

ARTICLE **OPEN**

Understanding the incremental value of novel diagnostic tests for tuberculosis

Nimalan Arinaminpathy¹ & David Dowdy²

Tuberculosis is a major source of global mortality caused by infection, partly because of a tremendous ongoing burden of undiagnosed disease. Improved diagnostic technology may play an increasingly crucial part in global efforts to end tuberculosis, but the ability of diagnostic tests to curb tuberculosis transmission is dependent on multiple factors, including the time taken by a patient to seek health care, the patient's symptoms, and the patterns of transmission before diagnosis. Novel diagnostic assays for tuberculosis have conventionally been evaluated on the basis of characteristics such as sensitivity and specificity, using assumptions that probably overestimate the impact of diagnostic tests on transmission. We argue for a shift in focus to the evaluation of such tests' incremental value, defining outcomes that reflect each test's purpose (for example, transmissions averted) and comparing systems with the test against those without, in terms of those outcomes. Incremental value can also be measured in units of outcome per incremental unit of resource (for example, money or human capacity). Using a novel, simplified model of tuberculosis transmission that addresses some of the limitations of earlier tuberculosis diagnostic models, we demonstrate that the incremental value of any novel test depends not just on its accuracy, but also on elements such as patient behaviour, tuberculosis natural history and health systems. By integrating these factors into a single unified framework, we advance an approach to the evaluation of new diagnostic tests for tuberculosis that considers the incremental value at the population level and demonstrates how additional data could inform more-effective implementation of tuberculosis diagnostic tests under various conditions.

Nature 528, S60–S67 (3 December 2015), DOI: 10.1038/nature16045

This article has not been written or reviewed by *Nature* editors. *Nature* accepts no responsibility for the accuracy of the information provided.

Every year, nearly three million people develop active tuberculosis (TB), but are not notified to health authorities¹. Some of these individuals may spontaneously resolve their disease, die or be treated in the private sector, but many remain infectious, fuelling ongoing transmission in the community. Reaching this 'missing three million' remains one of the top priorities for global TB control². A widely cited reason for the ongoing gap between incidence and cases notified is the lack of highly sensitive and deployable diagnostic tests for TB³. Sputum smear microscopy, the global cornerstone of TB diagnosis⁴, can miss half of all people with infectious TB⁵, whereas more sensitive tests cannot routinely be implemented at the point of treatment^{6,7}. Nevertheless, the link between improved diagnostic sensitivity and better TB detection remains uncertain. Studies^{8–11} in different settings have found little or no change in the number of pulmonary TB diagnoses or deaths when comparing sputum smear microscopy and Xpert MTB/RIF, a more sensitive molecular test¹². This result may reflect high levels of empirical treatment among people who test negative^{13–15}. Against this backdrop, a key question remains: if novel diagnostic tests are developed and implemented at scale, what impact can we expect on TB epidemiology within populations?

The impact of TB diagnostics on transmission reflects not only the accuracy of the test, but also the way in which patients with infectious TB interact with members of the community and with health systems over time^{16,17}. These infection pathways have at least three crucial dimensions: the transmission rate (number of transmission events per unit time), the frequency at which people contact health systems (often slower in subpopulations with poor access to care), and the probability of starting effective TB treatment after such contact¹⁸. Each of these dimensions varies through the duration of infectiousness (from onset to effective treatment, spontaneous recovery or death)¹⁹.

Mathematical models can be a useful tool in helping to demonstrate how these dimensions relate to the impact of diagnostic tests on TB transmission^{20–22}. Figure 1a depicts the simplest, and most commonly used^{23–25}, conceptualization of TB diagnosis in mathematical models so far. In this framework, on becoming infectious, people with TB experience a series of uniform processes. Specifically, they transmit TB at a constant rate, contact the health system at a constant rate and undergo a constant probability of successful diagnosis (leading to appropriate treatment) with each health-system contact. In this framework, the speed at which someone with TB gets treated — and the number of people they infect before that treatment — are strongly related to the sensitivity of the diagnostic algorithm. If, for example, people with TB contact the health system on average every 6 months with a 50% chance of being diagnosed at each visit, the mean duration of infectiousness will be 1 year (approximately the prevalence/incidence ratio estimated by the World Health Organization¹). If a more sensitive test (for example, replacing sputum smear microscopy with Xpert MTB/RIF^{26,27}) can increase that probability of diagnosis from 0.5 to 0.75, the mean duration of disease, and thus the transmission per active case, could be cut by one-third. As a result, the projected epidemiological impact of a more sensitive diagnostic test in this framework is tremendous. This conceptualization of the diagnostic process (constant transmission, constant health-system contact and constant probability of successful diagnosis) over time has permeated nearly all projections of expected epidemiological impact from novel diagnostic tests for pulmonary TB — and it is almost certainly wrong. Figure 1b shows an alternative conceptualization of the TB diagnostic process. In this framework, the transmission rate, frequency of health-system contact and probability of successful diagnosis can all change

¹MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College London, Norfolk Place, London W2 1PG, UK. ²Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA. Correspondence should be addressed to: N. A. e-mail: nim.pathy@imperial.ac.uk or D. D. e-mail: ddowdy1@jhmi.edu.

over time¹⁹. As an illustration, if patients remain infectious for an average of 10 months before seeking care and then begin to contact the health system once a month^{28,29}, a 50% chance of successful diagnosis per visit would still result in a mean duration of infectiousness of 1 year — but increasing the probability of diagnosis from 0.5 to 0.75 would only reduce that duration to 11.3 months. Worse still, if most transmissions occur in the first 10 months, then even a perfect diagnostic test at the health facility could not avert those events. Thus, the dynamic trajectories of transmission, health-care seeking and diagnostic index of suspicion over the course of TB disease are inextricably linked to the epidemiological impact of novel diagnostic tests^{19,30-33} — and overly simple depictions of those trajectories may systematically overestimate that impact. Adding complexity to these simple frameworks requires additional data to inform a more nuanced understanding of the impact of diagnostic tests. Without such data, and models with sufficient flexibility to incorporate them, it is likely that projections of the impact of novel diagnostic tests on TB transmission will continue to be biased, often dramatically so.

So far, test accuracy (sensitivity and specificity) — and to a lesser extent, feasibility of implementation in peripheral settings — has dominated thinking about the ‘value’ of new TB diagnostic tests. However, the impact of any novel TB diagnostic test will depend on how the health-care system incorporates it³⁴, as well as on the dynamics of patient interactions with that health-care system (Fig. 1). Epidemiologically, therefore, a novel diagnostic assay should be evaluated not by its sensitivity and specificity, but rather the extent to which it provides diagnostic information beyond earlier tests and practices³⁵ — its incremental value. This concept is similar to the classic concept of the expected value of diagnostic information (EVDI) promoted by Phelps and Mushlin³⁶, who also highlighted the need to combine the EVDI with estimates of cost or resource requirements. Subsequent work has expanded on this concept^{37,38}. In this paper, we use principles of infectious-disease modelling and diagnostic epidemiology to argue for a change in conceptual approach, from one that has focused primarily on a test’s sensitivity to one that centres on its incremental value.

METHODS

Quantifying the incremental value of diagnostic tests for TB. In the context of TB, there are a number of benefits that new diagnostics could provide. These include, but are not limited to, averting TB transmission, averting TB morbidity and mortality⁹, saving money^{39,40}, freeing up health-care capacity for other activities, enabling better treatment of other conditions by ruling out TB⁴¹ and improving patients’ economic situations⁴² or quality of life⁴³. We focus here on the use of novel diagnostic tests as tools to avert TB transmission; however, the intention of some tests may be to add value in one or more of these other areas — and each test’s utility should be evaluated according to its intended purpose.

To appropriately estimate the incremental value of a new diagnostic test for TB in terms of transmissions averted, one must consider its relationship to the diagnostic pathways outlined in Figure 1. Table 1 lists four defining features of TB disease and diagnosis (latency⁴⁴, gradual symptom onset^{45,46}, reliance on sputum⁴⁷ and concentration of transmission among ‘superspreaders’⁴⁸). These features highlight a number of potential diagnostic gaps, or elements along the TB diagnostic pathway, which, if filled by a novel diagnostic test, could generate substantial incremental value.

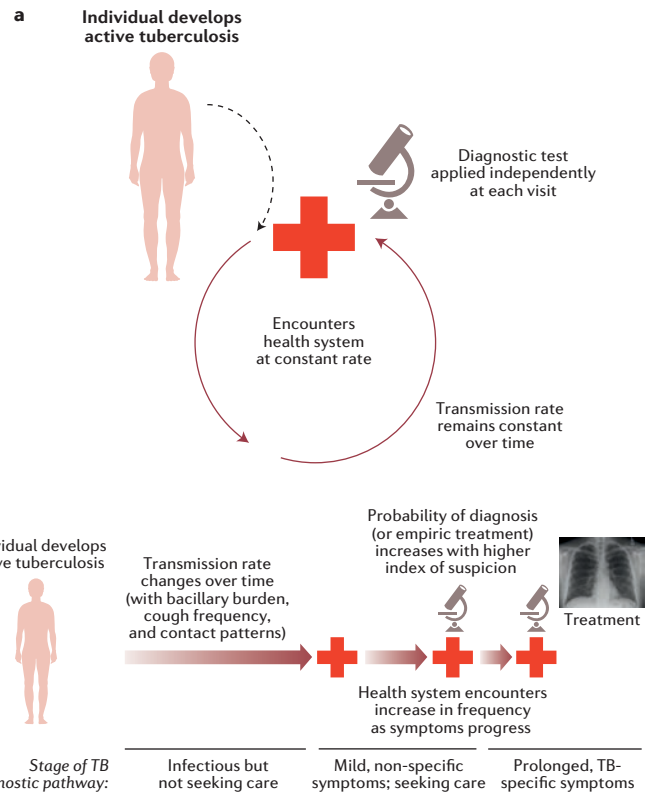


Figure 1 | Conceptual diagrams of different tuberculosis (TB) diagnostic models. a. The ‘standard’ model. So far, most models of TB diagnosis have assumed that, on becoming infectious, individuals with active TB transmit their disease at a constant rate, seek care at a constant rate and maintain a constant probability of diagnosis and treatment with each care-seeking attempt. In reality, the rate at which individuals with active TB transmit disease and seek care, as well as the probability of successful diagnosis and treatment, change over time with the disease course. This process can be more accurately represented by assuming three different stages in the TB diagnostic pathway, as represented at the bottom of **b**. This framework accommodates different types of variation that can be crucial in the potential impact of a test. For example, patients might transmit their disease at an increasing rate over time as bacillary burden increases, seek care more frequently as symptoms progress, and be more likely to receive ancillary diagnostic tests (or empiric treatment) as symptoms persist and other diagnoses become less likely.

Specifically, in any given setting, TB transmission may occur primarily from people who are not sufficiently ill to seek care^{49,50}; those who are seeking care, but have symptoms (for example, a mild cough) not specific to TB⁵¹; or those with severe or prolonged symptoms, but who test negative for TB and are therefore not treated (Fig. 1b). Alternatively, most transmission may occur from hard-to-reach populations in which the rate for seeking care is low⁵². Each of these

Table 1 | Four potential diagnostic gaps in tuberculosis (TB).

Feature of TB natural history	Description	Resultant source of transmission	Potential representation within models	Diagnostic test capable of filling gap
Latency	Prolonged latent period	Individuals who are asymptomatic or have only very mild symptoms	Asymptomatic (or mildly symptomatic) infectious state (I_0)	Test to identify who will progress to active disease, allowing targeted preventive therapy (‘progression biomarker’)
Slow clinical course	Early non-specific symptoms (for example, cough)	Individuals who are presenting to care, but for syndromic management	Infectious state with symptoms sufficient to drive care seeking, but with low index of suspicion for TB (I_1)	Test to rule out TB (or suggest further testing for TB) in people with a cough (‘cough triage test’)
Difficult microbiological confirmation	Bacilli often present in low numbers, and only in lungs or sputum; no specific antibody	Individuals who test false negative for TB	Active, care-seeking but undiagnosed state (I_2)	Test to supplant current tests with imperfect sensitivity (‘smear replacement test’)
Heterogeneous transmission and access to care	Transmission concentrated among those with poor access to care	Individuals who lack sufficient access to seek care rapidly	State with lower care-seeking rate (I')	Smear replacement test for use in peripheral settings with poor access (‘point-of-care test’)

Table 2 | Profiles of three illustrative diagnostic tests for tuberculosis (TB).

Illustrative name (see Table 1)	Descriptive profile	Mathematical representation	Approximate number needed to screen to identify 1 additional case, typical high-burden setting	Avertable transmission load
Progression biomarker	This test could be applied to a general population to identify people who would subsequently develop active TB; these people could be treated with highly effective preventive regimens	A proportion of individuals with latent TB infection are returned to the uninfected state if successfully identified and treated	100–500 (1/lifetime probability of incident TB × probability of completing effective preventive therapy)	Pre-care seeking, mild symptoms and prolonged symptoms (general population only)
Cough triage test	This test could be applied to all people presenting to care with a cough, even if suspicion for TB was low — those testing positive could have a highly sensitive test performed	The probability of successful diagnosis in the early symptomatic period increases	20–100 (prevalence of active TB among all patients with a cough)	Mild symptoms and prolonged symptoms (general population only)
Smear replacement test	This test would allow highly sensitive diagnosis among those already seeking care with high suspicion of TB	The probability of successful diagnosis in the late symptomatic period increases	10–20 (prevalence of smear-negative active TB among those with smears currently performed)	Prolonged symptoms (general population only)
Point-of-care test	As for the smear replacement test, but one that is possible to deploy in the poor-access population	As above	10–20 (prevalence of smear-negative active TB among those with smears currently performed)	Prolonged symptoms (low-access population)

gaps suggests a potential diagnostic solution that would have high incremental value. This may be a test to predict progression to active TB (and thus allow targeted preventive therapy), one optimized for diagnosing combinations of symptoms (such as cough and fever), one that is simply more sensitive, or one that is more deployable to peripheral and informal settings (Table 2)⁵³. We incorporate these possibilities more formally into a mathematical model of TB transmission.

Model description. Figure 2 presents a simple, illustrative model of TB diagnosis and transmission that expands the constant care-seeking approach shown in Figure 1a. In this model, the population is divided into different compartments that reflect the natural history of TB and incorporate both the stages of the diagnostic pathway shown in Figure 1b and the corresponding diagnostic gaps listed in Table 1. Movement of people between these compartments can be represented by a system of ordinary differential equations, with rates of transition between compartments (for example, γ_0 , the rate of initiating care seeking) that reflect the inverse of the mean duration of time spent in each phase (for example, the mean duration between onset of infectiousness and beginning to seek care). As most of these durations are currently unknown (and differ from one setting to the next), we assume — for the purposes of illustration — a population that is at equilibrium, with values of TB incidence, prevalence and mortality that reflect a setting of moderate TB burden (see Supplementary Information). We then use this simplified model to estimate, in this hypothetical setting, the incremental value of diagnostic tests with different profiles under different assumptions about the relative importance of each diagnostic gap. This simplified model divides the population of individuals with active TB into three categories (Figs 1b and 2): those who are infectious, but who are not actively seeking care (I_0), those who have early symptoms that trigger less frequent care seeking and who have a lower probability of correct diagnosis/empiric therapy (I_1), and those who have characteristic and prolonged symptoms that trigger frequent care seeking and a likely diagnosis with each attempt (I_2). We also assume a general population and a sub-population (I' , set at 10% for the purposes of illustration) with 'poor' access to care whose rate of care seeking is a specified fraction (k , set initially at 0.5) of the rate in the general population.

Importantly, this model captures the three dynamic processes of transmission, health-care seeking and empiric treatment shown in Figure 1b. First, the rate of transmission (the probability of a 'contact' resulting in TB transmission, multiplied by the number of potential contacts per unit time) can vary over time. For example, β_0 (the number of transmissions per person-month spent in the asymptomatic infectious state I_0) may be higher than β_1 and β_2 (transmission rate from the symptomatic states I_1 and I_2), because the contact rate with susceptible individuals may be highest early in the disease course (suggested by the high prevalence of TB infection in contact investigations⁵⁴). Alternatively, the inverse might be true because the bacillary burden grows over time⁵⁵. We capture this in the concept of the 'transmission load', which we define as the proportion of transmission events at the population level that occur in each of these three stages. Second, the rate of seeking care can increase over time as symptoms progress. Third, the probability of diagnosis

with each care-seeking attempt can also increase over time, as symptoms become more suggestive of underlying TB disease⁵⁶. These two processes can be combined into a single 'rate of successful diagnosis and treatment' (d) that increases over time from d_0 to d_1 to d_2 .

We explore three hypothetical settings for how transmission varies during the course of TB disease: late diagnostic gap, in which the transmission rate β is four-fold higher at each subsequent stage of TB disease (for example, constant contact rate with susceptible individuals with increasing bacillary burden); early diagnostic gap, in which β falls by a factor of four at each stage (for example, pool of susceptible individuals shrinks over time as household members and other close contacts are exposed); and high access disparity, in which those with least access to care are assumed to have a rate of diagnosis and treatment that is 10% (rather than 50%) that of the general population. Each setting is calibrated to have the same level of TB incidence (see Supplementary Information).

In the context of each of these settings, we explore the potential incremental

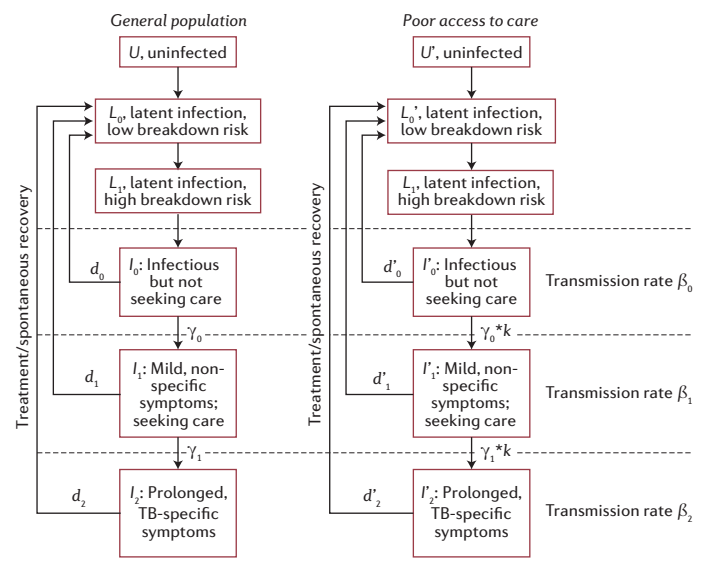


Figure 2 | Model structure relating diagnostic pathways to transmission load. A representation of a simple mathematical model that incorporates the three stages of diagnosis shown in Figure 1b. Relative rates of transmission, β , can vary from one stage to the next, with γ representing the inverse of the mean duration of each stage at the population level. Upward arrows denote removal of cases through diagnosis and curative treatment, d , as well as spontaneous resolution (not shown, for simplicity). We also assume a fixed proportion of the population (10% in the base case) have 'poor' access to care, defining an 'access disparity parameter' k to reflect the relative rates of diagnosis in this population. At baseline, we assume that $k = 0.5$. TB, tuberculosis.



Figure 3 | Tuberculosis (TB) transmission load under three alternative scenarios. The size of each bar denotes the transmission load, defined as the percentage of all tuberculosis transmission that occurs within a given diagnostic stage. Transmission from the general population is shown in darker colours, with that originating from the ‘poor-access’ population shown in lighter colours. Interrupting transmission at a given stage also averts transmission in subsequent stages (for example, diagnosing a case in stage I_1 also averts the transmission that this case could have caused in stage I_2); this effect can be calculated as the sum of the transmission load in the relevant stage plus all subsequent stages within that population.

value of four illustrative diagnostic tests: a ‘progression biomarker’ that predicts progression from latent to active TB (to facilitate preventive therapy)⁵⁷; a ‘triage’ test that facilitates syndromic diagnosis of people presenting with cough⁵⁸; a more sensitive ‘replacement test’ to supplant current sputum-based confirmatory tests for TB⁵⁹; and a ‘point-of-care test’ that can replace sputum smear in peripheral settings⁶⁰, thereby (unlike the other three tests) being accessible to those with poor access to care. These tests, along with their mathematical representation in our simplified modelling framework, are summarized in Table 2.

We focus on comparisons between these types of diagnostic tests when they are added to the standard of care. To illustrate the transmission contributions of different groups, we assume that progression biomarker, triage and replacement tests are deployed in the general population, whereas the point-of-care test is deployed in the poor-access population. We discuss below how different diagnostic gaps might cause each of these illustrative tests to be preferred over the others, thereby emphasizing the importance of quantifying (or at least estimating) the diagnostic gap in any given setting.

Incorporating resource constraints. Ultimately, discussions of a new diagnostic test’s incremental value must also consider any constrained resources — whether economic or otherwise — that would be required to implement the test. One method for evaluating the incremental value of a diagnostic test in a given setting is to first identify any constrained resources required for test implementation. The additional resources required to change from the existing standard of care to an algorithm that augments that standard of care with the new diagnostic test can then be estimated (the incremental resource requirement)⁶¹. Finally, this is combined with estimates of the incremental number of transmissions averted under this augmented algorithm, relative to the standard of care (incremental impact). Thus, tests that aim to avert TB transmission can be compared using an inverse incremental cost-effectiveness ratio⁶²: (incremental transmissions averted)/(incremental resource requirement), or

$$(T_1 - T_0)/(R_1 - R_0) \quad (1)$$

where 1 denotes the presence of the new test and 0 denotes its absence.

BOX 1 | ESTIMATING THE INCREMENTAL VALUE OF TUBERCULOSIS DIAGNOSTIC TESTS, PER UNIT OF CONSTRAINED RESOURCES.

In comparing diagnostic tests for tuberculosis (TB), it is important to consider both the incremental impact (presented here as transmissions averted) and the incremental resource requirements associated with implementing each new test. The following considerations are not meant to be an exhaustive list, but a demonstration of some of the complexity that must be considered (and corresponding data collected) to properly evaluate the incremental value of diagnostic tests for TB in the setting of constrained resources. Illustrative considerations therefore include:

Determinants of incremental impact (incremental transmissions averted)

1. Epidemiological setting/existing diagnostic gaps
2. Diagnostic test characteristics (accuracy, diagnostic gap targeted)
3. Existing diagnostic algorithms (incremental role of the new test)

Determinants of incremental resource requirements

1. Enumeration of constrained resources
2. Number of tests needed to identify one additional case
3. Per-test outlay of constrained resources (‘unit cost’)

In settings in which TB diagnostic tests are being compared with other interventions (for example, TB treatment or HIV diagnosis), transmissions averted can be converted into measures of health utility (such as disability-adjusted life years, or DALYs, averted)⁶³ to estimate resources in terms of economic costs and to report this incremental value as an incremental cost-effectiveness ratio. However, when only comparing diagnostic tests with the same primary aim (to avert transmission), the formulation of incremental value in Equation 1 may be more useful; this formulation places the emphasis on impact rather than cost and does not require additional model assumptions to convert transmissions into DALYs or constrained resources (for example, human capacity) into economic costs. Therefore, we use this more direct formulation in our model results.

RESULTS

Incremental value of TB diagnostic tests. Figure 3 shows how the transmission load at equilibrium (the proportion of population-level transmission contributed by each stage) differs in each transmission scenario. For example, in the late diagnostic gap scenario, 35% of all transmission originates from individuals with mild symptoms in the general population, whereas this percentage falls to 5% in the early diagnostic gap scenario. Importantly, averting transmission in the earlier stages (for example, preventing a case from developing, even before to care seeking) also averts that transmission in later stages — seen in Figure 3 by the combined value of the stacked bars. Thus, for example, preventing all transmission in the latter two care-seeking stages in the general population would avert 51% (35% + 16%) of all transmission in the late diagnostic gap scenario, compared with only 5% in the early diagnostic gap scenario — and a diagnostic test targeting these stages (for example, the ‘cough triage’ test) might be expected to have greater impact in settings that more closely resemble the late diagnostic gap scenario.

A notable feature of the late diagnostic gap scenario is that, despite transmission being substantially more intense⁶⁴ in the prolonged-symptom stage I_2 (16 times greater per unit time than in the pre-care-seeking stage I_0), the contribution of this stage to transmission remains relatively modest. This is largely due to the relatively short time that individuals spend in this late symptomatic stage. We assume here that, under the standard of care (typically using sputum smear microscopy), individuals are diagnosed on average after 1 month in this late symptomatic stage, compared with 6 months spent in the asymptomatic stage. However, the high access disparity scenario shows the

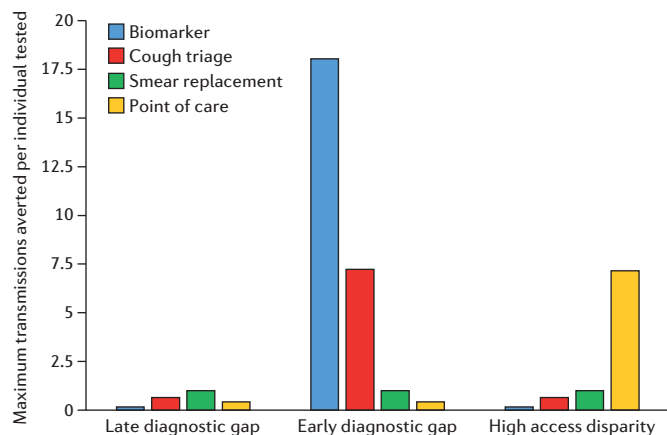


Figure 4 | Maximum incremental impact per unit of constrained resources for four illustrative diagnostic tests in three alternative scenarios. On the y-axis is the maximum incremental impact (number of tuberculosis (TB) transmissions averted) for each of four illustrative TB diagnostic tests, divided by the incremental resources required to implement each test (Equation 1). All measures are benchmarked to an incremental impact of 1 for the smear replacement test in the general population. Here, we assume that the constrained resources are simply proportional to the number of people needed to test to diagnose one additional case of active TB. The maximum incremental impact is the number of transmissions that would be averted if diagnosis averts all transmission associated with a given patient stage in Figure 2. Accordingly, the results presented here should be interpreted as an upper bound that are illustrative of the role of diagnostic gaps in each stage. In the cases illustrated here, the 'progression biomarker' (which identifies individuals at risk for progression to active TB) is clearly favoured in the early diagnostic gap scenario, whereas the point-of-care test (which replaces the smear test in the poor-access population, and is deployed only in the poor-access population) is strongly favoured in the high access disparity scenario.

potential importance of the late symptomatic stage when the rate of diagnosis is diminished. Here, transmission in the late symptomatic stage is sufficiently strong for 55% of the transmission load to occur from a high-risk (and symptomatic) subgroup that accounts for no more than 10% of the total population — a level of disproportionate transmission that is only modestly higher than has been suggested in some settings⁶⁵.

Incremental value of new diagnostic tests under constrained resources. Figure 4 shows results for the incremental value (Equation 1), comparing diagnostic tests that target different stages and under different transmission scenarios. For the denominator of Equation 1, Figure 4 assumes a simple, illustrative example for which the constrained resource is the number of individuals who can be tested with a novel test, irrespective of the test type or its unit cost (see Supplementary Table 2 for further details). This might, for example, reflect a setting in which donor funding could be obtained to implement a new test, but the equipment or human resources available to conduct those tests were extremely limited. For the numerator of Equation 1, Figure 4 assumes the maximum number of transmissions averted if the diagnostic test in question (such as the cough triage test) could avert all of the transmission occurring in the stage of disease targeted (for example, I_1 , mild symptoms, but seeking care). In practice, owing to factors such as imperfect sensitivity and incomplete population-level implementation, an actual test would only avert a portion of that maximum transmission load; the actual incremental value of each test would therefore be proportionally lower. Thus, in dividing the maximum incremental impact by the fixed incremental resources available, Figure 4 compares the maximum incremental value for each idealized test type, leaving it to subsequent work to estimate what proportion of that maximum could actually be achieved by a given test in practice.

Figure 4 illustrates that — where the primary diagnostic gap is early in the disease course — the maximum incremental value for tests that target earlier stages is higher than that of the smear-replacement test. By contrast, when the primary diagnostic gap is late in the disease course, the maximum

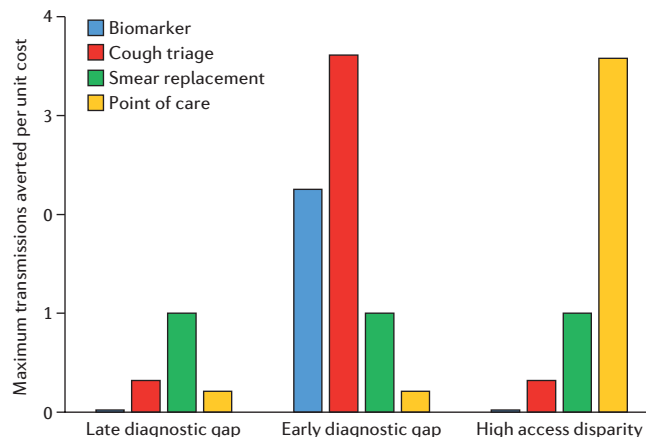


Figure 5 | Maximum incremental value per unit of constrained resources after incorporation of a cost function. The same illustrative tests are evaluated in the same alternative scenarios as in Figure 4, but in this case we apply a cost function that accounts for the fact that diagnosis earlier in the disease process, or among low-access populations, is generally more resource-intensive on a per-test basis (see Supplementary Table 2 for full assumptions). After considering this cost function, the 'progression biomarker' is no longer clearly favoured in the early diagnostic gap scenario, and the degree to which the point-of-care test is favoured over the smear replacement test in the high access disparity scenario is reduced by the same factor (8 in this case) by which the cost per person screened in the low-access population exceeds that in the general population. As in Figure 4, all measures are benchmarked to an incremental impact of 1 for the smear replacement test in the general population.

incremental value of the later-stage diagnostics is far greater (as represented by their markedly higher incremental value). Notably, where transmission is concentrated among a population with particularly poor access to care, the maximum incremental value for a test that can be implemented in this population can be considerably higher than for any other test (as in the access disparity scenario).

Figure 5 shows an alternative scenario for the denominator of Equation 1 in which the limiting resource is financial (for example, a fixed amount of money available), assuming that the cost per test is higher when applied earlier in the diagnostic pathway. (For example, it is more costly to screen a patient for TB in a prevalence survey⁶⁶ than it is in a clinic⁶⁷.) The 'unit cost' of a test is also assumed to be higher per person when applied in the poor-access population (see Supplementary Table 2), as these individuals are assumed to be harder to reach than the general population. In the early diagnostic gap scenario, for example, considering this unit cost dramatically lowers the maximum incremental value of the biomarker test that could be achieved per unit of the constrained resource, relative to the cough triage and smear replacement tests. As a result, under this alternative resource constraint, the cough triage test, rather than the biomarker, would be preferred.

DISCUSSION

In evaluating novel diagnostic tests for TB, it is crucial that we move beyond simple considerations of elements such as sensitivity, specificity and turnaround time — and instead begin to consider the incremental value of diagnostic tests that fit certain profiles. We use a simple mathematical model to demonstrate key trade-offs in an illustrative setting. This work demonstrates how diagnostic tests for TB can be quantitatively assessed in terms of their incremental value (incremental impact divided by incremental resource requirement), and moreover how this incremental value can vary from one setting to the next.

The prevailing diagnostic gap in a given setting has a profound effect on the potential incremental impact of each diagnostic test. When most transmission occurs before patients begin to seek care, diagnostic tests that require patients to access the health system are unlikely to have substantial epidemiological impact; thus, in the early diagnostic gap scenario (Fig. 3) the only diagnostic test capable of averting the bulk of the transmission load is the prevention biomarker. Similarly, when a substantial disparity exists between

the high-risk and general population, diagnostics that cannot be implemented in the high-risk group are limited in their potential.

Consideration of incremental impact must also include consideration of incremental resource requirements, however. For example, the resources required to avert a transmission are generally much greater when diagnostic tests are performed early in the disease course⁶⁶, or in hard-to-reach populations. As a result, those diagnostic tests with the largest maximum incremental impact may also be those that require the most resources. In estimating the incremental resource requirement of a given test, it is important to consider the resources for TB control that are constrained in a given setting. In many cases, these constrained resources will be purely financial, but in others, there may be limitations on the availability of trained staff or laboratory capacity to perform certain tests⁶⁸. The per-test incremental outlay of the most constrained resources is therefore also likely to vary from one setting to the next. Ultimately, the incremental value of a TB diagnostic test depends not on sensitivity and specificity, but also on multiple factors that will vary from one system to the next (Box 1). For any setting, all six of the elements in Box 1 should be evaluated to help to identify the type of novel test that is likely to have the greatest incremental value (avert the most TB transmission events, given the constrained resources). As assessments of these factors are performed across a variety of settings, consensus may emerge as which tests should be prioritized for development.

Unfortunately, we currently lack the empirical data in most settings to make such an informed assessment. Specifically, it is likely that different transmission loads and diagnostic gaps — early, late or among high-risk subpopulations — predominate in different settings, and that resource constraints vary widely from one setting to the next. How can this data gap be closed?

First, we require better evidence regarding how novel diagnostic tests function when implemented under field conditions. Such data would allow us to estimate the proportion of any diagnostic gap that a new TB test could close, as well as the number of tests required to make one additional diagnosis. Unfortunately, most diagnostic tests are evaluated primarily in well-funded trials and demonstration studies, without good evidence of how they perform in the real world. For example, Xpert MTB/RIF was recommended on the basis of high-quality data about its accuracy and cost-effectiveness under controlled conditions and in a large field trial²⁶; however, emerging evidence has suggested that, in many settings, the characteristics of Xpert may be different when implemented in the field — including its sensitivity⁶⁹, calibration⁷⁰, positive predictive value (owing to low pre-test probability)⁷¹, and accuracy for rifampin resistance⁷². To make accurate assessments of the incremental value of diagnostics, we should collect such data early after launch, and update expectations and recommendations as those data become available.

Second, we need better data on the performance of existing tests, including clinical judgement. These data would enable us to evaluate the incremental number of transmissions that a novel test might be able to avert, relative to the existing standard of care. A series of recent high-quality studies suggests that, when patients present with symptoms that are highly suggestive of TB in upper-middle income settings (for example, South Africa and Brazil), the probability of empirical diagnosis is reasonably high^{8–11} — but that a large number of people may be presenting to care with a cough without TB ever being considered⁸. Such studies are crucial to understand the likely diagnostic gaps for TB, but unfortunately, very few such analyses have been performed in settings with fewer resources (for example, most of sub-Saharan Africa^{73,74} and Southeast Asia) where empirical diagnosis rates (and the capacity to implement novel diagnostic tests) may be much lower. Characterizations of relative TB transmission from high-risk populations (akin to the ‘low-access’ population in Figure 3) compared with the general population are also sparse⁷⁵, and could potentially be informed by better use of surveillance data⁷⁶.

Third, and perhaps most challengingly, we need to prioritize characterizations of the transmission load and diagnostic gaps in a variety of settings. If we can describe the prevailing transmission loads in any given setting, we can then quantify the maximum incremental impact (transmissions averted) of any diagnostic test in that setting. Ultimately, for any setting, one should be able to delineate what proportion of the transmission load in each of the phases of TB (pre-care seeking, mildly symptomatic and prolonged symptomatic in the general population and in high-risk groups) is being averted using

existing tests, and therefore what proportion might still be amenable to implementation of a novel diagnostic. Molecular characterization of TB (for example, through whole-genome sequencing⁷⁷) in entire populations is becoming available and can be linked to conventional epidemiological investigations (for example, through contact investigations⁷⁸) using increasingly discriminatory tools for analysis and data collection⁷⁹. Thus, it may become possible to triangulate an infectious individual’s onset of symptoms, initiation of care-seeking activities and specific transmission events. Studies that merge data on transmission, contact patterns, symptom histories, care-seeking patterns and interactions with the health-care system on a population level should be prioritized in this regard. In the meantime, simple investigation of surveillance data can help to identify geographic hotspots of transmission, and operational analyses of diagnostic test implementation can demonstrate where diagnoses are probably being missed. Although estimating the duration of an infectious episode poses significant challenges, household cohorts using currently available tools could cast some light on the ‘transmission load’ that occurs early in the clinical course^{80,81}.

Finally, we need better investigations of constrained resources in specific settings to enumerate the resources that are genuinely constrained, and to quantify those resources per test performed (as the equivalent of a unit cost). Although conventional economic evaluations of interventions against diseases such as TB implicitly consider money to be the most constrained resource, other studies in low-income settings have shown that human resources, laboratory capacity, regulatory infrastructure or ability to implement new interventions may be the key limiting factors⁶⁸. This may be especially true in the modern era of direct assistance for health — which may supply money, but not resources in the form of trained personnel⁸². An understanding of the most constrained resources in any given setting must then be merged with data on the number of tests required to identify an incremental case, as well as the per-test resource outlay, for any given novel diagnostic test. Only if we truly understand the resources that are most constrained in a given setting, as well as the resource outlay for each type of diagnostic test, can we identify the diagnostic tests that will optimize epidemiological impact under existing resource constraints.

Ultimately, the only way to end TB is to diagnose and treat people with TB before transmission occurs — novel diagnostics are an essential component of any strategy with this aim. If we are to succeed in that endeavour, we must think of, and quantify, those tests not just in terms of sensitivity, specificity and turnaround time, but rather in terms of their incremental value across a variety of epidemiological settings. We present a framework for estimating this incremental value that also highlights the need for additional data in order to inform more appropriate prioritization of novel TB diagnostic tests, across settings that may differ in their existing diagnostic gaps and resource constraints. As we continue to develop diagnostic tests with the goal of curbing TB transmission, we must think beyond accuracy and consider the broader context of patient behaviour, health systems and TB natural history.

1. World Health Organization. *Global Tuberculosis Report, 2014* (WHO, 2014).
2. Herbert, N. et al. World TB Day 2014: finding the missing 3 million. *Lancet* **383**, 1016–1018 (2014).
3. Lawn, S. D. et al. Advances in tuberculosis diagnostics: the Xpert MTB/RIF assay and future prospects for a point-of-care test. *Lancet Infect. Dis.* **13**, 349–361 (2013).
4. TB CARE I. *International Standards for Tuberculosis Care 3rd Edn* (TB CARE I, 2014).
5. Steingart, K. R. et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect. Dis.* **6**, 570–581 (2006).
6. Cobelens, F. et al. Which new diagnostics for tuberculosis, and when? *J. Infect. Dis.* **205**, S191–S198 (2012).
7. Pho, M. T. et al. Optimizing tuberculosis case detection through a novel diagnostic device placement model: the case of Uganda. *PLoS ONE* **10**, e0122574 (2015).
8. Churchyard, G. et al. Effect of Xpert MTB/RIF on early mortality in adults with suspected TB: a pragmatic randomized trial. *Proc. Conference on Retroviruses and Opportunistic Infections* (IAS-USA, 2014).
9. Theron, G. et al. Feasibility, accuracy, and clinical effect of point-of-care Xpert MTB/RIF testing for tuberculosis in primary-care settings in Africa: a multicentre, randomised, controlled trial. *Lancet* **383**, 424–435 (2014).
10. Durovni, B. et al. Impact of replacing smear microscopy with Xpert MTB/RIF for diagnosing tuberculosis in Brazil: a stepped-wedge cluster-randomized trial. *PLoS Med.* **11**, e1001766 (2014).
11. Sachdeva, K. S. et al. Use of Xpert MTB/RIF in Decentralized Public Health Settings and Its Effect on Pulmonary TB and DR-TB Case Finding in India. *PLoS ONE* **10**, e0126065 (2015).

12. Steingart, K. R. et al. Xpert(R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst. Rev.* **1**, CD009593 (2013).
13. Theron, G. et al. Do high rates of empirical treatment undermine the potential effect of new diagnostic tests for tuberculosis in high-burden settings? *Lancet Infect. Dis.* **14**, 527–532 (2014).
14. Sun, A. Y., Denkinger, C. M. & Dowdy, D. W. The impact of novel tests for tuberculosis depends on the diagnostic cascade. *Eur. Respir. J.* **44**, 1366–1369 (2014).
15. Menzies, N. A., Cohen, T., Murray, M. & Salomon, J. A. Effect of empirical treatment on outcomes of clinical trials of diagnostic assays for tuberculosis. *Lancet Infect. Dis.* **15**, 16–17 (2015).
16. Dye, C. The potential impact of new diagnostic tests on tuberculosis epidemics. *Indian J. Med. Res.* **135**, 737–744 (2012).
17. Langley, I. et al. Assessment of the patient, health system, and population effects of Xpert MTB/RIF and alternative diagnostics for tuberculosis in Tanzania: an integrated modelling approach. *Lancet Glob. Health.* **2**, e581–e591 (2014).
18. Dowdy, D. W. & Chaisson, R. E. The persistence of tuberculosis in the age of DOTS: reassessing the effect of case detection. *Bull. World Health Organ.* **87**, 296–304 (2009).
19. Dowdy, D. W., Basu, S. & Andrews, J. R. Is passive diagnosis enough? The impact of subclinical disease on diagnostic strategies for tuberculosis. *Am. J. Respir. Crit. Care Med.* **187**, 543–551 (2013).
20. Zwerling, A. et al. Modeling of novel diagnostic strategies for active tuberculosis - a systematic review: current practices and recommendations. *PLoS ONE* **9**, e110558 (2014).
21. Garnett, G. P., Cousens, S., Hallett, T. B., Steketee, R. & Walker, N. Mathematical models in the evaluation of health programmes. *Lancet* **378**, 515–525 (2011).
22. Dowdy, D. W., Dye, C. & Cohen, T. Data needs for evidence-based decisions: a tuberculosis modeler's 'wish list'. *Int. J. Tuberc. Lung Dis.* **17**, 866–877 (2013).
23. Dye, C., Garnett, G. P., Sleeman, K. & Williams, B. G. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. *Lancet* **352**, 1886–1891 (1998).
24. Dowdy, D. W., Chaisson, R. E., Moulton, L. H. & Dorman, S. E. The potential impact of enhanced diagnostic techniques for tuberculosis driven by HIV: a mathematical model. *AIDS* **20**, 751–762 (2006).
25. Abu-Raddad, L. J. et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proc. Natl Acad. Sci. USA.* **106**, 13980–13985 (2009).
26. Boehme, C. C. et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* **377**, 1495–1505 (2011).
27. Boehme, C. C. et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N. Engl. J. Med.* **363**, 1005–1015 (2010).
28. Storla, D. G., Yimer, S. & Bjune, G. A. A systematic review of delay in the diagnosis and treatment of tuberculosis. *BMC Public Health* **8**, 15 (2008).
29. Sreeramareddy, C. T., Panduru, K. V., Menten, J. & Van den Ende, J. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. *BMC Infect. Dis.* **9**, 91 (2009).
30. Salje, H. et al. The importance of implementation strategy in scaling up Xpert MTB/RIF for diagnosis of tuberculosis in the Indian health-care system: a transmission model. *PLoS Med.* **11**, e1001674 (2014).
31. Sachdeva, K. S. et al. The potential impact of up-front drug sensitivity testing on India's epidemic of multi-drug resistant tuberculosis. *PLoS ONE* **10**, e0131438 (2015).
32. Lin, H. H., Dowdy, D., Dye, C., Murray, M. & Cohen, T. The impact of new tuberculosis diagnostics on transmission: why context matters. *Bull. World Health Organ.* **90**, 739–747A (2012).
33. Mandal, S. & Arinaminpathy, N. Transmission modeling and health systems: the case of TB in India. *Int. Health* **7**, 114–120 (2015).
34. McNERNEY, R., CUNNINGHAM, J., HEPPLE, P. & ZUMLA, A. New tuberculosis diagnostics and rollout. *Int. J. Infect. Dis.* **32**, 81–86 (2015).
35. Grobbee, D. E. & Hoes, A. W. *Clinical Epidemiology: Principles, Methods, and Applications for Clinical Research* (Jones and Bartlett, 2009).
36. Phelps, C. E. & Mushlin, A. I. Focusing technology assessment using medical decision theory. *Med. Decis. Making* **8**, 279–289 (1988).
37. Laking, G., Lord, J. & Fischer, A. The economics of diagnosis. *Health Econ.* **15**, 1109–1120 (2006).
38. Van den Bruel, A., Cleemput, I., Aertgeerts, B., Ramaekers, D. & Buntinx, F. The evaluation of diagnostic tests: evidence on technical and diagnostic accuracy, impact on patient outcome and cost-effectiveness is needed. *J. Clin. Epidemiol.* **60**, 1116–1122 (2007).
39. van't Hoog, A. H. et al. Optimal triage test characteristics to improve the cost-effectiveness of the Xpert MTB/RIF assay for TB diagnosis: a decision analysis. *PLoS ONE* **8**, e82786 (2013).
40. Millman, A. J. et al. Rapid molecular testing for TB to guide respiratory isolation in the U.S.: a cost-benefit analysis. *PLoS ONE* **8**, e79669 (2013).
41. Cobelens, F., van Kampen, S., Ochodo, E., Atun, R. & Lienhardt, C. Research on implementation of interventions in tuberculosis control in low- and middle-income countries: a systematic review. *PLoS Med.* **9**, e1001358 (2012).
42. Tanimura, T., Jaramillo, E., Weil, D., Raviglione, M. & Lonroth, K. Financial burden for tuberculosis patients in low- and middle-income countries: a systematic review. *Eur. Respir. J.* **43**, 1763–1775 (2014).
43. Bauer, M., Leavens, A. & Schwartzman, K. A systematic review and meta-analysis of the impact of tuberculosis on health-related quality of life. *Qual. Life Res.* **22**, 2213–2235 (2013).
44. Getahun, H., Matteelli, A., Chaisson, R. E. & Raviglione, M. Latent *Mycobacterium tuberculosis* infection. *N. Engl. J. Med.* **372**, 2127–2135 (2015).
45. Corbett, E. L. et al. Epidemiology of tuberculosis in a high HIV prevalence population provided with enhanced diagnosis of symptomatic disease. *PLoS Med.* **4**, e22 (2007).
46. Mao, T. E. et al. Cross-sectional studies of tuberculosis prevalence in Cambodia between 2002 and 2011. *Bull. World Health Organ.* **92**, 573–581 (2014).
47. Ho, J., Marks, G. B. & Fox, G. J. The impact of sputum quality on tuberculosis diagnosis: a systematic review. *Int. J. Tuberc. Lung Dis.* **19**, 537–544 (2015).
48. Gardy, J. L. et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N. Engl. J. Med.* **364**, 730–739 (2011).
49. Ayles, H. et al. Prevalence of tuberculosis, HIV and respiratory symptoms in two Zambian communities: implications for tuberculosis control in the era of HIV. *PLoS ONE* **4**, e5602 (2009).
50. Hoa, N. B. et al. Yield of interview screening and chest X-ray abnormalities in a tuberculosis prevalence survey. *Int. J. Tuberc. Lung Dis.* **16**, 762–767 (2012).
51. Banda, H. T. et al. Prevalence of tuberculosis in TB suspects with short duration of cough. *Trans. R. Soc. Trop. Med. Hyg.* **92**, 161–163 (1998).
52. Lonroth, K. et al. Towards tuberculosis elimination: an action framework for low-income countries. *Eur. Respir. J.* **45**, 928–952 (2015).
53. Denkinger, C. M. et al. Defining the needs for next generation assays for tuberculosis. *J. Infect. Dis.* **211** (Suppl 2), S29–S38 (2015).
54. Morrison, J., Pai, M. & Hopewell, P. C. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect. Dis.* **8**, 359–368 (2008).
55. Lawn, S. D., Kerkhoff, A. D. & Wood, R. Progression of subclinical culture-positive tuberculosis to symptomatic disease in HIV-infected individuals. *AIDS* **25**, 2190–2191 (2011).
56. Harries, A. D. et al. Management of pulmonary tuberculosis suspects with negative sputum smears and normal or minimally abnormal chest radiographs in resource-poor settings. *Int. J. Tuberc. Lung Dis.* **2**, 999–1004 (1998).
57. Wallis, R. S. et al. Biomarkers and diagnostics for tuberculosis: progress, needs, and translation into practice. *Lancet* **375**, 1920–1937 (2010).
58. Garcia-Basteiro, A. L. & Cobelens, F. Triage tests: a new priority for tuberculosis diagnostics. *Lancet Respir. Med.* **3**, 177–178 (2015).
59. Kik, S. V., Denkinger, C. M., Jefferson, C., Ginnard, J. & Pai, M. Potential market for novel tuberculosis diagnostics: worth the investment? *J. Infect. Dis.* **211** (Suppl 2), S58–S66 (2015).
60. Pai, N. P. & Pai, M. Point-of-care diagnostics for HIV and tuberculosis: landscape, pipeline, and unmet needs. *Discov. Med.* **13**, 35–45 (2012).
61. Adang, E., Voordijk, L., Jan van der Wilt, G. & Ament, A. Cost-effectiveness analysis in relation to budgetary constraints and reallocation restrictions. *Health Policy* **74**, 146–156 (2005).
62. Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J. & Stoddart, G. L. In *Methods for the Economic Evaluation of Health Care Programmes* 396 (Oxford Univ. Press, 2005).
63. Azman, A. S., Golub, J. E. & Dowdy, D. W. How much is tuberculosis screening worth? Estimating the value of active case finding for tuberculosis in South Africa, China, and India. *BMC Med.* **12**, 216 (2014).
64. Behr, M. A. et al. Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet* **353**, 444–449 (1999).
65. Dowdy, D. W., Golub, J. E., Chaisson, R. E. & Saraceni, V. Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. *Proc. Natl Acad. Sci. USA.* **109**, 9557–9562 (2012).
66. Glaziou, P. et al. Tuberculosis prevalence surveys: rationale and cost. *Int. J. Tuberc. Lung Dis.* **12**, 1003–1008 (2008).
67. Schnippel, K. et al. Scaling up Xpert MTB/RIF technology: the costs of laboratory- vs. clinic-based roll-out in South Africa. *Trop. Med. Int. Health* **17**, 1142–1151 (2012).
68. Travis, P. et al. Overcoming health-systems constraints to achieve the Millennium Development Goals. *Lancet* **364**, 900–906 (2004).
69. Sohn, H. et al. Xpert MTB/RIF testing in a low tuberculosis incidence, high-resource setting: limitations in accuracy and clinical impact. *Clin. Infect. Dis.* **58**, 970–976 (2014).
70. Cowan, J. et al. Implementing rapid testing for tuberculosis in Mozambique. *Bull. World Health Organ.* **93**, 125–130 (2015).
71. Page-Shipp, L. et al. Successes, challenges and lessons from a novel deployment of Xpert(R) MTB/RIF at a major South African public event. *Int. J. Tuberc. Lung Dis.* **18**, 438–440 (2014).
72. Kelly, J. D. et al. Xpert MTB/RIF false detection of rifampin-resistant tuberculosis from prior infection. *Am. J. Respir. Crit. Care Med.* **190**, 1316–1318 (2014).
73. Nakiyingi, L. et al. Clinical predictors and accuracy of empiric tuberculosis treatment among sputum smear-negative HIV-infected adult TB suspects in Uganda. *PLoS ONE* **8**, e74023 (2013).
74. Yoon, C. et al. Impact of Xpert MTB/RIF testing on tuberculosis management and outcomes in hospitalized patients in Uganda. *PLoS ONE* **7**, e48599 (2012).
75. Dowdy, D. W., Azman, A. S., Kendall, E. A. & Mathema, B. Transforming the fight against tuberculosis: targeting catalysts of transmission. *Clin. Infect. Dis.* **59**, 1123–1129 (2014).
76. Jenkins, H. E. et al. Assessing spatial heterogeneity of multidrug-resistant tuberculosis in a high-burden country. *Eur. Respir. J.* **42**, 1291–1301 (2013).
77. van Soolingen, D. Whole-genome sequencing of *Mycobacterium tuberculosis* as an epidemiological marker. *Lancet Respir. Med.* **2**, 251–252 (2014).
78. Walker, T. M., Monk, P., Smith, E. G. & Peto, T. E. Contact investigations for outbreaks of *Mycobacterium tuberculosis*: advances through whole genome sequencing. *Clin. Microbiol. Infect.* **19**, 796–802 (2013).
79. Merker, M. et al. Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature Genet.* **47**, 242–249 (2015).
80. Grandjean, L. et al. Transmission of multidrug-resistant and drug-susceptible tuberculosis within households: a prospective cohort study. *PLoS Med.* **12**, e1001843 (2015).

81. Zelner, J. L. *et al.* Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. *Am. J. Epidemiol.* **180**, 853–861 (2014).
82. Ravishankar, N. *et al.* Financing of global health: tracking development assistance for health from 1990 to 2007. *Lancet* **373**, 2113–2124 (2009).

SUPPLEMENTARY MATERIAL

Is linked to the online version of this paper at: <http://dx.doi.org/10.1038/nature16045>

ACKNOWLEDGEMENTS

The work of D.D. was partially funded by the B. Frank and Kathleen Polk Assistant Professorship in Epidemiology at the John Hopkins Bloomberg School of Public Health.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests. Financial support for this publication has been provided by the Bill & Melinda Gates Foundation.

ADDITIONAL INFORMATION



This work is licensed under the Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0>