)	
Model outputs	26.5	24.1	22.1	20.3	21.3	NA*	NA*	24.2 ± 2.1		0.897
Xpert-based algorithm										
Routine data	NA*	NA*	23.6	21.7	22.6	19.8	24.2	22.1 ± 1.7	0.9 (−0.9 to 2.9)	
Model outputs	NA*	NA*	23.2	21.4	21.1	20.9	20.2	21.2 ± 0.9		0.276
Overall										
Routine data	26.6	23.1	22.5	22.8	22.7	19.8	24.2	23.2 ± 2.0	0.7 (−1.7 to 3.1)	
Model outputs	26.5	24.1	22.6	21.1	21.2	20.9	20.2	22.4 ± 2.1		0.535

*The smear/culture-based algorithm was not in use in the TB programme during T6 and T7; and the Xpert-based algorithm was not in use during T1 and T2.
TB = tuberculosis; SD = standard deviation; CI = confidence interval; NA = not applicable.

Table 3 Model outputs from the scenarios comparing TB diagnostic yield (%) between algorithms and sensitivity of input parameters

	Smear/culture-based algorithm			Xpert-based algorithm			Change in yield between algorithms (relative % difference)
Routine and modelled data across all time periods (i.e., smear/culture T1–T5 and Xpert T3–T7)							
Diagnostic yield from routine data		20.9			17.9		−3.0 (−14.4)
Model with routine data as input parameters		20.8			18.0		−2.8 (−13.5)
Modelled scenarios: all input parameters identical between algorithms	True TB	Not TB	Yield	True TB	Not TB	Yield	
Scenario A: 85% adherence to algorithm and 50% of presumptive cases know their HIV status*							
TB detected	15.0	0.8	15.8	16.3	1.6	17.9	2.1 (13.3)
TB not detected	3.3	80.9		2.1	80.0		
Scenario B: increase estimated TB prevalence among presumptive cases by 10%							
TB detected	23.3	0.7	24.0	25.3	1.5	26.7	2.7 (11.3)
TB not detected	5.1	70.9		3.1	70.2		
Scenario C: increase additional culture testing for smear or Xpert-negative presumptive cases to that found in routine practice†							
TB detected	16.0	0.8	16.8	16.4	1.7	18.1	1.3 (7.7)
TB not detected	2.3	80.9		1.9	80.0		
Scenario D: increased proportion (85%) of presumptive cases know their HIV status							
TB detected	15.6	0.8	16.4	16.7	1.7	18.4	2.0 (12.2)
TB not detected	2.6	81.0		1.5	80.0		
Scenario E: 100% adherence to algorithms; 50% know their HIV status							
TB detected	15.3	0.8	16.1	16.5	1.8	18.3	2.2 (13.7)
TB not detected	3.0	80.9		1.8	79.9		
Scenario F: 100% adherence to algorithms; 100% know their HIV status							
TB detected	16.1	0.7	16.8	17.1	1.9	19.0	2.2 (13.1)
TB not detected	2.1	81.1		1.1	79.9		
Scenario G: remove culture as part of the sequence of tests required in each diagnostic algorithm							
TB detected	12.9	0.8	13.7	16.5	1.8	18.3	4.6 (33.6)
TB not detected	5.4	80.9		1.8	80.0		
Scenario H: lower the sensitivity of smear microscopy by 10%							
TB detected	13.9	0.8	14.7	16.3	1.6	17.9	3.2 (21.8)
TB not detected	4.4	80.9		2.1	80.0		

* In Scenario A, 85% of cases in each algorithm received the initial tests as required and 85% of smear- or Xpert-negative cases who were HIV-infected underwent culture; 50% of presumptive cases knew their HIV status. The same values for TB prevalence (18.8%), proportions of HIV−, HIV+ (status known and undiagnosed), new and previously treated cases were used in each algorithm. All scenario changes are in relation to Scenario A. Yield = TB detected (true TB + not TB).
† In Scenario C, culture in smear/culture-based algorithm increased new HIV− (0–30%), HIV+ (85–92%), previously treated HIV− (0–10%) HIV+ (85–95%); Xpert-based algorithm: new HIV− (0–5%), HIV+ (85–92%), previously treated HIV− (0–10%), HIV+ (85–95%).
TB = tuberculosis; HIV = human immunodeficiency virus; − = negative; + = positive.

Against Tuberculosis and Lung Disease, Paris, France (59/10). A waiver for informed consent was granted for use of routine data. The City Health Directorate, Western Cape Health Department, Cape Town, and the National Health Laboratory Services, Pretoria, South Africa, granted permission to use routine health data.

RESULTS

Model validation

The mean differences between observed yield from routine data and model outputs over the seven time points is shown in Figure 1 and Table 2. The TB yield from the model closely approximated that from routine data for both diagnostic algorithms. The mean model yield was 0.1% ($P = 0.928$) lower than observed values in the smear/culture-based algorithm and 0.1% ($P = 0.874$) higher in the Xpert-based algorithm overall (Figure 1; Table 2).

Simulated scenarios: comparing the smear/culture and Xpert-based algorithms

Figure 2 and Table 3 summarise the model outputs from the scenarios comparing the TB diagnostic yield from the smear/culture and Xpert-based algorithms. In Scenario A (detail in Appendix Table A.6), with 85% adherence to the diagnostic algorithms and where 50% of presumptive cases knew their HIV status, the overall TB diagnostic yield was 15.8% in the smear/culture-based algorithm compared to 17.9% in the Xpert-based algorithm (relative difference 13.3%), with respectively 3.3% and 2.1% of presumptive cases having a missed TB diagnosis. A lower proportion of cases were falsely diagnosed with TB in the smear/culture-based algorithm (0.8%) than in the Xpert-based algorithm (1.6%).

When the estimated TB prevalence among presumptive cases was increased by 10% (absolute) (Scenario B) (Appendix Table A.7), the yield was 24.0% and 26.7% in the respective algorithms. The relative increase in yield between algorithms was 11.3%. The proportion of missed cases was respectively 5.1% and 3.1%.

When Scenario A was adjusted so that the proportions of smear-negative and Xpert-negative cases who received a culture test were set to reflect the values found in routine practice (Scenario C), the

overall yield was 16.8% in the smear/culture-based algorithm compared to 18.1% in the Xpert-based algorithm, a relative increase of 7.7%. The proportion of missed cases was 2.3% and 1.9% in the respective algorithms (Appendix Table A.8). In comparison to Scenario A, the relative increase in TB yield was 6.3% in the smear/culture-based algorithm and 1.1% in the Xpert-based algorithm in Scenario C.

If HIV testing among presumptive TB cases was increased and 85% knew their HIV status (Scenario D) (Appendix Table A.9), the overall yield was 16.4% in the smear/culture-based algorithm compared to 18.4% in the Xpert-based algorithm (relative difference 12.2%), with 2.6% and 1.5% missed cases in the respective algorithms. In comparison to Scenario A, an increase in HIV testing resulted in a relative increase of 3.8% and 2.8% in TB yield in the respective algorithms.

If adherence to the testing protocol in each algorithm was increased to 100% but only 50% of presumptive TB cases knew their HIV status (Scenario E) (Appendix Table A.10), the TB yield was 16.1% in the smear/culture-based compared to 18.3% in the Xpert-based algorithm, with a relative increase of 10.9%. The proportion of missed cases was respectively 3.0% and 1.8%. In comparison to Scenario A, the relative increase in TB yield was 1.9% in the smear/culture-based algorithm compared to 2.2% in the Xpert-based algorithm.

If adherence to the testing protocol in each algorithm was increased to 100% and 100% of presumptive TB cases knew their HIV status (Scenario F) (Appendix Table A.11), the TB yield was 16.8% in the smear/culture-based compared to 19.0% in the Xpert-based algorithm, with a relative increase of 13.1% between algorithms. The proportion of missed cases was respectively 2.1% and 1.1%. In comparison to Scenario A, the relative increase in TB yield was 6.3% in the smear/culture-based algorithm compared to 6.1% in the Xpert-based algorithm.

Removing culture as part of the testing protocol from both algorithms (Scenario G) (Appendix Table A.12) resulted in a TB yield of 13.7% and 18.3% in the respective algorithms. The relative increase in yield between algorithms was 33.6%. The proportion of missed cases was 5.4% in the smear/culture algorithm compared to 1.8% in the Xpert algorithm.

If we assume the sensitivity of smear microscopy to be 10% lower than that estimated in our model (Scenario H) (Appendix Table A.13), the yield in the smear/culture algorithm would be 14.7% (missed cases 4.4%), with a relative increase in yield with the Xpert algorithm of 21.8%.

DISCUSSION

A strength of this study was the availability of detailed routine data and information collected on health and laboratory processes, which allowed us to develop a precise operational model to assess the impact of different diagnostic algorithms in Cape Town. The model input parameters were mostly based on these detailed routine data, and only a few assumptions were made. We assumed that prevalence among presumptive cases was higher among HIV-positive presumptive cases than among HIV-negative cases,[18,19] and among previously treated than among new cases;[20] we assumed a decrease in TB prevalence among presumptive cases over time based on the empiric yield data, which showed a decrease in yield over time despite similar proportions of the population being tested.[16] The latter assumption is supported by national data that showed a decrease in the number of laboratory-confirmed cases since 2011 (nationally and across the Western Cape Province).[21]

The availability of routine TB NHLS data collected through the PROVE IT study allowed us to validate the model by comparing TB yield observed in routine practice to model outputs using input parameters from seven different time points during the period when PHC facilities changed from the smear/culture algorithm to the Xpert algorithm. This comparison built confidence in the outputs from the model and confirmed that the outputs were credible. Overall model outputs closely resembled the TB yield observed in Cape Town over the seven time points, with a mean difference of 0.1% ($P = 0.951$) between routine data and the model outputs.

A direct comparison of TB yield in the Xpert and smear/culture-based algorithms in routine practice is difficult due to the variability in the population characteristic at each time point and different levels of adherence to testing protocols. When the Xpert-based algorithm was newly introduced, it took staff a period of time to adapt their clinical practice and become familiar with the new protocols. The global stock-out of the Xpert test during the study period also played a role in the extent to which testing protocols were followed. The operational model allows a direct comparison between the two algorithms with identical population characteristics and adherence to testing protocols. To understand the mechanisms and the extent to which they influenced TB yield in our setting, we used the validated model to compare various scenarios.

In Scenario A, with 85% adherence to algorithm and 50% of presumptive cases knowing their HIV status, the yield in the Xpert-based algorithm was higher than in the smear/culture-based algorithm, with a relative increase of 13.3%. Although the TB diagnostic yield was higher in the Xpert-based algorithm, the increase was lower than the predicted

increase with full roll-out of Xpert in South Africa[10] and reported by other studies.[11,12]

Scenarios B and C provide insights into the findings on TB yield from our empirical study. The TB yield in the smear/culture-based algorithm in Scenario B, where TB prevalence was 10% higher than in Scenario A, was 25% higher than in the Xpert-based algorithm in Scenario A, demonstrating the impact of a decline in prevalence among presumptive TB cases on TB yield. This helps explain findings from the empirical study that reported yields of 20.9% in the smear/culture-based and 17.9% in the Xpert-based algorithm.[16] It is likely that the change in prevalence in our setting during the study period was lower than the 10% value tested in the model.

Our model showed the impact of additional culture testing on reducing the difference in TB yield between algorithms. When the proportions of smear- and Xpert-negative cases who received culture tests were increased to reflect those found in routine practice (Scenario C), the relative increase in yield between the smear/culture and Xpert-based algorithms was reduced to 7.7%. This was attributable to a higher proportion of smear-negative than Xpert-negative cases undergoing culture. A cluster-randomised study in four other provinces also found that culture was more likely to be undertaken for smear-negative (32%) than Xpert-negative (14%) HIV-infected cases.[22] It was proposed that a greater belief in the efficacy of the Xpert test contributed to this.

Scenarios D to F provide insights into the potential benefits of interventions that strengthen the health system. Increasing the proportion of presumptive cases who knew their HIV status from 50% to 85% had a small influence on TB yield: the yield increased by 3.8% in the smear/culture-based algorithm, and by 2.8% in the Xpert-based algorithm in relative terms from Scenario A. It is interesting to note the modest benefits, considering the effort required to increase HIV testing of presumptive cases to this extent.

In Scenario E for the Xpert-based algorithm, increasing adherence to the algorithm to 100%, but with only 50% of presumptive TB cases knowing their HIV status, produced a 2.2% relative increase in yield compared to Scenario A. Increasing adherence to 100% and with 100% knowing their HIV status resulted in a relative increase in yield of only 6.1% in the Xpert-based algorithm in Scenario F compared to Scenario A. In addition to this disappointingly small benefit, 100% adherence is not realistic in routine practice due to the failure to request the correct test (due to new or locum staff who are unfamiliar with Xpert and costs concerns, for example), the availability of the Xpert test due to maintenance on Xpert machines or stock-outs and clinical decisions overriding the use of the testing algorithm. It is important to note that we started on a baseline of 85% adherence to the algorithms and with 50% knowing their HIV status; the increases in yield would be greater if these baseline values were lower. In a future study, we will model the effect of a more sensitive test than Xpert to assess the extent to which this can increase TB diagnostic yield.

We compared two scenarios that are pertinent to other settings. In Scenario G, where culture was removed from the algorithms, there was a 33.6% relative difference in TB yield between the smear/culture and the Xpert-based algorithm. The diagnostic benefits of Xpert are thus likely to be greater in areas that do not use or have very limited use of culture. The performance of smear microscopy in our central laboratory may also be much higher than reported by peripheral microscopy units. This is possibly due to greater proficiency and technical aspects in the central laboratory. It has been shown that an increase in yield for smear microscopy could be achieved by chemical treatment, centrifugation and fluorescence microscopy,[23–25] as used in our laboratory. If we did not have these benefits and smear microscopy sensitivity was 10% lower, the relative increase in TB yield in the Xpert-based algorithm would have been 21.8% compared to that in the smear/culture-based algorithm.

Limitations

We did not have data on TB prevalence for presumptive cases. This was estimated with a range of values tested in the model. The model was validated against data from Cape Town, a well-resourced urban setting where there is extensive use of culture in both algorithms. This may limit the generalisability of the findings to other settings.

As complete data are rarely available for any modelling study, assumptions are required for some input parameters. A model is also, by definition, a simplification of real-life processes. In this study, our model was validated by running the model with input parameters based on routine data for seven individual time points and comparing the output from the model against corresponding routine data (Figure 1; Table 2). Cost implications, treatment delay and rifampicin resistance were not addressed in the current study and will be reported in future studies.

CONCLUSION

We have developed and validated an operational model that can be used to directly compare the TB diagnostic yield between different algorithms, i.e., a smear/culture-based and an Xpert-based algorithm. Our model accounted for the variability found in routine practice and made it possible to eliminate the effect of a difference in population characteristics and adherence to testing protocols within algorithms on the TB diagnostic yield.

We were able to show that extensive use of culture in the smear/culture-based algorithm and decline in TB prevalence are the main factors likely to have contributed to our not finding an increased TB yield in the Xpert-based algorithm in our empiric study. The Xpert-based algorithm is likely to yield greater diagnostic benefits in areas without culture or with less sensitive TB microscopy.

We have demonstrated the benefits of using an operational model to determine the effect of scaling up a new diagnostic algorithm and investigate the mechanistic reasons that influence the yield of a new TB diagnostic algorithm. We would therefore recommend that policy makers use operational modelling to make appropriate decisions before new diagnostic algorithms are scaled up. The model could provide evidence as to how the greatest benefits could be obtained by using a new diagnostic test within a TB diagnostic algorithm and in a specific setting.

Acknowledgements

The authors wish to thank the National Health Laboratory Services (Pretoria), Cape Town Health Directorate and Western Cape Provincial Department of Health (Cape Town); and the City of Cape Town Health Directorate, Cape Town, South Africa, for providing permission for this data to be used.

Conflicts of interest: none declared.

References

1 World Health Organization. WHO End TB Strategy. Geneva, Switzerland: WHO, 2015. http://www.who.int/tb/post2015_strategy/en/ Accessed December 2016.
2 Botha E, Den Boon S, Verver S, et al. Initial default from tuberculosis treatment: how often does it happen and what are the reasons? Int J Tuberc Lung Dis 2008; 12: 820–823.
3 Claassens M M, du Toit E, Dunbar R, et al. Tuberculosis patients in primary care do not start treatment. What role do health system delays play? Int J Tuberc Lung Dis 2013; 17: 603–607.
4 Claassens M M, Jacobs E, Cyster E, et al. Tuberculosis cases missed in primary health care facilities: should we redefine case finding? Int J Tuberc Lung Dis 2013; 17: 608–614.
5 Sreeramareddy C T, Qin Z Z, Satyanarayana S, Subbaraman R, Pai M. Delays in diagnosis and treatment of pulmonary tuberculosis in India: a systematic review. Int J Tuberc Lung Dis 2014; 18: 255–266.
6 Nasehi M, Hassanzadeh J, Rezaianzadeh A, Zeigami B, Tabatabaee H, Ghaderi E. Diagnosis delay in smear positive tuberculosis patients. J Res Med Sci 2012; 17: 1001–1004.
7 World Health Organization. Rapid implementation of the Xpert MTB/RIF diagnostic test. WHO/HTM/TB/2011.2. Geneva, Switzerland: WHO, 2011. http://whqlibdoc.who.int/publications/2011/9789241501569_eng.pdf?ua=1 Accessed December 2016.
8 Scott L E, McCarthy K, Gous N, et al. Comparison of Xpert MTB/RIF with other nucleic acid technologies for diagnosing pulmonary tuberculosis in a high HIV prevalence setting: a prospective study. PLOS MED 2011; 8: e1001061.
9 Lessells R J, Cooke G S, Newell M-L, Godfrey-Faussett P. Evaluation of tuberculosis diagnostics: establishing an evidence base around the public health impact. J Infect Dis 2011; 204(suppl 4): S1187–S1195.
10 Meyer-Rath G, Schnippel K, Long L, et al. The impact and cost of scaling up GeneXpert MTB/RIF in South Africa. PLOS ONE 2012; 7: e36966.
11 Cox H S, Mbhele S, Mohess N, et al. Impact of Xpert MTB/RIF for TB diagnosis in a primary care clinic with high TB and HIV prevalence in South Africa: a pragmatic randomised trial. PLOS MED 2014; 11: e1001760.
12 Habte D, Melese M, Hiruy N, et al. The additional yield of GeneXpert MTB/RIF test in the diagnosis of pulmonary tuberculosis among household contacts of smear positive TB cases. Int J Infect Dis 2016; 49: 179–184.
13 Van Den Handel T, Hampton K H, Sanne I, Stevens W, Crous R, Van Rie A. The impact of Xpert® MTB/RIF in sparsely populated rural settings. Int J Tuberc Lung Dis 2015; 19: 392–398.
14 Mupfumi L, Makamure B, Chirehwa M, et al. Impact of Xpert MTB/RIF on antiretroviral therapy-associated tuberculosis and mortality: a pragmatic randomized controlled trial. Open Forum Infect Dis 2014; 1: ofu038.
15 Lin H-H, Langley I, Mwenda R, et al. A modelling framework to support the selection and implementation of new tuberculosis diagnostic tools. Int J Tuberc Lung Dis 2011; 15: 996–1004.
16 Naidoo P, Dunbar R, Lombard C, et al. Comparing tuberculosis diagnostic yield in smear/culture and Xpert® MTB/RIF-based algorithms using a non-randomised stepped-wedge design. PLOS ONE 2016; 11: e0150487.
17 Lanner. WITNESS Service and Process Performance Edition. Redditch, UK. http://www.lanner.com/technology/witness-simulation-software.html Accessed February 2017
18 Middelkoop K, Bekker L-G, Myer L, et al. Antiretroviral program associated with reduction in untreated prevalent tuberculosis in a South African township. Am J Respir Crit Care Med 2010; 182: 1080–1085.
19 Wood R, Middelkoop K, Myer L, et al. Undiagnosed tuberculosis in a community with high HIV prevalence: implications for tuberculosis control. Am J Respir Crit Care Med 2007; 175: 87–93.
20 den Boon S, van Lill S W P, Borgdorff M W, et al. High prevalence of tuberculosis in previously treated patients, Cape Town, South Africa. Emerg Infect Dis 2007; 13: 1189–1194.
21 Nanoo A, Izu A, Ismail N A, et al. Nationwide and regional incidence of microbiologically confirmed pulmonary tuberculosis in South Africa, 2004–12: a time series analysis. Lancet Infect Dis 2015; 15: 1066–1076.
22 McCarthy K M, Grant A D, Chihota V, et al. Implementation and operational research: what happens after a negative test for tuberculosis? Evaluating adherence to TB diagnostic algorithms in South African primary health clinics. J Acquir Immune Defic Syndr 2016; 71: e119–126.
23 Mase S R, Ramsay A, Ng V, et al. Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: a systematic review. Int J Tuberc Lung Dis 2007; 11: 485–495.
24 Steingart K R, Henry M, Ng V, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis 2006; 6: 570–581.
25 Steingart K R, Ng V, Henry M, et al. Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis 2006; 6: 664–674.

APPENDIX

Modelling as a framework to help with decision making is an attractive and viable option to guide policy makers in implementing new diagnostic tools and diagnostic algorithms. Two modelling approaches previously used in tuberculosis (TB) control are transmission (epidemiological) modelling and operational modelling.[1] Transmission modelling can be used to predict the long-term impact of interventions on the community by projecting TB incidence, prevalence and mortality. Operational modelling, on the other hand, can be used to project the impact of interventions on health system costs and infrastructure, as well as patient access and outcomes. Operational modelling can also be useful in identifying gaps within a health system and to identify ways to address the gaps within the health system.[2] An operational model is a simplified representation of complex real-life processes. The data sources usually used to drive operational models are derived from published literature (i.e., meta-analyses, randomised control trials, cohort studies, global reports, unpublished literature, expert opinion, field data), and from assumptions. Models are therefore only as good as the level of detail available to develop the logic of the model and the availability and accuracy of the data to drive the model.[3]

In industrial and commercial settings, operational models are widely used to plan and assess the performance and efficiency of processes.[4] Operational models have also increasingly been used to improve performance of the health sector.[5,6] The use of operational models in health systems is common in high-income countries, but not in middle- to low-income countries at this stage. Many operational models use a discrete event simulation approach, where the system modelled is first defined in terms of its most important elements, including the items or people processed through the facility, resources, activities, rules and the process flow. The required outputs of the model are defined (e.g., productivity, costs, identification of bottlenecks, capacity and sensitivity to changes), along with the key input parameters to be investigated. Once the system is defined and appropriate parameter inputs are assigned (e.g., the number of items entering the system, the quantity of resources and the time for completing particular activities), then simulations can be run to assess the relative effect of different input assumptions on the modelled outputs.[7]

Model development

The Witness package, a discrete event and continuous process simulator,[8] was used to develop a comprehensive model to represent the diagnosis of pulmonary TB (PTB) in Cape Town, South Africa. The model incorporated the TB diagnostic algorithms (Figure A.1), as well as patient pathways and sample flow (Figure A.2) from specimen collection, through laboratory test procedures, to a result being provided to the patient and treatment initiation.

The main elements in the model (Table A.1) were entities (representing patients, sputum samples), activities (representing patient reception, sputum collection from the patient, sample transport, sample registration at the laboratory, sample preparation and test procedures, review and return of results to primary health care [PHC] facilities), queues (representing delays before each activity, e.g., patient waiting in reception prior to clinical evaluation and sputum collection, batching and other processes in

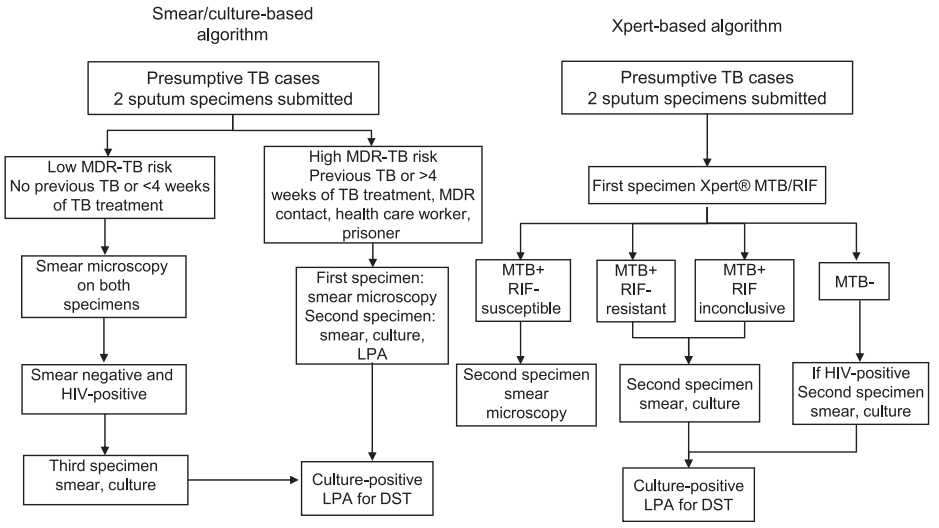

Figure A.1 TB diagnostic algorithms. The simplified sequence of diagnostic tests in each algorithm and the action taken based on test results is shown. TB = tuberculosis; MDR-TB = multidrug-resistant TB; LPA = line-probe assay; MTB = *Mycobacterium tuberculosis*; RIF = rifampicin; HIV = human immunodeficiency virus; DST = drug susceptibility testing.

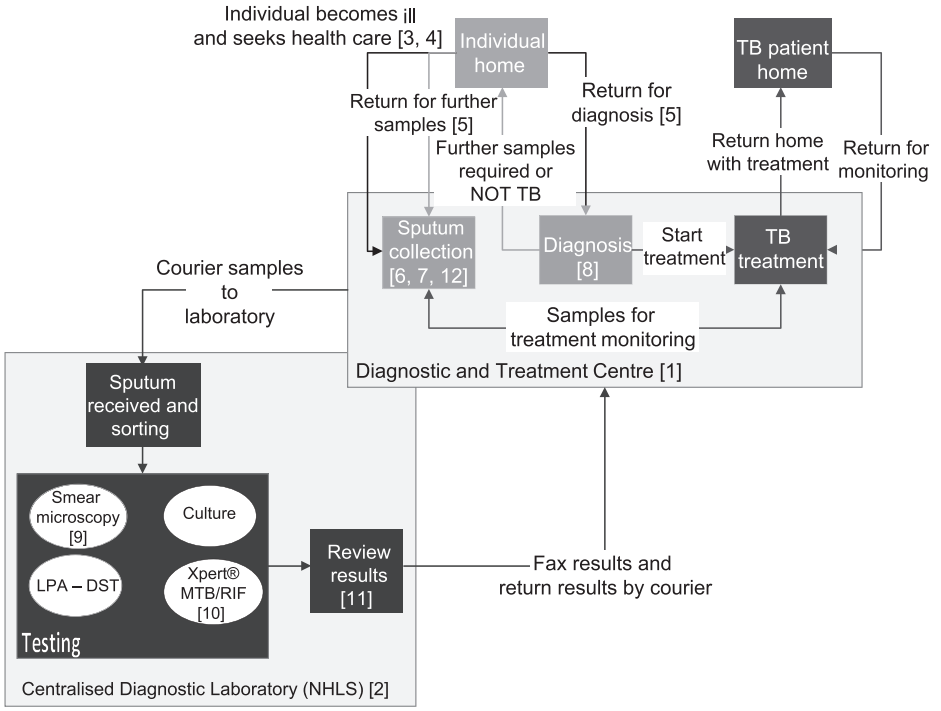

Figure A.2 A representation of the diagnostic pathway for the diagnosis of TB in Cape Town, South Africa.* * See Table A.2. TB = tuberculosis; LPA = lineprobe assay; DST = drug susceptibility testing; NHLS = National Health Laboratory Services.

the laboratory) and resources (representing health facility and laboratory staff and equipment).

The model structure follows the patient and sample pathways (Figure A.2), with the flow of entities between activities and queues dictated by rules, which are dependent on attributes for entities (patient or sample).

Input parameters

A detailed list of the input parameters, including processing times, patient and sample data as well as test sensitivity and specificity, is shown in Table A.2.

Input data sources

Health system and laboratory processes were mapped in three ways: through key informant interviews, through the review of standard operating procedures and through detailed observation and timing of clinic and laboratory processes, all undertaken as part of the PROVE IT study.

Characteristics of presumptive cases were derived from electronic laboratory TB test data received from the National Health Laboratory Services (NHLS) Data Warehouse. Data included demographic information, treatment history (new or previously treated cases) human immunodeficiency virus (HIV) status (HIV-negative, HIV-positive, status unknown), test type (smear, culture or Xpert® MTB/RIF), test results and date when sputum was collected, tested and results available.

As there were no unique identifiers in the NHLS data to link results belonging to an individual, matching was performed on personal identifying information (first name, family name, date of birth and facility folder number). After record matching, all personal identifying information was removed. These data were used to assess adherence to testing protocols in each algorithm and to calculate diagnostic yield (the proportion of presumptive cases diagnosed as TB).

Data on the sensitivity and specificity of tests were obtained from systematic reviews and published literature.[9–11] We did not have data on the proportion of true TB cases among presumptive cases (TB prevalence), and we tested a range of prevalence values, assuming that prevalence was lower among HIV-negative than among HIV-positive,[12,13] and among new than among previously treated presumptive cases[14]; the outputs in comparison to observed TB yield values from routine data were then assessed (see below).[15]

Model outputs

The output from the model indicated the proportion of cases diagnosed as TB (TB yield, i.e., true-positive and false-positive cases) and the proportion of cases missed (false-negative) using the diagnostic algorithm in different scenarios.

Model calibration and validation

To have confidence in the model and the outputs produced, the model was verified and validated.[16]

Table A.1 The key elements used in the model developed for this study

Model element type	Representation of
Entities	Patients, presumptive cases reporting to a TB diagnostic centre (clinic) for TB diagnosis
	Sputum sample
Attributes	Number of sputum samples required
	Number of sputum samples collected
	Previous history of anti-tuberculosis treatment (new, previous TB)
	HIV status (HIV−, HIV+, HIV+ with status not known)
	Test result
Activities	Clinic
	Reception
	TB room
	Sputum collection
	Patient return home
	Courier, transport of samples from clinic to central laboratory
	Laboratory
	Sample sorting/reception
	Smear
	Preparation
	Microscopy
	Review result
	Fax result
	Culture
	Preparation
	MGIT
	ZN smear
	Fax result
	Xpert
	Preparation
	Xpert test
	Review results
	Fax result
Queues	Clinic
	Waiting room at reception
	Waiting for sputum collection
	Patient
	Wait for result or provide further sputum samples
	Laboratory
	Sample waiting for preparation
	Sample waiting for testing/batching
	Microscopy
	Culture
	Xpert
	Result waiting for review
	Result waiting to be faxed

TB = tuberculosis; HIV = human immunodeficiency virus; − = negative; + = positive; MGIT = Mycobacteria Growth Indicator Tube; ZN = Ziehl-Neelsen.

Model verification to ensure that the coding and logic of the model and its execution were correct was performed through incremental model building and carefully scrutinising the structure and logic of the model at each stage. The distribution of input parameters was assessed against outputs to make sure that the model assigned patient categories correctly.

The model was validated using input parameters from routine data and by comparing TB yield from model outputs to routine data.[15] As part of the PROVE IT Study, NHLS data from presumptive cases had previously been collected and analysed to compare TB yield in the smear/culture-based algorithm to that in the Xpert-based algorithm over seven time periods (T1–T7), during which the PHC changed from the former algorithm to the latter.[15] The model used probability distributions derived from this analysis to assign patients to categories: diagnostic algorithm used, HIV status (known HIV-positive, undiagnosed HIV and HIV-negative) and treatment history (new or previously treated).

Data on HIV status were only available for time points T6 and T7 (50% knew their HIV status), and similar proportions were assumed for T1–T5. The extent to which testing protocols was followed in each diagnostic algorithm was derived from these data. As we did not have data on TB prevalence among presumptive cases, a range of values were tested. We made the assumption that prevalence among presumptive cases was higher among HIV-positive presumptive cases than among HIV-negative cases,[12,13] and among previously treated than among new cases.[14] We

Table A.2 Main input parameters for operational model*

Input parameter	Description	Values		Source
Processing times				
TB diagnostic and treatment facility processing times	Duration of activities within the diagnostic centre for reception, sputum collection and returning results	Not defined for current model		
Laboratory process times Duration of and batching process for microscopy, culture and drug susceptibility testing. Preparation and processing times for Xpert	Sample collection times by courier Sample sorting time Xpert preparation time Xpert test time Smear preparation time Microscopy reading Culture preparation Culture test time	First sputum delivery: between 2 and 3 pm 25 min in batches of 96 50 min in batches of 16 1 h 50 min 2 h for 96 samples 1 h per batch of 96 2 h per batch of 750 5–36 days		Interviews with NHLS staff, review of SOPs and direct observations of laboratory procedures
Patient data				
Number of new patients seeking diagnosis Number and arrival rate of individuals seeking diagnosis	Number of presumptive cases per day	Uniform distribution: Minimum = 11 Maximum = 690 Mean = 55		NHLS data warehouse and sampled from a uniform distribution
Arrival time of patients[4]	The time during the day that the patient arrives at the diagnostic centre	5 days: Monday to Friday 8 am to 5 pm Working 540 min per day		Sampled from user-defined distribution starting at the opening time of the health facility with all patients arriving by closing time.
Return probability[5]	The probability that an individual returns to the diagnostic centre for the next stage of the diagnostic process	Not defined for current model		Published literature[18,19]
HIV status of presumptive cases[6]	Proportion of presumptive cases who are identified as HIV+	New presumptive cases = 18% Previous history of TB = 35%		Estimated from NHLS Data Warehouse data for 2013
History of previous TB treatment for presumptive cases[7]	Proportion of presumptive cases with previous TB treatment	Average over all observed time points = 24% (breakdown by time point in Table A.3)		Estimated from NHLS Data Warehouse data and stepped wedge analysis[15]; 90% of presumptive cases with missing previous TB treatment status were assumed to be new cases
Proportion of presumptive cases with diagnostic test performed[8] Average over all observed time points (breakdown by time point in Table A.3)	Smear/culture algorithm: Two smears Smear-negative with culture Xpert algorithm: Xpert Xpert-negative with culture	New cases 84% 36% 77% 18%	Previous history of TB 82% 85% 65% 41%	Estimated from NHLS Data Warehouse data and stepped wedge analysis[15]
Diagnostic test accuracy				
Accuracy of smear microscopy[9]	Sensitivity and specificity of LED fluorescence microscopy	Two sputum samples HIV−: sensitivity 75%, specificity 99% HIV+: sensitivity 65%, specificity 99%		Published literature[9,10]
Accuracy of Xpert[10]	Sensitivity and specificity of Xpert in identifying TB from sputum samples	HIV−: sensitivity 89%, specificity 98% HIV+: sensitivity 80%, specificity 98%		Published literature[11]
Proportion of tests by test type that give no result[11]	Level of retesting required for smear, culture or Xpert	2%		Estimated from NHLS Data Warehouse data and stepped wedge analysis[15]

Table A.2 *(continued)*

Input parameter	Description	Values	Source
Number of sputum tests required per patient with suspected TB[12]	The number of sputum samples required for each diagnostic algorithm	Two sputum samples	South African National TB Guidelines[20]

* See Figure A.2.
TB = tuberculosis; NHLS = National Health Laboratory Services; SOPs = standard operating procedures; HIV = human immunodeficiency virus; + = positive; LED = light-emitting diode; − = negative.

assumed a decrease in TB prevalence among presumptive cases based on the routine yield data, which showed a decrease in yield over time despite a similar proportion of the population being tested.[15] This is supported by national data that show a reduction in the number of laboratory-confirmed cases since 2011 (nationally and across the Western Cape Province).[17] A summary of the population characteristics used for model validation by time period is provided in Table A.3.

The availability of TB test data collected through the PROVE IT study allowed us to validate the model by comparing TB yield observed in routine practice to model outputs using the input parameters from seven different time points during the period when PHC facilities changed from the smear/culture algorithm to

Table A.3 Population characteristics used for model validation by time period (%)

	T1*	T2*	T3*	T4*	T5*	T6*	T7*
History of previous anti-tuberculosis treatment							
Proportion of previously treated cases	29	24	24	25	25	24	19
HIV status							
New presumptive cases							
HIV-positive	36	36	36	36	36	36	36
HIV-negative	64	64	64	64	64	64	64
Previously treated presumptive cases							
HIV-positive	53	53	53	53	53	53	53
HIV-negative	47	47	47	47	47	47	47
Proportion who know their HIV status							
Estimated TB prevalence among presumptive cases							
New presumptive cases							
HIV-positive	26	23	21	19	19	18	17.8
HIV-negative	25	22	19	18	17.8	17	17
Previously treated presumptive cases							
HIV-positive	27	24	23	21.5	21.5	21	20.8
HIV-negative	26	23	22	20	19.8	19.8	19.5
Proportion of presumptive cases tested using algorithm							
Smear/culture-based	100	100	57	30	19	0	0
Xpert-based	0	0	43	70	81	100	100
Adherence to smear/culture-based algorithm							
New presumptive cases with two smears	85	85	85	85	85	—	—
Previously treated presumptive cases with culture	88	91	90	92	75	—	—
Adherence to Xpert-based algorithm							
All presumptive cases with Xpert test done	—	—	57	67	63	75	80
Proportion of patients who were smear or Xpert-negative with culture							
Smear/culture-based algorithm							
New presumptive cases							
HIV-positive	92	92	92	92	92	92	92
HIV-negative	30	30	30	30	30	30	30
Previously treated presumptive cases							
HIV-positive	95	95	95	95	95	95	95
HIV-negative	10	10	10	10	10	10	10
Xpert-based algorithm							
New presumptive cases							
HIV-positive	92	92	92	92	92	92	92
HIV-negative	2	2	2	2	2	2	2
Previously treated presumptive cases							
HIV-positive	95	95	95	95	95	95	95
HIV-negative	10	10	10	10	10	10	10

*T1–T7 reflect the time points evaluated as part of a non-randomised stepped-wedge evaluation of TB yield with a transition from a smear/culture to an Xpert-based algorithm in Cape Town.[15] At T1 and T2, all facilities used the smear/culture-based algorithm; this decreased to 65% of facilities at T3, 38% at T4 and 23% at T5. At T6 and T7, all facilities used the Xpert-based algorithm: T1, November 2010; T2, May 2011; T3, November 2011; T4, May 2012; T5, November 2012, T6, May 2013; T7, November 2013. Values are derived from routine data.[15]
HIV = human immunodeficiency virus; TB = tuberculosis.

Table A.4 Input parameters used in base-case (Scenario A) comparing the smear/culture and Xpert-based algorithms (%)*

History of previous anti-tuberculosis treatment	25
HIV status	
New presumptive cases	
HIV-positive	36
HIV-negative	64
Previously treated presumptive cases	
HIV-positive	53
HIV-negative	47
Proportion who know their HIV status	50
Estimated TB prevalence among presumptive cases	
New presumptive cases	
HIV-positive	17.8
HIV-negative	17
Previously treated presumptive cases	
HIV-positive	20.8
HIV-negative	19.5
Adherence to smear/culture-based algorithm	
New presumptive cases with two smears	85
Previously treated presumptive cases with culture	85
Adherence to Xpert-based algorithm	
All presumptive cases with Xpert test done	85
Proportion smear or Xpert-negative with culture testing	
Smear/culture-based algorithm	
New presumptive cases	
HIV-positive	85
HIV-negative	0
Previously treated presumptive cases	
HIV-positive	85
HIV-negative	0
Xpert-based algorithm	
New presumptive cases	
HIV-positive	85
HIV-negative	0
Previously treated presumptive cases	
HIV-positive	85
HIV-negative	0

* Input parameters used in all scenarios except where specific parameters are changed for a scenario (Table A.6).
HIV = human immunodeficiency virus; TB = tuberculosis.

the Xpert algorithm. This comparison built confidence in the outputs from the model, and confirmed that the outputs were credible. Overall model outputs closely resembled TB yield observed in Cape Town over the seven time points, with a mean difference of 0.0% ($P = 0.988$) between the routine data and the model outputs (Table 2).

Model sensitivity analysis

We selected five parameters to test the sensitivity of our results to uncertainty in the parameter values. The parameters evaluated were the estimated TB prevalence among presumptive cases, the proportion of presumptive cases with previous anti-tuberculosis treatment, the test sensitivity of smear microscopy and Xpert, the proportion of adherence to testing algorithms and the extent of use of culture. These parameters were selected because they have a direct impact on the probability of being correctly tested with TB, and therefore an impact on the primary outputs, i.e., diagnosed as TB (TB yield), missed cases (false-negative) and unnecessarily treated cases (false-positive).

This analysis is summarised in Tables A.14 and A.15 as well as in Figures A.3 and A.4. The analysis shows that TB diagnostic yield is sensitive to TB prevalence among presumptive cases and, to a lesser extent, to test sensitivity and previous history of TB. The proportion of HIV-positive cases among presumptive cases had more of an effect on the Xpert-based algorithm, with a decrease in yield as the proportion of HIV-positive cases increased.

Table A.5 Input parameters used in other simulated scenarios comparing the smear/culture and Xpert-based algorithms (%)

	Smear/culture-based algorithm		Xpert-based algorithm	
Scenario B: increase estimated TB prevalence among presumptive cases by 10%				
	Scenario A	Scenario B	Scenario A	Scenario B
New presumptive cases				
HIV-negative	17	27	17	27
HIV-positive	17.8	27.8	17.8	27.8
Previously treated presumptive cases				
HIV-negative	19.5	29.5	19.5	29.5
HIV-positive	20.8	30.8	20.8	30.8
Scenario C: increase additional culture testing for smear or Xpert-negative presumptive cases				
	Scenario A	Scenario C	Scenario A	Scenario C
New presumptive cases				
HIV-negative	0	30	0	5
HIV-positive	85	92	85	92
Previously treated presumptive cases				
HIV-negative	0	10	0	10
HIV-positive	85	95	85	95
Scenario D: increased proportion of presumptive cases know their HIV status				
	Scenario A	Scenario D	Scenario A	Scenario D
Proportion of presumptive cases	50%	85%	50%	85%
Scenario E: adherence to algorithms				
	Scenario A (85%)	Scenario E (100%)	Scenario A (85%)	Scenario E (100%)
New presumptive cases				
HIV-negative	2 smear	2 smear	Xpert test	Xpert test
HIV-positive	2 smear	2 smear	Xpert test	Xpert test

Table A.5 *(continued)*

	Smear/culture-based algorithm		Xpert-based algorithm	
Previously treated presumptive cases				
HIV-negative	Culture test	Culture test	Xpert test	Xpert test
HIV-positive	Culture test	Culture test	Xpert test	Xpert test
As part of follow-up testing if smear or Xpert is negative				
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	100	85	100
Previously treated presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	100	85	100
Scenario F: increased proportion of presumptive cases know their HIV status (100%) and adherence to algorithm (100%)				
	Scenario A	Scenario F	Scenario A	Scenario F
Proportion of presumptive cases	50% Scenario A (85%)	100% Scenario F (100%)	50% Scenario A (85%)	100% Scenario F (100%)
New presumptive cases				
HIV-negative	2 smear	2 smear	Xpert test	Xpert test
HIV-positive	2 smear	2 smear	Xpert test	Xpert test
Previously treated presumptive cases				
HIV-negative	Culture test	Culture test	Xpert test	Xpert test
HIV-positive	Culture test	Culture test	Xpert test	Xpert test
As part of follow-up testing if smear or Xpert is negative				
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	100	85	100
Previously treated presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	100	85	100
Scenario G: remove culture test as part of the sequence of tests required in each diagnostic algorithm				
	Scenario A	Scenario G	Scenario A	Scenario G
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	0	0	0	0
Previously treated presumptive cases				
HIV-negative	85	0	0	0
HIV-positive	85	0	0	0
As part of follow-up testing if smear or Xpert test is negative				
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	0	85	0
Previously treated presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	0	85	0
Scenario H: lower test sensitivity of smear microscopy by 10%				
	Scenario A	Scenario H		
HIV-negative	75	65		
HIV-positive	65	55		

TB = tuberculosis; HIV = human immunodeficiency virus.

Table A.6 Scenario A: comparison of smear/culture and Xpert-based algorithm model outputs (with 85% adherence to algorithms)*

	Smear/culture-based algorithm			Xpert-based algorithm			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	14.8	17.3	15.8	17.1	19.2	17.9	2.3 ($P < 0.001$)	1.9 ($P < 0.001$)	2.1 ($P < 0.001$)
FP†	0.9	1.0	0.9	1.9	2.2	2.0	1	1.2	1.1
FN†	20.7	14.4	18.0	12.6	9.5	11.3	8.1	4.9	6.7
New presumptive cases									
Yield†	13.6	15.2	14.2	16.5	17.9	17.0	2.9 ($P < 0.001$)	2.7 ($P < 0.001$)	2.8 ($P < 0.001$)
FP†	0.9	0.9	0.9	2.0	2.1	2.0	1.1	1.2	1.1
FN†	25.3	21.3	23.8	13.4	11.5	12.7	11.9	9.8	11.1
Previously treated presumptive cases									
Yield†	19.5	21.5	20.6	19.1	21.7	20.5	0.4 ($P = 0.476$)	0.2 ($P = 0.644$)	0.0 ($P = 0.894$)
FP†	0.9	1.0	0.9	1.8	2.3	2.1	0.9	1.3	1.2
FN†	3.9	2.2	3.0	9.5	6.1	7.6	5.6	3.9	4.6

* In Scenario A, 85% of cases in each algorithm received the initial tests as required and 85% of smear- or Xpert-negative cases who were HIV-infected underwent culture; 50% of presumptive cases knew their HIV status. The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
† TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.7 Scenario B: increase in estimated TB prevalence among presumptive cases by 10%*

	Smear/culture-based algorithm % (95%CI)			Xpert-based algorithm % (95%CI)			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	22.7	26.0	24.0	25.7	28.1	26.7	3.0 ($P < 0.001$)	2.1($P < 0.001$)	2.7 ($P < 0.001$)
FP†	1.0	1.0	1.0	1.9	2.2	2.0	0.9	1.2	1
FN†	20.5	13.9	17.8	12.1	9.6	11.0	8.4	4.3	6.8
New presumptive cases									
Yield†	21.1	23.4	21.9	25.1	26.6	25.6	4.0 ($P < 0.001$)	3.2 ($P < 0.001$)	3.7 ($P < 0.001$)
FP†	1.0	1.1	1.0	2.0	2.2	2.0	1	1.1	1
FN†	25.0	20.3	23.3	12.9	11.7	12.5	12.1	8.6	10.8
Previously treated presumptive cases									
Yield†	29.2	31.4	30.4	28.3	31.2	29.9	0.9 ($P = 0.125$)	0.2 ($P = 0.771$)	0.1 ($P = 0.212$)
FP†	0.9	0.9	0.9	1.8	2.2	2.0	0.9	1.3	1.1
FN†	3.6	2.4	3.0	8.8	5.8	7.1	5.2	3.4	4.1

* In Scenario B, the estimated TB prevalence among presumptive case was increased by 10; 50% of presumptive cases knew their HIV status. The same values for proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
† TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
TB = tuberculosis; CI = confidence interval; HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TP = true-positive.

Table A.8 Scenario C: increased proportion of smear- or Xpert-negative presumptive cases with additional culture*

	Smear/culture-based algorithm % (95%CI)			Xpert-based algorithm % (95%CI)			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	15.9	18.2	16.8	17.2	19.3	18.1	1.3 ($P < 0.001$)	1.2 ($P < 0.001$)	1.3 ($P < 0.001$)
FP†	0.9	1.0	0.9	2.0	2.2	2.1	1.1	1.2	1.2
FN†	14.5	9.8	12.5	11.9	8.8	10.6	2.6	1	1.9
New presumptive cases									
Yield†	15.0	16.4	15.5	16.7	18.1	17.2	1.7 ($P < 0.001$)	1.6 ($P < 0.001$)	1.7 ($P < 0.001$)
FP†	0.9	0.9	0.9	2.0	2.1	2.0	1.1	1.2	1.1
FN†	17.5	14.5	16.4	12.7	10.8	12.0	4.8	3.7	4.4
Previously treated presumptive cases									
Yield†	19.6	21.6	20.7	19.4	21.9	20.7	0.2 ($P = 0.632$)	0.3 ($P = 0.543$)	0.0 ($P = 0.895$)
FP†	0.9	1.0	0.9	1.9	2.4	2.1	1	1.4	1.2
FN†	3.5	1.6	2.4	8.7	5.2	6.8	5.2	3.6	4.4

* In Scenario C, 85% of cases in each algorithm underwent the initial tests as required; 50% of presumptive cases knew their HIV status. Additional culture was based on values found in routine practice for each patient category (smear/culture algorithm by 14.3%, Xpert algorithm by 8%). The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
† TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
CI = confidence interval; HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.9 Scenario D: increased proportion (85%) of presumptive cases who know their HIV status*

	Smear/culture-based algorithm			Xpert-based algorithm			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	14.8	18.7	16.4	17.1	20.5	18.4	2.3 ($P < 0.001$)	1.8 ($P < 0.001$)	2.1 ($P < 0.001$)
FP†	0.9	0.9	0.9	1.9	2.4	2.1	1	1.5	1.2
FN†	20.7	5.9	14.4	12.6	2.7	8.4	8.1	3.2	6
New presumptive cases									
Yield†	13.6	17.1	14.9	16.5	19.2	17.5	2.9 ($P < 0.001$)	2.2 ($P < 0.001$)	2.6 ($P < 0.001$)
FP†	0.9	0.9	0.9	2.0	2.4	2.1	1.1	1.5	1.2
FN†	25.3	8.9	19.3	13.4	3.4	9.7	11.9	5.5	9.6
Previously treated presumptive cases									
Yield†	19.5	22.0	20.8	19.1	23.0	21.2	0.4 ($P = 0.476$)	1.0 ($P = 0.043$)	0.0 ($P = 0.297$)
FP†	0.9	0.9	0.9	1.8	2.4	2.1	0.9	1.5	1.2
FN†	3.9	0.7	2.1	9.5	1.5	5.1	5.6	0.8	3

*Scenario D, 85% of cases in each algorithm underwent the initial tests as required and 85% of smear- or Xpert-negative cases that were HIV-infected underwent culture; 85% of presumptive cases knew their HIV status. The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
†TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.10 Scenario E: increased adherence to smear/culture and Xpert algorithm to 100%*

	Smear/culture-based algorithm % (95%CI)			Xpert-based algorithm % (95%CI)			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	14.9	17.7	16.1	17.5	19.5	18.3	2.5 ($P < 0.001$)	1.8 ($P < 0.001$)	2.2 ($P < 0.001$)
FP†	0.9	1.0	0.9	2.1	2.4	2.2	1.2	1.4	1.3
FN†	19.8	12.1	16.5	10.9	9.0	10.1	8.9	3.1	6.4
New presumptive cases									
Yield†	13.6	15.6	14.4	17.1	18.4	17.6	3.5 ($P < 0.001$)	2.7 ($P < 0.001$)	3.2 ($P < 0.001$)
FP†	0.9	0.9	0.9	2.1	2.4	2.2	1.2	1.5	1.3
FN†	25.3	19.0	22.9	11.0	10.5	10.8	14.3	8.5	12.1
Previously treated presumptive cases									
Yield†	20.3	21.9	21.2	19.0	21.8	20.5	1.3 ($P = 0.013$)	0.2 ($P = 0.744$)	0.7 ($P = 0.057$)
FP†	0.9	1.0	0.9	1.9	2.5	2.2	1	1.5	1.3
FN†	0.0	0.0	0.0	10.6	6.2	8.2	10.6	6.2	8.2

*In Scenario E, adherence to the full range of tests required in each algorithm was set at 100% from 85%; 50% of presumptive cases knew their HIV status. The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
†TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
CI = confidence interval; HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.11 Scenario F: increased proportion of presumptive cases who know their HIV status to 100% and adherence to algorithms to 100%*

	Smear/culture-based algorithm % (95%CI)			Xpert-based algorithm % (95%CI)			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	14.9	19.7	16.8	17.5	21.2	19.0	2.5 ($P < 0.001$)	1.5 ($P < 0.001$)	2.1 ($P < 0.001$)
FP†	0.9	0.8	0.9	2.1	2.7	2.3	1.2	1.9	1.4
FN†	19.8	0.0	11.5	10.9	0.0	6.3	8.9	0	5.2
New presumptive cases									
Yield†	13.6	18.5	15.4	17.1	20.1	18.2	3.5 ($P < 0.001$)	1.5 ($P < 0.001$)	2.8 ($P < 0.001$)
FP†	0.9	0.9	0.9	2.1	2.8	2.3	1.2	1.9	1.4
FN†	25.3	0.0	16.0	11.0	0.0	7.0	14.3	0	9
Previously treated presumptive cases									
Yield†	20.3	22.0	21.2	19.0	23.5	21.4	1.3 ($P = 0.013$)	1.5 ($P = 0.004$)	0.2 ($P = 0.623$)
FP†	0.9	0.8	0.8	1.9	2.6	2.3	1	1.8	1.5
FN†	0.0	0.0	0.0	10.6	0.0	4.7	10.6	0	4.7

*In Scenario F, increase in the percentage of presumptive cases who knew their HIV status from 50% to 100% and adherence to the full range of tests required in each algorithm to 100% from 85%. The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
†TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
CI = confidence interval; HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.12 Scenario G: remove culture test as part of the sequence of tests required in smear/culture and Xpert algorithms*

	Smear/culture-based algorithm % (95%CI)			Xpert-based algorithm % (95%CI)			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	14.0	13.3	13.7	17.5	19.5	18.3	3.5 ($P < 0.001$)	6.2 ($P < 0.001$)	4.6 ($P < 0.001$)
FP†	0.9	1.0	0.9	2.1	2.4	2.2	1.2	1.4	1.3
FN†	25.3	35.2	29.4	10.9	9.0	10.1	14.4	26.2	19.3
New presumptive cases									
Yield†	13.6	12.5	13.2	17.1	18.4	17.5	3.5 ($P < 0.001$)	5.9 ($P < 0.001$)	4.3 ($P < 0.001$)
FP†	0.9	0.9	0.9	2.1	2.4	2.2	1.2	1.5	1.3
FN†	25.3	36.1	29.3	11.0	20.3	10.5	14.3	15.8	18.8
Previously treated presumptive cases									
Yield†	15.4	14.8	15.1	19.0	21.8	20.5	3.6 ($P < 0.001$)	6.9 ($P < 0.001$)	5.4 ($P < 0.001$)
FP†	0.9	1.0	0.9	1.9	2.5	2.2	1	1.5	1.3
FN†	25.2	33.6	29.9	10.6	6.2	8.2	14.6	27.4	21.7

*In Scenario G, culture testing was removed as part of the sequence of tests required in each algorithm; 50% of presumptive cases knew their HIV status. The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
†TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
CI = confidence interval; HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.13 Scenario H: lower test sensitivity of smear microscopy by 10% in smear/culture algorithm*

	Smear/culture-based algorithm % (95%CI)			Xpert-based algorithm % (95%CI)			Change in yield between algorithms % difference *P* value		
	HIV−	HIV+	Overall	HIV−	HIV+	Overall	HIV−	HIV+	Overall
All presumptive cases									
Yield†	13.4	16.6	14.7	17.1	19.1	17.9	3.6 ($P < 0.001$)	2.6 ($P < 0.001$)	3.2 ($P < 0.001$)
FP†	1.0	1.0	1.0	1.9	2.2	2.0	0.9	1.2	1
FN†	28.6	18.0	24.1	12.6	9.5	11.3	16	8.5	12.8
New presumptive cases									
Yield†	12.0	14.2	12.8	16.5	17.9	17.0	4.6 ($P < 0.001$)	3.7 ($P < 0.001$)	4.2 ($P < 0.001$)
FP†	1.0	1.1	1.0	2.0	2.1	2.0	1	1	1
FN†	35.1	26.8	32.0	13.4	11.5	12.7	21.7	15.3	19.3
Previously treated presumptive cases									
Yield†	19.2	21.4	20.4	19.1	21.7	20.5	0.1 ($P = 0.881$)	0.3 ($P = 0.493$)	0.1 ($P = 0.681$)
FP†	0.9	0.9	0.9	1.8	2.3	2.1	0.9	1.4	1.2
FN†	5.4	2.7	3.9	9.5	6.1	7.6	4.1	3.4	3.7

* In Scenario H, the test sensitivity of smear microscopy was reduced by 10%. The Xpert algorithm is set as per Scenario A; 50% of presumptive cases knew their HIV status. The same values for TB prevalence, proportions of HIV–, HIV+ (status known and undiagnosed), and new and previously treated cases were used in each algorithm.
† TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
CI = confidence interval; HIV = human immunodeficiency virus; − = negative; + = positive; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

Table A.14 Summary of input and output parameters for model sensitivity analysis*

	Base value	High value	Low value
Estimated TB prevalence among presumptive cases			
Input	18.9	28.9	8.9
Output			
Smear/culture-based algorithm			
Yield	15.8	24.0	7.6
FP	0.9	1.0	1.0
FN	17.6	17.4	16.9
Xpert-based algorithm			
Yield	17.5	26.1	8.7
FP	1.9	1.9	1.9
FN	12.5	12.5	12.4
Proportion of presumptive cases with previous anti-tuberculosis treatment			
Input	25	50	15
Output			
Smear/culture-based algorithm			
Yield	15.8	17.3	15.2
FP	0.9	1.0	1.0
FN	17.6	11.7	19.3

Table A.14 *(continued)*

	Base value	High value	Low value
Xpert-based algorithm			
Yield	17.5	18.2	17.0
FP	1.9	2.0	1.9
FN	12.5	11.1	13.1
Test sensitivity of smear microscopy and Xpert			
Input			
Smear microscopy (two samples)			
HIV-negative	75	85	65
HIV-positive	65	75	55
Xpert			
HIV-negative	89	94	81
HIV-positive	80	88	67
Output			
Smear/culture-based algorithm			
Yield	15.8	17.0	14.7
FP	0.9	1.0	1.0
FN	17.6	11.4	23.6
Xpert-based algorithm			
Yield	17.5	18.1	16.5
FP	1.9	1.9	1.9
FN	12.5	9.3	17.9

* TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
TB = tuberculosis; FP = false-positive; FN = false-negative; HIV = human immunodeficiency virus; TP = true-positive.

Table A.15 Proportion of culture as part of the sequence of tests required in each diagnostic algorithm and 100% adherence to algorithms*

	Smear/culture-based algorithm		Xpert-based algorithm	
	Base	0% culture	Base	0% culture
Use of culture				
Input				
As part of the initial sequence of tests				
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	0	0	0	0
Previously treated presumptive cases				
HIV-negative	85	0	0	0
HIV-positive	85	0	0	0
As part of follow-up testing if smear or Xpert test is negative				
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	0	85	0
Previously treated presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	0	85	0
Output				
Yield	15.8	13.7	17.5	17.2
FP	0.9	0.9	1.9	2.0
FN	17.6	29.4	12.5	14.7
Adherence to algorithms				
Input				
As part of the initial sequence of tests	85% adherence	100% adherence	85% adherence	100% adherence
New presumptive cases				
HIV-negative	2 smears	2 smears	Xpert test	Xpert test
HIV-positive	2 smears	2 smears	Xpert test	Xpert test
Previously treated presumptive cases				
HIV-negative	Culture test	Culture test	Xpert test	Xpert test
HIV-positive	Culture test	Culture test	Xpert test	Xpert test
As part of followup testing if smear or Xpert test is negative	Base	0% culture	Base	0% culture
New presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	100	85	100
Previously treated presumptive cases				
HIV-negative	0	0	0	0
HIV-positive	85	100	85	100
Output	85% adherence	100% adherence	85% adherence	100% adherence
Yield	15.8	16.1	17.5	18.3
FP	0.9	0.9	1.9	2.2
FN	17.6	16.5	12.5	10.1

* TB diagnostic yield (yield) = (TP + FP)/presumptive cases; TP = correctly diagnosed with TB/culture-positive TB cases; FP = incorrectly diagnosed with TB/culture-negative TB cases; FN = incorrectly NOT diagnosed with TB (missed)/culture-positive TB cases.
HIV = human immunodeficiency virus; FP = false-positive; FN = false-negative; TB = tuberculosis; TP = true-positive.

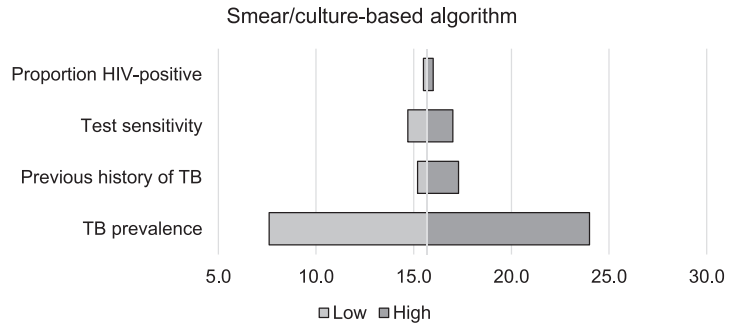

Figure A.3 Oneway sensitivity analysis with results on the effect of change in input parameters on TB diagnostic yield for the smear/culturebased algorithm. HIV = human immunodeficiency virus; TB = tuberculosis.

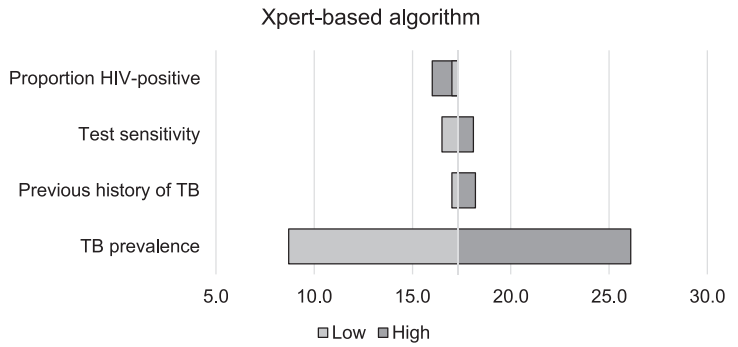

Figure A.4 One-way sensitivity analysis with results on the effect of change in input parameters on TB diagnostic yield for the Xpert-based algorithm. HIV = human immunodeficiency virus; TB = tuberculosis.

References

1 Dowdy D W, Houben R, Cohen T, et al. Impact and cost-effectiveness of current and future tuberculosis diagnostics: the contribution of modelling. Int J Tuberc Lung Dis 2014; 18: 1012–1018.
2 Lin H-H, Langley I, Mwenda R, et al. A modelling framework to support the selection and implementation of new tuberculosis diagnostic tools. Int J Tuberc Lung Dis 2011; 15: 996–1004.
3 Waaler H T, Piot M A. The use of an epidemiological model for estimating the effectiveness of tuberculosis control measures. Sensitivity of the effectiveness of tuberculosis control measures to the coverage of the population. Bull World Health Organ 1969; 41: 75–93.
4 Robinson S. Simulation: the practice of model development and use. J Field Ornithol 2004; 67: 336.
5 Günal M M, Pidd M. Discrete event simulation for performance modelling in health care: a review of the literature. J Simulation 2010; 4: 42–51.
6 Fone D, Hollinghurst S, Temple M, et al. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. J Public Health Med 2003; 25: 325–335.
7 Langley I, Doulla B, Lin H-H, Millington K, Squire B. Modelling the impacts of new diagnostic tools for tuberculosis in developing countries to enhance policy decisions. Health Care Manag Sci 2012; 15: 239–253.
8 Lanner. WITNESS Service and Process Performance Edition. Redditch, UK: http://www.lanner.com/technology/witness-simulation-software.html Accessed February 2017
9 Mase S R, Ramsay A, Ng V, et al. Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: a systematic review. Int J Tuberc Lung Dis 2007; 11: 485–495.
10 Steingart K R, Henry M, Ng V, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis 2006; 6: 570–581.
11 Steingart K R, Sohn H, Schiller I, et al. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev 2013; (1): CD009593.
12 Middelkoop K, Bekker L-G, Myer L, et al. Antiretroviral program associated with reduction in untreated prevalent tuberculosis in a South African township. Am J Respir Crit Care Med 2010; 182: 1080–1085.
13 Wood R, Middelkoop K, Myer L, et al. Undiagnosed tuberculosis in a community with high HIV prevalence: implications for tuberculosis control. Am J Respir Crit Care Med 2007; 175: 87–93.
14 den Boon S, van Lill S W P, Borgdorff M W, et al. High prevalence of tuberculosis in previously treated patients, Cape Town, South Africa. Emerg Infect Dis 2007; 13: 1189–1194.
15 Naidoo P, Dunbar R, Lombard C, et al. Comparing tuberculosis diagnostic yield in smear/culture and Xpert® MTB/RIF-based algorithms using a non-randomised stepped-wedge design. PLOS ONE 2016; 11(3): e0150487.
16 Sargent G R. Verification and validation of simulation models. J Simul 2013; 7: 12–24.
17 Nanoo A, Izu A, Ismail N A, et al. Nationwide and regional incidence of microbiologically confirmed pulmonary tuberculosis in South Africa, 2004–12: a time series analysis. Lancet Infect Dis 2015; 15: 1066–1076.
18 Botha E, Den Boon S, Verver S, et al. Initial default from tuberculosis treatment: how often does it happen and what are the reasons? Int J Tuberc Lung Dis 2008; 12: 820–823.
19 Claassens M M, du Toit E, Dunbar R, et al. Tuberculosis patients in primary care do not start treatment. What role do health system delays play? Int J Tuberc Lung Dis 2013; 17: 603–607.
20 Department of Health. National tuberculosis management guidelines 2014. Pretoria, South Africa: DoH, 2014.

RESUME

CONTEXTE : Le Cap, Afrique du Sud.

OBJECTIF : Comparer le rendement diagnostique des algorithmes de frottis/culture et d'Xpert® MTB/RIF et étudier les mécanismes influençant le rendement de la tuberculose (TB).

MÉTHODE : Nous avons élaboré et validé un modèle opérationnel du processus de diagnostic de la TB, d'abord avec l'algorithme de frottis/culture et ensuite avec l'algorithme de l'Xpert. Nous avons modélisé les scénarios en variant la prévalence de la TB, l'adhésion aux algorithmes de diagnostic et le statut du virus de l'immunodéficience humaine. Ceci a permis de faire des comparaisons directes du rendement diagnostique dans les deux algorithmes.

RÉSULTATS : Les données de routine ont montré que le rendement diagnostique avait diminué pendant la période de lancement de l'algorithme Xpert par rapport à la période où l'algorithme frottis/culture était en place. Cependant, le rendement de la modélisation dans des conditions identiques a mis en évidence une augmentation de 13,3% du rendement diagnostique de l'algorithme Xpert comparé au frottis/culture. Le modèle a démontré que l'utilisation extensive de la culture dans l'algorithme frottis/culture et le déclin de la prévalence de la TB étaient les principaux facteurs contribuant à ne pas trouver d'augmentation du rendement diagnostique dans les données de routine.

CONCLUSION : Nous avons démontré les bénéfices d'un modèle opérationnel afin de déterminer l'effet de l'expansion d'un nouvel algorithme de diagnostic et de recommander que les décideurs politiques utilisent la modélisation opérationnelle pour prendre des décisions appropriées avant que de nouveaux algorithmes de diagnostic ne soient étendus.

RESUMEN

MARCO DE REFERENCIA: La Ciudad del Cabo en Suráfrica.

OBJETIVO: Comparar el desempeño de un algoritmo diagnóstico basado en la baciloscopia y el cultivo y un algoritmo con la prueba Xpert® MTB/RIF e investigar los mecanismos que influyen en su eficacia.

MÉTODOS: Se creó un modelo operativo del proceso diagnóstico de la tuberculosis (TB) y se evaluó inicialmente con el algoritmo de la baciloscopia y el cultivo y luego con el algoritmo que incluía la prueba Xpert. Se simularon modelos con diferentes hipótesis de prevalencia de TB, adhesión a los algoritmos y situación frente a la infección por el virus de la inmunodeficiencia humana. Estos modelos permitieron una comparación directa del rendimiento diagnóstico de ambos algoritmos.

RESULTADOS: Los datos de la práctica corriente pusieron de manifiesto que el rendimiento diagnóstico disminuyó durante el período de despliegue del algoritmo con la prueba Xpert en comparación con el rendimiento que se lograba cuando se aplicaba el algoritmo de la baciloscopia y el cultivo. Sin embargo, al utilizar la modelización en idénticas condiciones, se obtuvo un aumento de 13,3% del rendimiento diagnóstico del algoritmo con la prueba Xpert en comparación con el algoritmo de la baciloscopia y el cultivo. La modelización reveló que un uso extenso del cultivo en el algoritmo de la baciloscopia y el cultivo y la disminución de la prevalencia de TB fueron los principales factores que explicaban el hecho de no haber logrado un mejor rendimiento diagnóstico en los datos de la práctica corriente con la prueba Xpert.

CONCLUSIÓN: En el presente estudio se demuestra la utilidad de un modelo operativo diseñado con el propósito de determinar el efecto de la ampliación de escala de un nuevo algoritmo diagnóstico y se recomienda que las instancias normativas apliquen la modelización operativa a fin de adoptar las decisiones apropiadas, antes de ampliar la escala de nuevos algoritmos diagnósticos.