HHS Public Access

Author manuscript

Health Care Manag Sci. Author manuscript; available in PMC 2018 December 01.

Optimal timing of drug sensitivity testing for patients on first-line tuberculosis treatment

Sze-chuan Suen, PhD[1], **Margaret L. Brandeau, PhD**[2], and **Jeremy D. Goldhaber-Fiebert, PhD**[3]

[1]Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA, USA

[2]Department of Management Science and Engineering, Stanford University, Stanford, CA, USA

[3]Stanford Health Policy, Centers for Health Policy and Primary Care and Outcomes Research, Stanford University, Stanford, CA, USA

Abstract

Effective treatment for tuberculosis (TB) patients on first-line treatment involves triaging those with drug-resistant (DR) TB to appropriate treatment alternatives. Patients likely to have DR TB are identified using results from repeated inexpensive sputum-smear (SS) tests and expensive but definitive drug sensitivity tests (DST). Early DST may lead to high costs and unnecessary testing; late DST may lead to poor health outcomes and disease transmission. We use a partially observable Markov decision process (POMDP) framework to determine optimal DST timing. We develop policy-relevant structural properties of the POMDP model. We apply our model to TB in India to identify the patterns of SS test results that should prompt DST if transmission costs remain at status-quo levels. Unlike previous analyses of personalized treatment policies, we take a societal perspective and consider the effects of disease transmission. The inclusion of such effects can significantly alter the optimal policy. We find that an optimal DST policy could save India approximately $1.9 billion annually.

Keywords

POMDP; Optimal Testing; Tuberculosis; Drug Resistance

1 Introduction

In 2013, 1.5 million people died from tuberculosis (TB) and, of the estimated 9 million people who developed TB, 24% were in India [1]. Most deaths from TB are preventable through treatment. Drug-sensitive (DS) TB can be effectively treated with first-line regimens using rifampicin, isoniazid, pyrazinamide, and ethambutol antibacterial drugs [2]. Drug-resistant (DR) strains of TB may not respond to first-line treatment, and patients with DR TB may need more toxic and costly DR TB treatment to be cured. In India, 1.53 million patients were put on TB treatment in 2009, and approximately 40% of those were new cases on first-line treatment [3]. While the majority of these cases are drug-sensitive, the WHO estimates that 3% of new TB cases in India are drug-resistant [4].

Standard sputum-smear (SS) tests used for diagnosing and monitoring TB are inexpensive and easy to perform but have low sensitivity and are unable to distinguish bacteriological strain. Drug sensitivity testing (DST) has traditionally been done through bacteriological culture, which can take months to perform. New advances in diagnosis systems such as the Xpert MTB/RIF assay can detect rifampicin resistance within hours, but the optimal time to administer these tests for drug strain identification has not been established. The Indian government is interested in the diagnostic improvements that Xpert may be able to provide, mentioning Xpert in its national strategic plan for TB elimination for 2017–2025 [5] as well as listing it as a new tool under evaluation in the most recent TB training module [2].

Several cost-effectiveness analyses have found Xpert adoption to be cost-effective despite its high cost [6,7,8,9]. These cost-effectiveness studies use mathematical models of TB to project the effect of Xpert adoption. We also use a mathematical model of TB to estimate the benefits and costs of TB diagnosis strategies.

Both dynamic compartmental models and individual-level simulations have been used to estimate the effect of disease dynamics in the presence of heterogeneous demographic characteristics [10,11,12,13,14]. These models can be used to quantify the uncertainty around natural disease parameters with scarce data, such as transmission patterns [11]. Keeling and Rohani describe how to use who-mixes-with-whom matrices to capture these transmission patterns [15], and recently, additional data on mixing patterns for respiratory illnesses have been collected [16,17]. However, no studies address when in the TB testing algorithm the Xpert test should be used, nor how transmission influences the appropriate timing of such testing.

For new first-line patients in India, treatment regimens call for DST only if the patient is still SS positive after four months or at the end of treatment [2]. During this delay, patients with DR TB may continue to transmit disease and experience declining health and high mortality rates. However, this delay also reduces the pool of patients who require relatively expensive DST as patients with DS TB respond to treatment.

In this paper we use a finite-horizon, partially observable Markov decision process (POMDP) model to determine the optimal timing and frequency of SS test information collection and DST testing for smear-positive, treatment-naive pulmonary TB patients on first-line treatment. A POMDP is appropriate for this problem because the physician cannot observe the patient's health state directly but can only infer it from SS test result history when making a decision about whether to continue first-line treatment or test the patient for drug resistance. We maximize net monetary benefits to reflect the desire to reduce costs and achieve good health outcomes, and we take a societal perspective, including all costs and health outcomes regardless of source or beneficiary. The POMDP allows us to identify a social welfare maximizing strategy that provides the expected best action for all possible test outcomes in every period.

Our contributions in this research are threefold. First, we provide policy-relevant structural properties of the POMDP model that allow our framework to be useful in contexts outside our example case of TB in India. To draw general conclusions about when observations (in

this case, SS test results) are valuable, we examine the cases where observations are not available versus when they are always available for every time period. We determine analytical properties of when observations are valueless for changing future decisions, which offers insight into when observations should be foregone if they are costly to acquire.

Second, we determine optimal DST policies for India. We examine policies over age- and sex-stratified input parameters to determine how demographic differences influence the optimal policy, and we provide diagnostic algorithms for regions with little to no TB transmission (such as remote rural areas) and regions with transmission rates that reflect the national average transmission rate in India. This is the first analysis that uses information from repeated SS tests taken over the course of first-line treatment to identify which TB patients should undergo DST, and when.

Third, we establish optimal strategies from the societal perspective in both the presence and absence of disease transmission. Previous work rarely considers population effects in dynamic decision models for personalized treatment. While a number of POMDP studies address optimal treatment regimens for noninfectious diseases such as cancer [18,19], Parkinson's disease [20], and dementia [21], studies using a POMDP framework to optimize treatment for infectious diseases are rare, and do not include population effects such as disease transmission. We find that the inclusion of such effects can significantly alter the societal optimal policy, even when costs associated with transmission are statically incorporated into the objective function rewards (i.e., dynamic effects are not considered). Inclusion of population effects is particularly important for a disease such as TB where transmission of DR TB is an increasing threat.

In the next sections, after describing the model structure and notation, we analytically examine two classes of problems: cases where no observations (SS test results) are available and cases where observations are taken at every time period. We identify structural properties of when observations will not change decisions. Using this result, we expand our action space in an empirical example in India to include the possibility of not taking observations. We determine the optimal testing policy for regions of India with and without significant levels of TB transmission. We provide a diagnostic testing flowchart that can be used by physicians when treating TB patients. We estimate that the optimal policy will save approximately $3000 per patient on first-line treatment, which translates to an annual potential savings of approximately $1.9 billion. We conclude with discussion.

2 Model structure

We formulate a POMDP model to determine the optimal time to administer SS tests and DST for patients on first-line TB treatment. We assume that SS tests may be administered multiple times but DST is used at most once due to its high cost. Since Xpert has high sensitivity and specificity, DST results are treated as definitive: a DST-positive patient is immediately triaged to DR TB treatment while a DST-negative patient completes the rest of the first-line treatment regimen with no further testing. International standards for TB treatment currently do not allow for early completion of the standard regimen even if the physician believes the patient to be TB-free [2]. We seek to determine when and how many

SS tests to administer, and whether and when to administer DST. In this section, we introduce our model framework. All notation is summarized in Appendix Table 1.

We use a societal perspective and measure health outcomes using total lifetime discounted quality-adjusted life years (QALYs) and total discounted lifetime healthcare costs in US dollars. Our objective is to maximize the incremental net monetary benefit (NMB) for each policy relative to a reference policy in which no first-line or DR TB treatment is offered. Net monetary benefit is defined as QALYs, converted into monetary units using a willingness-to-pay conversion factor (λ), minus costs.

Since the first-line treatment regimen is six months long, and some time needs to elapse between observations for additional treatment to take effect, we allow the decision maker to take an action once in every month the patient is on treatment. Thus we index the monthly periods by $t = 0, 1, \ldots, 6$ where $t = 0$ denotes the time of initial diagnosis, $t = 1$ denotes the end of the first month, etc.

We let $s_t \in S = \{1, 2, 3, 4, 5, 6, 7\}$ denote the true health of patients on first-line treatment, where 1 represents a healthy (cured of TB) patient, 2 represents a patient with DS TB, and 3 represents a patient with DR TB. States 4, 5, and 6 represent healthy, DS, and DR TB patients who have have defaulted from treatment (do not return for subsequent visits; lost to followup), and state 7 represents death. The physician cannot directly observe the health states 1, 2, or 3 of the patient prior to DST, but can only observe SS test results – which may be false positive or false negative. Thus, states 1, 2, and 3 are partially observed health states. We assume that the physician can observe if the patient has defaulted (but not whether the defaulted patient is healthy, has DS TB, or has DR TB), or if the patient has died.

At the beginning of each time period, the decision maker can choose one of two actions for patients on first-line treatment (states 1, 2, and 3): continue on first-line treatment for another month with monitoring (using an SS test) or administer DST. We denote by $a(s)_t$ the action taken at time t, where $a(s)_t \in A(s)_t = \{DST, wait\}$ for states $s = 1, 2, 3$. No action is possible for patients in states 4, 5, 6, and 7 because such patients are no longer in treatment. In our theoretical results, we prove that the option of having the patient remain on treatment without taking an SS test observation can sometimes be a better option than taking an SS test observation, so for our numerical analysis we expand the action space to incorporate this option: thus, $A(s)_t = \{DST, \textit{wait and see}, \textit{wait and not see}\}$ for the numerical analysis.

After each time period, patients can stay in the same health state or transition to another state. If the physician chooses the *wait* action at time t, patients in state 2 (patients with DS TB who are on first-line treatment) will be cured and move to state 1 with probability $c_{DS,t}$, and patients in state 3 (patients with DR TB who are on first-line treatment) will be cured and move to state 1 with probability $c_{DR,t}$, where $0 \leq c_{DR,t} \leq c_{DS,t} \leq 1$. The patient has some probability of default d from states 1, 2, and 3 and some probability of death ν from all alive states. Defaulted and dead patients remain in their current state. The probabilities of default and death can be allowed to change over time. We represent the treatment dynamics in the following matrix (we omit the subscript t for readability):

$$f_t = \begin{bmatrix} 1-d_H-\nu_H & c_{DS} & c_{DR} & 0 & 0 & 0 & 0 \\ 0 & 1-c_{DS}-d_{DS}-\nu_{DS} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-c_{DR}-d_{DR}-\nu_{DR} & 0 & 0 & 0 & 0 \\ d_H & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & d_{DS} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & d_{DR} & 0 & 0 & 1 & 0 \\ \nu_H & \nu_{DS} & \nu_{DR} & \nu_{dH} & \nu_{dDS} & \nu_{dDR} & 1 \end{bmatrix}$$

If the physican chooses the DST action for a patient in treatment, no further actions are possible and the problem terminates, as we assume that the test is definitive. At the end of the time horizon, patients still on treatment have completed the treatment regimen, and exit treatment with no DST.

When deciding whether to wait or administer DST, the physician does not have perfect knowledge of a patient's health state. We denote by β_t the belief at time t about the patient's health state. We represent the belief as $\beta_t \in B$, the set of all possible beliefs. Since the physician can observe if the patient has defaulted or died, we have:

$$B = \left\{ \begin{bmatrix} 1-b_{DS}-b_{DR} \\ b_{DS} \\ b_{DR} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1-b_{dDS}-b_{dDR} \\ b_{dDS} \\ b_{dDR} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

where b_{DS} and b_{DR} are the probabilities the patient has DS or DR TB, respectively, and b_{dDS} and b_{dDR} are the probabilities the patient has defaulted with DS or DR TB, respectively. The last vector represents the belief if the physician observes that the patient has died. Each vector sums to one, since the vectors represent the patient's true health state as a probability. For instance, if the physician believes that the patient who is on treatment has a 10% chance of already being cured and a 3% chance of having DR TB, the belief state is: $\beta_t = [0.10\ 0.87\ 0.03\ 0\ 0\ 0\ 0]'$.

We assume that the physician has accurate knowledge of the treatment dynamics, and thus the beliefs change after the patient undergoes an additional period of treatment. We model this by multiplying the pre-treatment belief vector β_t with the transition matrix f_t. Thus, we obtain post-treatment beliefs equal to $f_t\beta_t$.

The vector calculated above provides an update based on treatment dynamics but does not take into account any observation obtained during period t. Observations about the patient state can come from default or death information, or from SS tests, which are either negative or positive. SS tests cannot be performed if the patient is defaulted or dead, and receiving more information on the patient's state does not change the probability of default or death.

We let z_t denote the observation at time t, where $z_t \in Z_a$, the set of possible observations given action a. Thus, $Z_{wait} = \{SS+, SS-, default, death\}$, while $Z_{DST} = \{no\ default\ or\ death, default, death\}$ because no SS tests are performed.

The SS tests do not provide perfect information. We let M denote a matrix that provides the probability of observing z given that the patient is in state s. These probabilities are calculated from the sensitivity of the SS test (probability of a positive test result for a patient with TB), which we denote by *sens*, and its specificity (probability of a negative test result for a patient who does not have TB), which we denote by *spec*. In the matrix M below, the first row corresponds to $z = SS+$, the second row corresponds to $z = SS-$, the third to $z = default$, and the fourth to $z = dead$ while the columns correspond to health states 1, …, 7, respectively.

$$M = \begin{bmatrix} 1-spec & sens & sens & 0 & 0 & 0 & 0 \\ spec & 1-sens & 1-sens & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

We assume $0.5 \leq sens \leq 1$ and $0.5 \leq spec \leq 1$, which implies $(1 - spec) \leq spec$ and $(1 - sens) \leq sens$.

Given a prior belief β_t, treatment dynamics f_t, observation z_t, and knowledge of M, the belief state is updated according to a function $\overline{M}_{z_t}(f_t\beta_t)$ if the patient is not defaulted or dead. This function, composed of a transformation matrix (denoted M_+ when $z_t = SS+$ and M_- when $z_t = SS-$) and a normalization factor (denoted N_+ when $z_t = SS+$ and N_- when $z_t = SS-$), is calculated using Bayes rule, and provides a new belief state β_{t+1}. Let $\widetilde{\beta}_t(i)$ denote element i in the post-dynamics belief vector: $\widetilde{\beta}_t = f_t\beta_t$. In the case where the physician has pre-treatment belief β_t and observes SS+, we have:

$$N_+(\widetilde{\beta}_t) = \frac{1}{(1-spec)\widetilde{\beta}_t(1) + (sens)(\widetilde{\beta}_t(2) + \widetilde{\beta}_t(3))}$$

$$M_+ = \begin{bmatrix} (1-spec) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & (sens) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (sens) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore the updated belief β_{t+1} is given by:

$$(\beta_{t+1} | f_t\beta_t, z_t = SS+) = \overline{M}_+(\widetilde{\beta}_t) = N_+(\widetilde{\beta}_t)M_+\widetilde{\beta}_t = \frac{1}{(1-spec)\widetilde{\beta}_t(1) + (sens)(\widetilde{\beta}_t(2) + \widetilde{\beta}_t(3))}.$$

$$\begin{bmatrix} (1-spec) & 0 & 0 & 0\ 0\ 0\ 0 \\ 0 & (sens) & 0 & 0\ 0\ 0\ 0 \\ 0 & 0 & (sens) & 0\ 0\ 0\ 0 \\ 0 & 0 & 0 & 0\ 0\ 0\ 0 \\ 0 & 0 & 0 & 0\ 0\ 0\ 0 \\ 0 & 0 & 0 & 0\ 0\ 0\ 0 \\ 0 & 0 & 0 & 0\ 0\ 0\ 0 \end{bmatrix} \tilde{\beta}_t \quad (1)$$

The updated belief for an *SS*− observation is similar, with *spec* instead of 1−*spec* and 1−*sens* instead of *sens*. We use a similar notation for the updated belief for a *default* observation:

$$(\beta_{t+1} | f_t \beta_t, z_t = defaulted) = P(\tilde{\beta}_t) Q \tilde{\beta}_t \; where \; P(\tilde{\beta}_t) = \frac{1}{\tilde{\beta}_t(4) + \tilde{\beta}_t(5) + \tilde{\beta}_t(6)}$$

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

The updated belief for a *dead* observation is $[0, 0, 0, 0, 0, 0, 1]'$.

We will use $pr(+ | \tilde{\beta}_t)$ to denote the probability of seeing a positive observation given post-dynamics belief $\tilde{\beta}_t = [p_1, p_2, p_3, 0, 0, 0, 0]'$, where $pr(+ | \tilde{\beta}_t) = (1 - spec)p_1 + (sens)(p_1 + p_3) = (N_+(\tilde{\beta}_t))^{-1}$ and, similarly for a negative observation, $pr(- | \tilde{\beta}_t) = (N_-(\tilde{\beta}_t))^{-1}$. Similarly, $pr(default | \tilde{\beta}_t) = [0, 0, 0, 1, 1, 1, 0]\tilde{\beta}_t$ and $pr(death | \tilde{\beta}_t) = [0, 0, 0, 0, 0, 0, 1]\tilde{\beta}_t$ are the probability of observing default or death.

Our objective is to maximize total societal net monetary benefit. We define monthly and discounted lifetime NMB values. We let $r_{wait,a}$ be the monthly NMB (relative to the reference case of no treatment) for being on first-line treatment for an additional month, given action *a*. This includes costs of treatment (drugs, labor, infrastructure), transmission (for the states where the patient has DS or DR TB), as well as test costs if action *a* involves testing. Since all patients incur the same costs to receive the same care regardless of health state, and QALY differences over one month are small, we simplify $r_{wait,a}$ to be a negative scalar that does not vary over health state or beliefs.

We define rewards r as the vector of NMB values that capture all expected future discounted lifetime costs and QALYs. We use r_{DST} if a_t is DST. If DST is chosen, we assume that patients in state 3 (patients with DR TB who are on first-line treatment) get triaged to DR TB treatment, and previously unconfirmed cured patients on first-line treatment (those in state 1) and patients with DS TB on first-line treatment (those in state 2) continue on first-line treatment without further testing. We assume that DST is 100% sensitive and specific, so the health states of patients will be correctly identified. Expected discounted lifetime costs and QALYs associated with these life trajectories are used to calculate the NMB. We denote these NMB values as:

$$r_{DST} = [NMB_{DST,H}\ NMB_{DST,DS}\ NMB_{DST,DR}\ NMB_{dH}\ NMB_{dDS}\ NMB_{dDS}\ NMB_{dead}] \quad (3)$$

Thus, for instance, $NMB_{DST,DR}$ includes the cost of the DST test, the expected costs and QALYs associated with getting a positive DST test, undergoing DR TB treatment, defaulting or completing DR TB treatment, any transmission costs for this life trajectory, and the costs and health outcomes for life thereafter. We discount all values back to the present at a rate δ per period.

We also include a terminal NMB that captures all expected future discounted lifetime costs and QALYs if a patient does not undergo DST by the end of the time horizon, denoted $r_{\sim DST}$. We assume all patients exit first-line treatment at this time, regardless of cure. Note that the NMB for defaulted and dead patients is the same between DST and $\sim DST$ rewards, as these patients cannot undergo DST. We denote these NMB values as:

$$r_{\sim DST} = [NMB_{\sim DST,H}\ NMB_{\sim DST,DS}\ NMB_{\sim DST,DR}\ NMB_{dH}\ NMB_{dDS}\ NMB_{dDS}\ NMB_{dead}] \quad (4)$$

For instance, $NMB_{\sim DST,H}$ captures the lifetime QALYs and health costs of a healthy individual who has completed first-line treatment and does not undergo DST. The reward $r_{\sim DST}$ includes costs and health outcomes associated with TB transmission (if the patient has DS or DR TB). For simplicity, we will assume that these rewards are time independent for our theoretical analyses in Section 3. We allow the rewards to vary over time (i.e., we use $r_{DST,t}$ and $r_{\sim DST,t}$) in our numerical analysis for India (Section 4).

3 Theoretical results

We now develop theoretical results for the case when no SS test results are available (Section 3.1), and the case when SS test results are available (Section 3.2). All proofs are in the appendix.

3.1 Optimal actions when no SS test results are available

We first examine the simplified case in which no SS test observation is available for the duration of treatment (e.g., due to lack of lab technicians or SS test equipment). The objective is then to maximize the following:

$$V_T(\beta_T) = max\left[r'_{DST}\beta_T, r'_{\sim DST}\beta_T\right] \text{ for } t = T$$

$$V_t(\beta_t) = max\left[r'_{DST}\beta_t, r_{wait} + \delta\left(V_{t+1}(f_t\beta_t)\right)\right] \text{ for } t = 0, 1, 2, \ldots, T-1 \quad (5)$$

In this case, r_{wait} does not include the SS test cost. We show that the optimal action is either to administer DST right away or not at all. This result holds when we make some reasonable assumptions about the relative value of DST and treatment to patients of different types (Assumption 1).

Assumption 1—NMB values: Administering DST to a patient with DS or DR TB generates more NMB than administering DST to a healthy patient; administering DST to a healthy patient generates more NMB than having the patient defaulted or dead; and a sick patient is better off in treatment than defaulted or dead.

When Assumption 1 does not hold, there are still conditions under which the optimal action is to administer DST right away or not at all.

Assumption 2—$r_{DST}(I - \delta f_{t-1})\beta_{t-1} \geq r_{wait}$ for all $t = 0, 1, \ldots, T-1$ where I is the identity matrix.

Assumption 2 considers the reward tradeoffs looking one period ahead. Assumption 2 holds when the expected loss in rewards by waiting to administer DST one period later is greater than the stage cost of waiting one period.

Theorem 1—When no observations are available for any time periods and Assumption 1 or 2 holds, the optimal policy is to administer DST immediately to all patients if the population estimate of DR TB prevalence generates initial beliefs that fall in the half-space

$$H_0 = \left\{\beta_0 : \left(r_{DST} - \delta^T r_{\sim DST} f_{T-1} \cdots f_0\right)\beta_0 \geq r_{wait} \sum_{i=0}^{T-1} (\delta^i)\right\}$$

Otherwise, DST should never be administered.

Theorem 1 provides a rule to guide policymakers in determining whether to administer DST to patients immediately – or not at all – in situations where no monitoring is available. Since no additional information is acquired during the treatment process, it is intuitive that the optimal action is to either administer DST right away or never. If the initial belief lies in the region H_0 as given in the theorem statement, then the optimal action is to administer DST right away.

In practice, the belief about a patient's health state at treatment initiation should be consistent with estimates of DR TB prevalence in that geographical area: thus, $\beta_0 = [0, 1 - b_{DR}, b_{DR}]'$ where b_{DR} is set equal to the regional (or national) estimate of DR TB

prevalence among first-line treatment patients. Theorem 1 tells us that there is a threshold value of β_0 – and in particular b_{DR} – above which the optimal policy is to administer DST right away and below which it is optimal to never administer DST. This threshold value depends on the rewards r_{DST}, $r_{\sim DST}$, r_{wait}, the discount factor δ, and treatment dynamics f_0, f_1, …, f_{T-1}. The rewards for administering DST (r_{DST}), for allowing an additional month of first-line treatment (r_{wait}), and for never administering DST ($r_{\sim DST}$) depend on factors such as treatment costs and willingness-to-pay.

3.2 Optimal actions when SS tests are available

We next examine the case when SS tests results are available for every time period. In the conventional POMDP model formulation [22], an action is taken, a transition occurs based on this action, and then an observation is seen. For our problem, if the decision at the beginning of the month is to wait (and thus continue first-line treatment), the patient may be cured or remain with DS TB or DR TB, and SS+ or SS- observations will be obtained. If the decision is to administer DST, the patient will be triaged into the appropriate treatment (no SS tests obtained or needed). Using this classical formulation in which observations are always obtained after every decision to continue first-line treatment, we have the following objective function:

$$V_T(\beta_T) = max\left[r'_{DST}\beta_T, r'_{\sim DST}\beta_T\right]$$

$$V_t(\beta_t) = max\left[r'_{DST}\beta_t, r_{wait} + \delta \underset{z \in Z}{\mathbb{E}}\left(V_{t+1}(\beta_t, z_t, f_t)\right)\right] \text{ for } t = 0, 1, 2, \ldots, T-1 \quad (6)$$

For example, at time $T-1$, the objective function is:

$$V_{T-1}(\beta_{T-1}) = max\left[r'_{DST}\beta_{T-1}, r_{wait} + \delta \underset{z \in Z}{\mathbb{E}}\left(V_T(\beta_t, z_t, f_t)\right)\right] \quad (7)$$

$$= max\Big[r_{DST}(\beta_{T-1}), r_{wait} + \delta\big(pr(+|f_{T-1}\beta_{T-1})max[r_{DST}\overline{M}_+(f_{T-1}\beta_{T-1}), r_{\sim DST}\overline{M}_+(f_{T-1}\beta_{T-1})] + pr(-|f_{T-1}\beta_{T-1})max[r_{DST}\overline{M}_-(f_{T-1}\beta_{T-1}), r_{\sim DST}\overline{M}_-(f_{T-1}\beta_{T-1})] + pr(default|f_{T-1}\beta_{T-1}) + pr(die|f_{T-1}\beta_{T-1})\big)\Big]$$

We first identify conditions under which the set of beliefs for which DST is optimal is a half-space. A threshold solution will occur when r_{wait} is within a particular range. Let I be the identity matrix.

Theorem 2—*Threshold solution. If Assumption 1 holds and* $r_{wait} \leq (r_{DST} - \delta(r_{DST}M_+ - K_{t+1}M_- + C_{t+1})f_t)\beta_t$ *holds at* $\beta^* = [1, 0, 0, 0, 0, 0, 0]'$ *and* $[0, 1, 0, 0, 0, 0, 0]'$*, then there exists some vector* A_t *and scalar* k_t *at time t such that the optimal action is to administer DST only if* $A\beta_t \geq k_t$.

Here, $K_{t+1} = \sum_{k=t+1}^{T-1} \delta^{T-1-k} \mathbf{r}_{wait} + \delta^{T-t-1} r_{\sim DST} \prod_{i=t+1}^{T-1} f_i$ is the discounted future reward vector for not administering DST and $C_{t+1}\widetilde{\beta}_t = \sum_{k=t+1}^{T-1} \delta^{T-1-k} (r_{\sim DST} Q + r_{\sim DST} S)\widetilde{\beta}_t$ is the future reward from observing death or default at time t, where S is a 7×7 matrix of zeros with a single 1 in the lower right corner (thus $S\beta = [0, 0, 0, 0, 0, 0, \beta(7)]$, where $\beta(7)$ is the seventh element of the β vector).

Theorem 2 shows that if Assumption 1 and a condition on the expected rewards when the patient definitely does not have DR TB are true, we can draw a line on the belief space that separates beliefs where one should administer DST and beliefs where one should wait. This means that the decision rule to administer DST can be expressed simply as a linear equation and that there is at most a single threshold at which the optimal action switches if one moves from an initial belief to another belief.

We know that if the decision maker has the additional option of waiting without administering an SS test, the NMB cannot decrease. The decision maker may not wish to take observations, particularly if they are costly and/or if they do not provide information that informs the treatment plan. For which beliefs do test results not change future decisions?

We will show that if a linear function separates the beliefs at time t for which it is optimal to administer DST from those beliefs where it is optimal to wait, as we discuss in Theorem 2, then the set of beliefs at time $k < t$ for which SS test observations are useless for determining the optimal decision at time t lie in two convex regions of the belief space that must include at least one edge and one corner of the belief space, and the region has linear borders. This implies that linear combinations of beliefs in this region must also be in the region, and "stronger" beliefs (those closer to the corner within the region) will also be in the region.

Our results in this section are illustrated graphically in Figure 1. We define $S_{t-1, H_t, z_{t-1}}$ as the set of beliefs at time $t-1$ that map to half-space H_t at time t with dynamics $f_t\beta_t$ and observation z_{t-1}. Figure 1 panel a schematically illustrates half-space H_t, a belief region where the optimal action at time t is to administer DST, and half-space $\sim H_t$, the belief region where the optimal action at time t is to continue on first-line treatment.

In the appendix, we prove (Proposition 3) that $S_{t-1, H_t, z_{t-1} = SS+}$ and $S_{t-1, H_t, z_{t-1} = SS-}$ are both half-spaces, which implies that a half-space of beliefs will update onto a half-space after disease dynamics and accounting for the SS observation: these beliefs all lie in a convex region of the belief space and can be represented using a linear inequality. Then we prove (Proposition 4) that one of $S_{t-1, H_t, z_{t-1} = SS+}$ or $S_{t-1, H_t, z_{t-1} = SS-}$ is a subset of the other, as illustrated in Figure 1 panels b and c.

This allows us to define $U_{k,t}$ as the set of beliefs at time $k < t$ such that the optimal action at time t will be the same regardless of an SS+ or SS− observation at times k, $k+1$, …, t.

These are the beliefs at time k where observations do not add value to the decision at time t and thus will not change the optimal action at time t. For instance, suppose the optimal action at time t is to administer DST for a half-space set of beliefs $H_t = \{\beta_t: C\beta_t \geq 0, \mathbf{1}\beta_t = 1\}$ and to wait for the complementary set $\sim H_t$. The set $U_{t-1,t}$ is then $(S_{t-1,H_t,SS+} \cap S_{t-1,H_t,SS-}) \cup (S_{t-1,\sim H_t,SS+} \cap S_{t-1,\sim H_t,SS-})$ (Figure 1 panel d). S_{t-1,H_t} and $S_{t-1,\sim H_t}$ are both half-spaces on the belief space.

Theorem 3—If the set of beliefs at time t where DST is the optimal action is a half-space, then $U_{k,t}$ for $k < t$ can be represented by the union of two half-spaces.

Theorem 3 tells us that the set of beliefs at time $k < t$ for which SS test observations are useless for determining the optimal decision at time t lie in two 'nice' regions of the belief space: each region is convex, must include at least one edge and one corner of the belief space (must include 'more certain' beliefs), and has linear borders.

If a belief at time k lies within $U_{k,t}$ for all $t > k$, it does not inform the optimal action for any future time period, and observations should not be taken. We define U_k as the set of beliefs at time k for which observations will not inform any future action.

Theorem 4—*If the set of beliefs where DST is the optimal action is a half-space for all times* $t = \{0, 1, 2, \ldots, T\}$, *then* U_k *is the intersection of* $U_{k,k+1}$, $U_{k,k+2}$, …, $U_{k,T}$, *and is the union of at most* 2^{T-k} *distinct convex regions.*

Theorem 4 provides properties of U_k, the set of beliefs at time k for which observations should not be taken. These beliefs are the union of possibly one or more convex regions (i.e., U_k can be discontinuous). Interestingly, this implies that even if a belief lies between two other beliefs known to be in U_k, it may not necessarily be in U_k, and it may be useful to collect information despite not needing to collect information for the beliefs it lies between.

4 Empirical example: India

4.1 Model and inputs

In this section we apply our model to the example of patients on first-line TB treatment in India. We assume that after the last month of treatment, all patients on first-line treatment (those in states 1, 2, 3) enter states 4, 5, or 6 according to their health state, as they have completed the treatment regimen.

From our theoretical analyses, we know that an additional month of treatment without an SS test can be valuable. This is the default action, which is to follow standard guidelines and continue first-line treatment to wait and see how the patient responds. We include this option in the action space, which becomes $A(s)_t = \{DST, \textit{wait and see}, \textit{wait and not see}\}$ for states 1, 2, and 3, and $A(s)_t = \{\textit{wait and not see}\}$ for all other states (no action is possible if the patient already had a DST test or has defaulted from treatment (states 4, 5, 6) or has died (state 7)). The observation space becomes {*SS+, SS–, default, death, no observation taken*}.

The sequence of events is illustrated schematically in Figure 2. At the beginning of period t, the physician has a belief β_t about the patient's state and chooses to administer DST or have the patient continue on first-line treatment, with or without an SS test. During that period, patients who are on first-line treatment may be cured, continue with DS or DR TB, default from treatment, or die. If the action in the previous month included administering an SS test, then an SS test is administered at the end of the month for those patients who have not defaulted or died. The physician then calculates an updated belief β_{t+1} about the patient's state and, for patients still on first-line treatment, the physician will select a new action a_{t+1} based on the updated belief.

We assume that at treatment initialization (time 0) the patient has either DS or DR TB, with the probability of DR TB equal to the national estimate for the fraction of new cases of TB in India that are drug-resistant [4]: thus, $\beta_0 = [0, 0.97, 0.03, 0, 0, 0, 0]'$. We collapse the belief set to only feasible beliefs given the prior belief at time 0 and solve the resulting dynamic program that has a finite number of beliefs. Because the state space, observation space, and number of periods are small, we use value iteration to find the optimal action associated with each feasible belief.

We obtained transition probabilities for the model from the literature (Table 1 and Appendix Table 2). We used published cost and QALY values to estimate one-month rewards (Appendix Table 3). All costs are reported here in 2015 USD. Costs and health outcomes were calculated from a published and validated microsimulation model of TB in India [14], which we used to generate discounted lifetime costs and QALYs for the terminal rewards (Appendix Table 3). The WHO defines a very cost-effective health intervention as one that has positive NMB with λ equal to the GDP per capita in the country under consideration [23], so we used this λ value in our analysis for India. We also examined the financial costs of both the status quo and optimal policies (to do so, we assigned a value of zero to QALYs when calculating NMB).

TB transmission in India may occur at high levels in densely populated urban areas but may be close to non-existent in rural communities where few new contacts happen. Significant numbers of people live in both environments in India, so we examined both cases. In particular, we examined a scenario where TB transmission occurs at nationally estimated levels and a scenario where no TB transmission occurs. In the scenario with transmission, we hold the probability of transmission constant, as the optimal policy that we identify applies only to patients on first-line treatment – a small fraction of the total TB cases in India – and is unlikely to change national transmission rates.

Patients on appropriate treatment are generally non-infectious, so for the case in which we considered the effects of TB transmission we assumed that transmission costs and QALY losses were only incurred for patients defaulting from treatment or DR TB patients on first-line treatment. We estimated these costs and QALY values for each age group and sex using a previously published microsimulation model [14] as follows: we exposed a population of 3 million susceptible individuals to 8000 infectious TB cases (reflecting the overall observed TB prevalence level in India [4]) of the specified age and sex for one month and found the net present per-case cost and QALY differences over the next 100 years compared to a

scenario where the exposure did not occur. Over the 100 years, individuals transitioned between health and treatment states and could potentially infect others according to the validated microsimulation transition probabilities, which reflect national estimates for India [14]. Finally, we obtained stable estimates of transmission costs and QALY losses by averaging over 10 such runs for each age- and sex-specific exposure to reduce the effects of stochastic noise.

The model was implemented and solved in MATLAB, 2015a (The Mathworks, Natick, MA).

4.2 Results: optimal testing policy

We find that India's current testing policy, in which patients undergo DST only if they have a positive SS test after four months or at the end of first-line treatment, is suboptimal given India's relatively high national estimates of DR TB prevalence and transmission. For areas where TB transmission occurs at nationally representative levels, the optimal policy is to administer DST to all patients upon initial TB diagnosis, before their first month of treatment. While many more patients undergo expensive DST under the optimal policy compared to India's current policy, the reduction in downstream costs and increase in health due to averted downstream TB transmission makes the optimal policy superior.

For regions of India where it is reasonable to assume there is no TB transmission, the optimal testing protocol is given in Figure 3. The optimal action is to wait without taking SS test observations in the first two months, demonstrating that there are belief regions in which it is advantageous to not collect costly observations, as discussed in our theoretical results. SS tests are administered only after the third, fourth, and fifth months. While the majority of patients (80%) are expected to test negative on all three SS tests, patients with a positive SS test after the third month of treatment should be given an additional month of first-line treatment with no monitoring, and then undergo DST. This is because some DR TB cases may be cured on first-line treatment during the fourth month of treatment, and it is cost-effective to attempt to cure them on first-line treatment despite the risk of delaying DR TB treatment. Patients with a positive SS test after the fourth or fifth month of treatment should immediately undergo DST. Patients with a negative SS test after the fifth month of treatment should finish first-line treatment without further SS testing. This policy incorporates DST earlier and gives patients more chances to test positive than India's current policy, where patients undergo DST only if they have a positive SS test after the fourth month or at the end of treatment (thus no chance of DST after the third or fifth month). A total of 7.9% of all patients who start treatment are expected to undergo DST under the optimal testing protocol.

Expected NMB for an individual patient is shown in Table 2 for three different policies: the status quo (India's current testing policy), the optimal policy (for the case of no TB transmission and the case of average TB transmission, as described above), and a policy which assumes that DST is not available at diagnosis but is administered after the first month of treatment. For a value of λ equal to India's GDP per capita, the NMB of the optimal policy when there is no TB transmission is larger, although very similar, to the NMB of India's current policy ($6368 vs. $6365). This is not surprising, as the optimal testing algorithm for regions of India with insignificant levels of TB transmission is somewhat

similar to India's current policy. Both policies have positive incremental NMB, which means that they would be considered cost-effective for this value of λ (see appendix for further discussion of the relationship between NMB and cost-effectiveness). However, when TB transmission rates are at levels similar to India's national average, India's current policy is not cost-effective (the NMB is negative) due to the costs and health effects associated with potential downstream transmission, nor is it optimal. In this case, the optimal policy (to administer DST at the time of TB diagnosis) would generate more than $7000 in NMB per patient compared to the status quo ($6335 vs. −$1157).

The rightmost columns of Table 2 show the costs of each policy in an environment with and without TB transmission; with $\lambda = 0$, NMB incorporates only costs. While the status quo and optimal policies with and without TB transmission are not cost-saving (the NMB is negative), the optimal policy when there is TB transmission would save almost $3000 in health expenditures per patient compared to the status quo ($1971 vs. $5015) due to reduced downstream TB transmission. This is particularly significant given the hundreds of thousands of patients who enter first-line TB treatment in India every year.

In some circumstances, it may not be possible to administer DST at initial diagnosis: for example, if the initial diagnosis site does not have the appropriate equipment or lacks properly trained personnel. In this case, the optimal policy in areas with average levels of TB transmission is to administer DST to patients as soon as possible, which we assume would be after the first month of treatment. For this policy, the NMB declines by approximately $1900 per patient to $4406, but is still a significant improvement over the status quo. Similarly, costs are larger than those of the optimal policy ($2752 vs. $1971), but smaller than those of the current policy ($5015).

Table 3 shows the additional annual number of SS and DST tests, NMB, and incremental cost for the optimal testing policy in regions with and without significant levels of TB transmission. Using 2014 data from India, we estimate that approximately 1,174,000 SS tests and 30,000 DST tests are administered annually to first-time TB treatment patients with smear-positive pulmonary TB [24]. Under the optimal policy without transmission, these values would increase to 1,623,000 and 52,000 tests, respectively (thus, 449,000 and 22,000 incremental tests, respectively). In the case with transmission, each patient would not receive any SS tests and would only receive one DST test, totaling 628,000 DST tests per year (thus 598,000 additional tests). However, while the numbers of SS and DST tests are higher than the status quo in both cases, the annual net monetary benefits increase substantially.

In regions with transmission, the optimal policy requires spending an additional $16M annually on administering additional DSTs compared to the status quo ($598,000 \times \$26.57$). While this expense is substantial, this policy generates a net savings ($1.89 billion in net present dollars) once averted TB cases are taken into account: preventing the transmission of DR TB averts large future TB treatment costs.

Although the optimal policy would ultimately be cost-saving for regions with transmission, it would require a substantial initial investment to implement, and initial per-person total costs would be higher than long-run per-person costs. In the appendix we explore the

amount of time needed for an upfront investment in Xpert testing to break even given the cost savings from onward transmission. If total costs per patient to administer Xpert at the optimal time are $100, and 3% of patients have DR TB, it would take between 1–3 years to recover the original investment; at $500 per patient, it would take 5–10 years to recover the initial investment. These time frames are within the typical planning periods for TB control divisions, which often publish strategic 5- or 10-year plans (e.g., see [5]). In practice, a staggered roll-out where investment in a few Xpert systems is made every few years until the majority of patients are covered could be feasible.

In some sparsely populated regions it may be difficult to identify whether DR TB transmission occurs and which policy should be used. We find that in most cases, early DST is preferable. If DST is used at initial diagnosis in a region with no transmission, an additional $22 per patient is spent compared to the optimal no-transmission policy (and $30 more than the status quo policy), reducing NMB by $33 per patient ($30 less than the status quo). In contrast, if the no-transmission optimal policy is used in a region with transmission, the policy costs $2820 extra per patient, mainly due to costly downstream transmission, and generates $6926 less per patient in NMB. While this is still an improvement over using the status quo policy in transmissive environments (the status quo would cost $3044 extra per patient and generate $7492 less NMB per patient), it would be better to err on the side of using the transmission-optimal policy when the level of transmission is uncertain as the early DST policy is less suboptimal when used in an inappropriate region. Indeed, only if the probability of transmission is lower than about 0.5% would using the optimal policy without transmission have higher expected NMB than using the optimal policy for regions with transmission. This is underscored by our finding that early DST is optimal even for low levels of transmission, as described in the sensitivity analysis section below.

4.3 Sensitivity analysis

We performed extensive sensitivity analysis on all model parameters. In this section we describe parameters to which the optimal testing policies were most sensitive.

We find that the optimal policy for regions with transmission is generally robust to changes in uncertain parameters while the optimal policy for regions without transmission is more sensitive. For instance, variation in the timing and probability of DR cure on first-line treatment can change the optimal time to administer DST in cases when there is no transmission, but DST at initial diagnosis remains the optimal policy when there is TB transmission even with unrealistically high values for the probability of DR cure (90% of DR patients cured in one month of treatment).

We examined how the optimal policy changes due to age-related variations in death and default rates for patients on first-line treatment. Using data from the Indian state of Bihar, and reported national TB program performance measures [3,26], we estimated that approximately 5.6% of patients who enter first-line treatment will die while on treatment and 5.6–12.6% will default, with variation over age and sex; this is consistent with national estimates that report 4.1% of new smear-positive patients died while on treatment and 5.8% defaulted [3]. When we use age- and sex-specific death and default rates (rather than averages as in the baseline analysis), we find that the optimal policy in regions with average

TB transmission is unchanged, whereas the optimal policy for regions without TB transmission can change. Appendix Figure 1 shows the optimal testing policy by age and sex for regions without TB transmission. Compared to the baseline policy, older individuals, who have higher death and default rates and who have less to gain from DST and DR TB treatment, are not monitored as often using the SS test (the SS test is omitted at month 4, then at month 2, and then entirely as age increases) and correspondingly are not identified as needing DST.

Transmission costs and the proportion of patients entering first-line treatment with DR TB are particularly uncertain and can vary by region. Two-way sensitivity analysis on these parameters shows that the optimal policy is to administer DST at initial diagnosis even with very small transmission costs and low DR TB prevalence (Appendix Figure 2). Specifically, DST should be administered at initial diagnosis if the transmission cost is above $422 (which is approximately 4% of the $10,550 average transmission cost calculated by the microsimulation) or if DR TB prevalence among patients on first-line treatment is 0.9% or greater (which is approximately one-third as high as the 3.0% average DR TB prevalence in India). Below these values, the optimal action depends on a patient's SS test result history, as illustrated in the optimal policy when there are no transmission costs. We conclude from this sensitivity analysis that if there is doubt about whether DR TB prevalence among first-line patients is zero, DST should be administered as soon as possible.

Our baseline analysis assumed that the DST test is 100% sensitive and specific. In sensitivity analysis we reduced these test characteristics to the lower bounds of the 95% confidence interval given by a WHO survey of Xpert performance in regions around the world (97% sensitivity, 90% specificity) [27]. We estimated cost and QALY parameter values for patients who are misdiagnosed (see appendix). We find that decreased Xpert sensitivity and specificity does not change the optimal policy in regions with TB transmission: physicians should still administer DST as soon as possible, at initial diagnosis. However, in regions without transmission, additional confirmation of DR TB is required before a patient undergoes DST compared to the baseline case: if negative SS tests are observed after the second and third months, the patient completes treatment without further testing, but if a positive test is observed, observations should be taken after the fourth, fifth, and sixth months of treatment, at which point any positive SS test triggers DST at the next available opportunity (Appendix Figure 3). With perfect DST outcomes, one positive SS test triggers the decision to administer DST, whereas with imperfect DST sensitivity and specificity, two positive SS tests must occur before it is optimal to administer DST (one after months 2 or 3, and one after month 4).

5 Discussion

In our theoretical analysis, we found that under some conditions when observations are not available, there is a half-space H_t at each time period such that the optimal action is to administer DST if β_t is in that half-space and to never administer DST otherwise. When the prior belief at treatment initiation – which may reflect national estimates of DR prevalence – falls in H_0, the optimal action is to administer DST at initial diagnosis.

We also found that when observations are always taken, there are two half-spaces of beliefs for which observations will be uninformative for determining the optimal action at time t in the future. The intersection of all these half-spaces for all times t greater than the current time forms a union of convex regions where observations will not be useful for determining decisions at any time. This result furthers our understanding of the value of information over repeated decisions, as it shows that costly information should not be collected on certain belief regions (a union of convex belief sets with linear borders). Interestingly, this implies that even if a belief lies between two other beliefs for which no information should be collected, it may still be useful to collect information despite not needing to collect information for the beliefs it lies between.

In our empirical analysis for India, we found that in regions with TB transmission at nationally representative levels, the optimal action is to administer DST to all patients as soon as possible after initial TB diagnosis, before their first month of treatment. We project that this would save more than $3000 per patient on first-line TB treatment compared to India's current testing protocol, due to health care savings from averted downstream TB transmission. This result demonstrates the importance of incorporating transmission effects into policies for individualized treatment of infectious diseases: externalities from transmission can have an important impact on the resulting societal benefit. If transmission rates are known to be close to negligible, the optimal policy is SS-test-history dependent with SS tests taken after the second, third, and fourth months of treatment. In this case, information from the repeated SS tests allows decision makers to separate probable-DR patients from cured and DS patients, thus reducing unnecessary DST testing.

Our proposed policy of administering DST to patients at diagnosis in regions with non-negligible TB transmission has significant budget implications. Expansion of DST testing will increase the up-front costs of diagnosis. Additionally, to implement such a policy, equipment would have to be purchased and personnel trained. Implementing the optimal policy nationwide would require substantial capital investment and might not be affordable, even though such a policy is ultimately cost-saving. For this reason, India may wish to gradually roll out the change, prioritizing areas that are likely to have high DR TB prevalence and transmission, such as densely populated urban environments.

We note that, even with broad implementation of the optimal DST testing policy, DR TB transmission may not decrease significantly. Many DR TB patients in India do not enter into federal, first-line TB treatment and thus would remain outside the reach of these improvements [30]. Therefore, implementation of the optimal testing policy would likely not be sufficient to control the growing DR TB epidemic in India, although it would be an important positive step in that direction.

Our analysis has several limitations. Our numerical analysis is limited to first-line treatment in the federal Indian TB treatment program, and TB patients in India may use other treatment programs (private clinics, second-line treatment). However, federally funded, first-line TB treatment remains an important regimen for analysis as many patients use this service and it may be the first time TB patients are exposed to World Health Organization-approved TB treatment regimens. We do not consider false positive tests in the initial

diagnosis that may incorrectly triage healthy patients into first-line TB treatment, as we assume all patients have either DS or DR TB. This is a reasonable assumption as the specificity of the SS test is relatively high (98% [30]).

We assume that no drug resistance is acquired during treatment. Inclusion of this complication could mean that it is optimal to administer DST multiple times (e.g., once at the beginning of treatment and again later on to identify patients with treatment-generated DR) or that it is optimal to administer DST later if DST can only be performed once. However, since treatment-acquired drug resistance is much more likely to occur in defaulted patients, who are generally lost to follow-up, even expanding the decision space to allow multiple DST times may not improve NMB significantly from only allowing one DST.

We also do not consider resistance to DR TB treatment, as information on acquiring such TB strains is lacking. The NMB and costs of the policies with transmission may decline in the future if the probability of transmission decreases over time due to reductions in demographic and social factors such as crowding and sanitation and improvements in TB treatment. However, our current analysis indicates that significant improvements in NMB and cost savings may accrue with current national estimates of transmission.

The optimal policies we have identified for regions with and without transmission are based on average values observed over age and sex groups, and thus are optimal for patients who behave similarly to their group average. If patients do not behave similarly to their group average, our policies may not be optimal. However, given the robustness of the results in regions with transmission, it is unlikely that deviating from the early DST policy will be beneficial in regions with transmission, even when additional information about patients is available.

Despite these limitations, our findings are robust for typical scenarios for first-line TB treatment in India. For most metropolitan areas in India, where transmission is close to the nationally estimated levels, and where the majority of TB patients are seen, we find that early DST will be a significant improvement on the status quo and could save India up to $1.89 billion annually.

Our model can also be used to determine the optimal timing of drug sensitivity testing for TB in other settings. Such analysis would be particularly valuable for countries in regions such as Southeast Asia, Eastern Europe, and southern Africa with high TB prevalence and where drug resistance is an emerging threat [28]. While our empirical example is country-specific, as it includes measures (such as default, death, diagnosis characteristics) and costs specific to India, our framework can be readily adapted to other settings if the requisite data are available.

Finally, our model could be applied to other diseases where identification of drug resistance during first-line treatment is important. For example, HIV drug resistance is an increasing concern as antiretroviral drugs are used more ubiquitously for pre-exposure prophlyaxis [29], as well as for shorter-term treatments, such as those for hospital-acquired infections. To apply our model to another disease, the simulation model that generated the lifetime costs

and health outcomes would need to be restructured to reflect the natural history of that disease.

APPENDIX

Relationship between Net Monetary Benefits and Incremental Cost-Effectiveness Ratios

We aim to maximize net monetary benefit (NMB), which is defined as NMB = λ * QALYs - Costs. In our analysis for India, we set λ, the willingness-to-pay threshold, equal to India's per capita GDP. The WHO defines per capita GDP as a "very cost-effective" willingness-to-pay threshold [23]. We observe that, if the NMB for one policy is larger than another, that policy is cost-effective relative to the other policy.

Cost-effectiveness is typically defined using incremental cost-effectiveness ratios (ICERs). To show that this is equivalent to our NMB analysis, let Q_i and C_i be the QALYs and costs, respectively, associated with policy *i*. If policy 2 is cost-effective relative to policy 1, its ICER, defined as the ratio of the difference in costs to the difference in QALYs, will be less than the willingness-to-pay threshold, λ. Assuming all costs, QALYs, and differences to be positive, we have:

$$\text{ICER} < \lambda$$

$$\frac{C_2 - C_1}{Q_2 - Q_1} < \lambda$$

$$C_2 - C_1 < \lambda(Q_2 - Q_1)$$

$$\lambda(Q_1) - C_1 < \lambda(Q_2) - C_2$$

$$\text{NMB for policy 1} < \text{NMB for policy 2}$$

In our analysis, we find the NMB for the optimal policy to be larger than that of the status quo. We also find that the absolute costs for the optimal policy are larger than those of the status quo, which implies that the absolute QALYs in the optimal policy are larger since the NMB is larger with the same λ. So all costs, QALYs, and differences are positive, and this relationship between ICERs and NMB holds. Using this rationale, we conclude that the optimal policy is very cost-effective.

Additional discussion of cost-effectiveness thresholds, NMB, and ICERs can be found elsewhere [23,31,33].

Xpert Initial Costs: Breakeven Calculations

Our results indicate that the optimal policy in regions with transmission would be cost-saving due to averted transmission costs. However, widespread implementation of this policy would incur substantial initial startup costs, and initial per-patient costs may be high due to low initial patient volume as the system rolls out. Here we estimate how long it would take to recoup initial costs.

Appendix Figure 4 shows expected cost savings from onward transmission from detecting one DR TB case four months early over the 10 years after the case is detected, broken down by age of the detected case. These values take into account the likelihood of onward transmission, the likelihood of seeking care, treatment default, competing mortality, and other population dynamics using the simulation described in the main manuscript. Over time, the expected cost savings increase, as secondary, tertiary, and higher-order transmissions are averted. The cost savings vary, as some age groups are more likely to contact and transmit TB to others, and different ages have different probabilities of seeking care or dying. In some cases, averting a case of DR TB may not necessarily save costs in the short term, as an early death can also save costs compared to the situation where the patient has a longer life but incurs large medical expenses (this explains the slightly negative values for age group 20 in the first year).

Appendix Table 4 shows the number of years until an initial investment of $100/patient, $500/patient, or $1000/patient to implement Xpert could be recovered through averted costs from downstream transmission, broken down by the age of the detected DR TB case. In most cases, the initial investment cost would be recovered within 10 years, and in many cases within 5 years.

Cost and QALY Calculations for Patients Misdiagnosed by Xpert

We assume that non-DR patients misdiagnosed to DR TB treatment behave similarly to those in first-line treatment: cure, death, and default rates are the same as for non-DR in first-line treatment, with the same long-term health outcomes, in expectation, as DR TB treatment can still cure DS patients. We assign a QALY decrement and cost increase due to misdiagnosis of non-DR patients due to being exposed to DR TB treatment (which is more expensive and uses harsher drugs than first-line treatment). Misdiagnosed DR TB patients, who are incorrectly identified as not having DR TB, continue on first-line treatment and exit treatment with DR TB upon default or treatment completion.

The QALY decrement and cost increase associated with misdiagnosed non-DR patients are calculated as follows. The cost increase is the difference between the cost of DR TB treatment and first-line treatment. The cost of a full DR TB treatment regimen (24 months) is $3649.58 [14], although the average duration of exposure to DR TB treatment is only 19.5 months due to deaths and default (death and default rates taken from [14]). A full first-line treatment regimen costs $183.33. We therefore approximate the expected cost increase due to using DR TB treatment as $2781.95 (= ($3649.58)(19.5/24) − $183.33). We assume that patient QALY values decrease to levels similar to having untreated TB due to the harsh

drugs used for DR TB treatment, which can cause severe side effects. This results in a monthly decrease of ((0.843–0.663)/12), as the yearly QALY value of a TB patient on first-line treatment is 0.843 and without treatment is 0.663. This monthly QALY decrease is applied for 19.5 months, the average number of months of exposure to DR TB treatment, for a lifetime QALY decrement of 0.2925 from being exposed to DR TB treatment.

Proofs

To prove Theorem 1, we first establish several preliminary results (Lemma 1 and Propositions 1 and 2) for patients still on treatment. Lemma 1 establishes that when DST yields greater NMB for patients with DR TB than for patients with DS TB or healthy patients (Assumption 1), rewards associated with DST cannot increase over the course of treatment.

Lemma 1

If administering DST to non-healthy patients generates more NMB than administering DST to healthy patients ($NMB_{DST,DR} \geq NMB_{DST,H}$, $NMB_{DST,DS} \geq NMB_{DST,H}$); administering DST to a healthy patient generates more NMB than having the patient defaulted or dead ($NMB_{DST,H}$ $d_H NMB_{dH}$ and $NMB_{DST,H} \geq NMB_{dead}$); and sick patients are better off in treatment ($NMB_{DR} \geq NMB_{dDR}$, $NMB_{DR} \geq NMB_{dead}$, $NMB_{DS} \geq NMB_{dDS}$, and $NMB_{DS} \geq NMB_{dead}$); then $r_{DST}(f_t\beta_t) < r_{DST}(\beta_t)$.

Proof—We omit the subscript t for $c_{DR,t}$, $c_{DS,t}$, and f_t as we are only concerned with the transition between time t to $t+1$. We prove the result by contradiction.

Suppose $r_{DST}(f\beta) \geq r_{DST}(\beta)$.

$$\Rightarrow \begin{bmatrix} NMB_{DST,H} \\ NMB_{DST,DS} \\ NMB_{DST,DR} NMB_{dH} \\ NMB_{dDS} \\ NMB_{dDS} \\ NMB_{dead} \end{bmatrix}, \begin{bmatrix} 1-d_H-\nu_H & c_{DS} & c_{DR,t} & 0 & 0 & 0 & 0 \\ 0 & 1-c_{DS}-d_{DS}-\nu_{DS} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-c_{DR}-d_{DR}-\nu_{DR} & 0 & 0 & 0 & 0 \\ d_H & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & d_{DS} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & d_{DR} & 0 & 0 & 1 & 0 \\ \nu_H & \nu_{DS} & \nu_{DR} & \nu_{dH} & \nu_{dDS} & \nu_{dDR} & 1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} 1-b_{DS}-b_{DR} \\ b_{DS} \\ b_{DR} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \geq \begin{bmatrix} NMB_{DST,H} \\ NMB_{DST,DS} \\ NMB_{DST,DR} NMB_{dH} \\ NMB_{dDS} \\ NMB_{dDS} \\ NMB_{dead} \end{bmatrix}, \cdot \begin{bmatrix} 1-b_{DS}-b_{DR} \\ b_{DS} \\ b_{DR} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

However, this cannot hold. The weighted sum of $NMB_{DST,H}$, NMB_{dH}, and NB_{dead} must be less than $NMB_{DST,H}$, and the weighted sum of $NMB_{DST,DS}$, $NMB_{DST,H}$, NMB_{dDS}, and

NMB_{dead} must be less than $NMB_{DST,DS}$, and the weighted sum of $NMB_{DST,DR}$, $NMB_{DST,H}$, NMB_{dDR}, and NMB_{dead} must be less than $NMB_{DST,DR}$. Therefore we have reached a contradiction and the assumption must not be true. Thus, $r_{DST}(f_t\beta) < r_{DST}(\beta)$.

Proposition 1

When no observations are available for any time periods and Assumption 1 holds, then the optimal action is to either administer DST immediately or not at all.

Proof—Proof by induction. For a patient still on treatment at time T, we can either administer DST or not ever administer DST. If the optimal action at time T given belief β_T is to administer DST ($V_T = r_{DST}\beta_T$), then at time $T-1$ the value function is:

$$max(r_{DST}\beta_{T-1}, r_{wait} + \delta(V_T(\beta_T))) \Rightarrow max(r_{DST}\beta_{T-1}, r_{wait} + \delta(V_T(f_{T-1}\beta_{T-1}))) \Rightarrow max(r_{DST}\beta_{T-1}, r_{wait} + \delta(V_{DST}(f_{T-1}\beta_{T-1})))$$

Lemma 1 states that $r_{DST}\beta_{t-1} \geq r_{DST}(f_{t-1}\beta_{t-1})$. The term r_{wait} is negative and the discount factor δ is less than one, so the optimal action at time $T-1$ is DST and $V_{T-1} = r_{DST}\beta_{T-1}$. This follows similarly for time $T-2$, $T-3$, …, 0. Therefore if it is optimal to administer DST at all, the best time to do so is as soon as possible, before the first month of treatment.

Corollary 1

If Assumption 1 does not hold, then Proposition 1 can still hold if Assumption 2 holds.

Proof—This ensures that the induction step from the proof of Proposition 1 holds without Lemma 1. DST is the optimal action at time t if:

$$max(r_{DST}\beta_{t-1}, r_{wait} + \delta(r_{DST}(f_{t-1}\beta_{t-1}))) = r_{DST}\beta_{t-1} \Rightarrow r_{DST}\beta_{t-1} \geq r_{wait} + \delta(r_{DST}(f_{t-1}\beta_{t-1})) \Rightarrow r_{DST}(I - \delta f_{t-1})\beta_{t-1} \geq r_{wait}$$

where I is the identity matrix.

Proposition 2

When no observations are available for any time periods and Assumption 1 or 2 holds, then there is a half-space H_t at each time period such that the optimal action is to administer DST if β_t is in that half-space and to wait otherwise. H_t is given by:

$$H_t = \left\{ \beta_t : \left(r_{DST} - \delta^{T-t} r_{\sim DST} f_{T-1} f_{T-2} \cdots f_t\right)\beta_t \geq r_{wait} \sum_{i=0}^{T-t-1} (\delta^i) \right\}$$

Proof—From Proposition 1 or Corollary 1 we know it will never be optimal to wait and then administer DST, so the only choices at any time t are to either administer DST or wait forever. Therefore it will only be optimal to administer DST in time period t if:

$$r_{DST}(\beta_t \geq r_{wait} + \delta(r_{wait} + \delta(r_{wait} \cdots \delta(r_{\sim DST} f_{T-1} \cdots f_t \beta_t))$$
$$= \sum_{i=0}^{T-t-1} (\delta^i r_{wait}) + \delta^{T-t} r_{\sim DST} f_{T-1} \cdots f_t \beta_t \Rightarrow H_t$$
$$= \left\{ \beta_t : \left(r_{DST} - \delta^{T-t} r_{\sim DST} f_{T-1} f_{T-2} \cdots f_t \right) \beta_t \geq \sum_{i=0}^{T-t-1} (\delta^i r_{wait}) \right\}$$

Note that this is a half-space as the left-hand side is linear in β_t, making the expression a single affine inequality.

Proof of Theorem 1—Theorem 1 follows directly from Proposition 2.

Proof of Theorem 2—We use a proof by induction. Initialization: for all patients still on treatment at time T, there is a threshold solution at time T, where the optimal action is to DST if $(r_{DST} - r_{\sim DST})\beta_T \geq 0$.

Induction step: if there is a threshold solution at time t+1, at time t there can be at most five belief regions, on which the optimal actions will be to DST at time t, wait and DST at time t +1, wait and DST at t+1 only if there is a positive observation, wait and DST at t+1 only if there is a negative observation, and wait and not DST (and if death or default is observed, no action is possible).

We first show that if the left inequality in the threshold condition holds, there cannot be a region where the optimal action is to wait and DST at t+1 only if there is a negative observation ("the counterintuitive action"). Let $\mathrm{r}_{wait} = r_{wait}[1, 1, 1, 0, 0, 0, 0]$ be the vectorized form of r_{wait}, and $K_{t+1} = \sum_{k=t+1}^{T-1} \delta^{T-1-k} \mathbf{r}_{wait} + \delta^{T-t-1} r_{\sim DST} \prod_{i=t+1}^{T-1} f_i$ be the discounted future reward vector for waiting (no DST).

The expected future reward from default or death for the post-dynamics belief state $\tilde{\beta}_t$ is $pr(z_t = default | \tilde{\beta}_t)$(future NMB from defaulting) $P(\tilde{\beta}_t) Q \tilde{\beta}_t + pr(z_t = dead | \tilde{\beta}_t)$ (future NMB from dying) $R \tilde{\beta}_t$. This simplifies to (future NMB from defaulting) $Q \tilde{\beta}_t$+(future NMB from dying$R$)S) $\tilde{\beta}_t$, where S is a 7×7 matrix of zeros with a single 1 in the lower right corner (thus $S\beta = [0, 0, 0, 0, 0, 0, \beta(7)]$). The future NMB from defaulting and dying can be found in the appropriate elements of $\sum_{k=t+1}^{T-1} \delta^{T-1-k} r_{\sim DST}$, since $r_{\sim DST}$ contains NMB_{dH}, …, NMB_{dead}. Therefore let $C_{t+1} \tilde{\beta}_t = \sum_{k=t+1}^{T-1} \delta^{T-1-k} r_{\sim DST} Q + \sum_{k=t+1}^{T-1} \delta^{T-1-k} r_{\sim DST} S) \tilde{\beta}_t$, or $\sum_{k=t+1}^{T-1} \delta^{T-1-k} (r_{\sim DST} Q + r_{\sim DST} S) \tilde{\beta}_t$ be all the future rewards from observing death or default at time t.

Then generally, the reward at time t is:

$$r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + \sum_{i=+,-} pr(\text{see i}|\widetilde{\beta}_t)[r_{DST}\overline{M}_i\widetilde{\beta}_t + K_{t+1}\overline{M}_i\widetilde{\beta}_t])$$

$$r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + \sum_{i=+,-} N_i^{-1}[r_{DST}N_iM_i\widetilde{\beta}_t + K_{t+1}N_iM_i\widetilde{\beta}_t])$$

$$r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + \sum_{i=+,-} [r_{DST}M_i\widetilde{\beta}_t + K_{t+1}M_i\widetilde{\beta}_t])$$

So, the reward for wait and not DST is $r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + K_{t+1}(M_+ + M_-)\widetilde{\beta}_t])$, and similarly, the reward for wait and DST at t+1 only if there is a negative observation is $r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + r_{DST}M_-\widetilde{\beta}_t + K_{t+1}M_+\widetilde{\beta}_t)$, and the reward for wait and DST at time t+1 is: $r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + r_{DST}(M_+ + M_-)\widetilde{\beta}_t)$.

Using a proof by contradiction, we will show that the reward for the counterintuitive action cannot be better than both the rewards for wait and DST regardless of observation or wait and not DST. Therefore we begin by assuming the reward for the counterintuitive action is better than both.

The reward for the counterintuitive action is better than wait and not DST, so:

$$r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + r_{DST}M_-\widetilde{\beta}_t + K_{t+1}M_+\widetilde{\beta}_t) > r_{wait} + \delta(C_{t+1}\widetilde{\beta}_t + K_{t+1}(M_+ + M_-)\widetilde{\beta}_t) \Rightarrow r_{DST}M_-\widetilde{\beta}_t + K_{t+1}M_+\widetilde{\beta}_t > K_{t+1}(M_+ + M_-)\widetilde{\beta}_t \Rightarrow (r_{DST} - K_{t+1})M_-\widetilde{\beta}_t > 0 \Rightarrow (r_{DST,H} - K^1_{t+1})(spec)\widetilde{\beta}^1_t + (r_{DST,DS} - K^2_{t+1})(1-sens)\widetilde{\beta}^2_t + (r_{DST,DR} - K^3_{t+1})(1-sens)\widetilde{\beta}^3_t > 0$$

where the superscripts denote the element number. If the left inequality in the threshold condition holds, then $r_{DST,H} - K^1_{t+1} \leq 0$, and $\widetilde{\beta}^1_t \geq 0$, and $1 \leq spec > 0.5$ so $1-sens < spec$ and is positive, so $(r_{DST,H} - K^1_{t+1})(spec)\widetilde{\beta}^1_t \leq (r_{DST,H} - K^1_{t+1})(1-sens)\widetilde{\beta}^1_t$. Then:

$$(r_{DST,H} - K^1_{t+1})(1-sens)\widetilde{\beta}^1_t + (r_{DST,DS} - K^2_{t+1})(1-sens)\widetilde{\beta}^2_t + (r_{DST,DR} - K^3_{t+1})(1-sens)\widetilde{\beta}^3_t > 0 \Rightarrow (r_{DST} - K_{t+1})(1-sens)\widetilde{\beta}_t > 0 \Rightarrow (r_{DST} - K_{t+1})\widetilde{\beta}_t > 0$$

Also, the reward for the counterintuitive action is better than the reward for wait and DST regardless of observation, so:

$$r_{wait} + \delta\left(C_{t+1}\tilde{\beta}_t + r_{DST}M_-\tilde{\beta}_t + K_{t+1}M_+\tilde{\beta}_t\right) > r_{wait} + \delta\left(C_{t+1}\tilde{\beta}_t + r_{DST}\left(M_+ + M_-\right)\tilde{\beta}_t\right)$$
$$\Rightarrow \left(r_{DST,H} - K_{t+1}\right)(1 - spec)\tilde{\beta}_t^1 + \left(r_{DST,DS} - K_{t+1}^2\right)(sens)\tilde{\beta}_t^2 + \left(r_{DST,DR} - K_{t+1}^3\right)(sens)\tilde{\beta}_t^3 < 0$$
$$\Rightarrow \left(r_{DST} - K_{t+1}\right)\tilde{\beta}_t < 0$$

using a very similar procedure but noting that $1 - spec < sens$. However, $\left(r_{DST} + K_{t+1}\right)\tilde{\beta}_t$ cannot be both greater than and less than zero, so there is a contradiction and our assumption is false.

This shows that the counterintuitive action cannot be the optimal action. We next show that DST at time t is better than waiting to DST at time t+1. Since $M_+ + M_-$ equals the identity, it must be that $r_{DST}\beta_t \geq r_{wait} + \delta r_{DST} f_t\beta_t$ since r_{wait} is negative, δ is less than 1, and from lemma 1 we know $r_{DST}\beta_t > r_{DST} f_t\beta_t$.

Finally, we need to find a condition such that the reward at DST at time t is greater than the reward for waiting and only DST after a positive observation.

$$r_{DST}\beta_t \geq r_{wait} + \delta\left(C_{t+1}\tilde{\beta}_t + r_{DST}M_+ f_t\beta_t + K_{t+1}M_- f_t\beta_t\right) \Rightarrow r_{wait} + \delta K_{t+1}M_- f_t\beta_t + \delta C_{t+1} f_t\beta_t$$
$$\leq \left(r_{DST} - \delta r_{DST}M_+ f_t\right)\beta_t \Rightarrow r_{wait} \leq \left(r_{DST} - \delta\left(r_{DST}M_+ - K_{t+1}M_- + C_{t+1}\right)f_t\right)\beta_t$$

We only need this last inequality to hold on the "corners" of the feasible space. This is because if it is true at [1,0,0,0,0,0,0]' and [0,1,0,0,0,0,0]' it holds for all beliefs since Assumption 1 ensures DST gives the largest reward at [0,0,1]' and the reward $r_{DST}\beta_t$ is a plane on the belief space.

This means that if this condition is met, the only possible action regions left are to DST at time t or to wait and not DST, and therefore there must exist a threshold solution: it is optimal to DST if $A_t\beta_t \geq k_t$ where $A_t = (r_{DST} - \delta K_{t+1} f_t)$ and $k_t = r_{wait}$.

We now establish several preliminary results that we use to prove Theorem 3.

Proposition 3

For a patient who is still on treatment at time t, beliefs on a half-space at time $t-1$ map to a half-space at time t after accounting for disease dynamics and sputum smear observations, and vice versa (beliefs on a half-space at time t can only be reached by beliefs on a half-space at time $t-1$).

Proof—Suppose a belief $\beta_{t-1} = [x_1\ x_2\ x_3\ 0\ 0\ 0\ 0]'$ is in a half-space S, defined as all β_{t-1} such that $A\beta_{t-1} \leq 0$ for some 3×1 vector $A = [a_1\ a_2\ a_3\ 0\ 0\ 0\ 0]$. Since we are concerned with only patients who are in treatment at time t, and must therefore have been in treatment at time $t-1$, there is no possibility of the patient being in states 4–7. We therefore restrict our notation to show states 1 to 3 for clarity and omit the zeros in the states and transitions associated with states 4 to 7. The belief state therefore becomes $\beta_t = [1 - b_{DS} - b_{DR},\ b_{DS},\ b_{DR}]$.

Note that we can use the homogeneous inequality $A\beta_{t-1} \leq 0$ to express the half-space (instead of $\bar{A}\beta_{t-1} \leq c$ for 3×1 vector $\bar{A}$ and scalar c) without loss of generality. We can express any $\bar{A}\beta_{t-1} \leq c$ as $A\beta_{t-1} \leq 0$: we know that $\mathbf{1}\beta_{t-1} = 1$ since β_{t-1} is a belief and its elements must sum to one. Subtracting c from both sides, we see that $[\bar{a}_1 - c\ \bar{a}_2 - c\ \bar{a}_3 - c]\ \beta_{t-1} \leq 0$, and we define $[a_1\ a_2\ a_3] = [\bar{a}_1 - c\ \bar{a}_2 - c\ \bar{a}_3 - c]$. We therefore use the homogeneous inequality throughout.

At time t, the updated belief will be $\overline{M}_+(f_{t-1}\beta_{t-1})$ if we observe an SS+ test result, or $\overline{M}_-(f_{t-1}\beta_{t-1})$ if we observe an SS- test result. Recall that

$$\left(\beta_t | f_{t-1}\beta_{t-1}, z_{t-1} = SS+\right) = \overline{M}_+\left(f_{t-1}\beta_{t-1}\right) = N_+\left(f_{t-1}\beta_{t-1}\right) M_+ f_{t-1}\beta_{t-1}$$

where $N_+(f_{t-1}\beta_{t-1})$ is a positive scalar, and M_+ and f_{t-1} are matrices:

$$N_+\left(\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}\right) = \frac{1}{(1-spec)p_1(sens)\left(p_1 + p_3\right)}$$

$$M_+ = \begin{bmatrix} (1-spec) & 0 & 0 \\ 0 & (sens) & 0 \\ 0 & 0 & (sens) \end{bmatrix}$$

$$f_{t-1} = \begin{bmatrix} 1 & c_{DS,t-1} & c_{DR,t-1} \\ 0 & 1-c_{DS,t-1} & 0 \\ 0 & 0 & 1-c_{DR,t-1} \end{bmatrix}$$

Therefore we have:

$$M_+^{-1} = \begin{bmatrix} \frac{1}{1-spec} & 0 & 0 \\ 0 & \frac{1}{sens} & 0 \\ 0 & 0 & \frac{1}{sens} \end{bmatrix}$$

$$f_{t-1}^{-1} = \begin{bmatrix} 1 & \frac{c_{DS,t-1}}{c_{DS,t-1}-1} & \frac{c_{DR,t-1}}{c_{DR,t-1}-1} \\ 0 & \frac{1}{1-c_{DS,t-1}} & 0 \\ 0 & 0 & \frac{1}{1-c_{DR,t-1}} \end{bmatrix}$$

Letting $C = Af_{t-1}^{-1}M_{+}^{-1}$, we see that:

$$
\begin{aligned}
& A\beta_{t-1} && \leq 0 && \text{Start with the half-space in time } t-1 \\
\Leftrightarrow\ & Af_{t-1}^{-1}M_{+}^{-1}M_{+}f_{t-1}\beta_{t-1} && \leq 0 && \text{Substitute } AI = A\left(f_{t-1}^{-1}M_{+}^{-1}M_{+}f_{t-1}\right) \\
\Leftrightarrow\ & CM_{+}f_{t-1}\beta_{t-1} && \leq 0 && \text{Use definition of } C \\
\Leftrightarrow\ & N_{+}(f_{t-1}\beta_{t-1})CM_{+}f_{t-1}\beta_{t-1} && \leq (N_{+}(f_{t-1}\beta_{t-1}))(0) && \text{Multiply both sides by positive scalar} \\
\Leftrightarrow\ & CN_{+}(f_{t-1}\beta_{t-1})M_{+}f_{t-1}\beta_{t-1} && \leq 0 && \text{Move scalar} \\
\Leftrightarrow\ & C\beta_{t} && \leq 0 && \text{Use definition of } (\beta_t|\beta_{t-1},\ z_{t-1} = SS_{+})
\end{aligned}
$$

Thus, beliefs at time $t-1$ in half-space S described by $\{\beta_{t-1}\colon A\beta_{t-1}\ 0\}$ will map to beliefs in half-space H_t at time t described by $\{\beta_t\colon C\beta_t \leq 0\}$, where $C = Af_{t-1}^{-1}M_{+}^{-1}$, given an SS+ observation, and a half-space at time t can only be reached after a positive observation by a half-space of beliefs at time $t-1$. The proof is identical for a negative SS test observation except we use N_{-} and M_{-}.

Since the proof is valid in the opposite direction, the opposite is also true: a half-space at time t, $H_t = \beta_t\colon C\beta_t \leq 0\}$ can only be reached by beliefs at time t–1 on a half-space $S_{t-1,H_t,z_{t-1}=SS+} = \left\{\beta_{t-1}\colon A\beta_{t-1} \leq 0\right\}$ where $A = CM_{+}f$ after a positive observation and half-space $S_{t-1,H_t,z_{t-1}=SS-} = \left\{\beta_{t-1}\colon \widetilde{A}\beta_{t-1} \leq 0\right\}$ where $\widetilde{A} = CM_{-}f$ after a negative observation.

Proposition 4

$$S_{t-1,H_t,z_{t-1}=SS+} \cap S_{t-1,H_t,z_{t-1}=SS-} = S_{t-1,H_t,z_{t-1}=SS+} \text{ or } S_{t-1,H_t,z_{t-1}=SS-}.$$

Proof—Since we are again concerned with the set of beliefs for which there are positive or negative SS test results, we again only need to consider patients who are in treatment at time t, and must therefore have been in treatment at time $t-1$. There is no possibility of the patient being in states 4–7, and we therefore restrict our notation to show states 1 to 3 for clarity and omit the zeros in the states and transitions associated with states 4 to 7, as given in the previous proof (Proposition 3).

Suppose we have a set of beliefs at time t in a half-space $H_t = \{\beta_t\colon C\beta_t \leq 0\}$. The line of demarcation $C\beta_t = 0$ intersects at least twice with the "edges" of the feasible belief space $-\begin{bmatrix} k \\ 0 \\ 1-k \end{bmatrix}, \begin{bmatrix} k \\ 1-k \\ 0 \end{bmatrix}$, or $\begin{bmatrix} 0 \\ 1-k \\ k \end{bmatrix}$ – for some value k, i.e., points for which at least one element of the 3×1 vector is 0 (points on the red triangle in Appendix Figure 5). The line $C\beta_t = 0$ can therefore be expressed as the linear combination of any two of these intersection points.

Proposition 3 established that the set of beliefs at time $t-1$ that map to half-space H_t at time t is a half-space itself, which we denote $S_{t-1,H_t,SS+}$ or $S_{t-1,H_t,SS-}$. The demarcation line

that describes $S_{t-1,H_t,SS+}$ and the line for $S_{t-1,H_t,SS-}$ can therefore also be expressed as a linear combination of two edge points (see Appendix Figure 5).

Let a be in the intersection of the line of demarcation ($C\beta_t \leq 0$) with the line segment $\begin{bmatrix} k \\ 0 \\ 1-k \end{bmatrix}$, point b the intersection with $\begin{bmatrix} k \\ 1-k \\ 0 \end{bmatrix}$, and the point c the intersection with $\begin{bmatrix} 0 \\ 1-k \\ k \end{bmatrix}$ where $k \in (0, 1)$. Let a'_+ be the point that maps to a after disease dynamics $f_t\beta_t$ and a positive SS observation, similarly for a'_- except after a negative SS observation, and similarly for b and (b'_+ and b'_-) and c (c'_+ and c'_-). Additionally, to simplify terminology, let us define the "DR direction" as the direction to travel from $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ to $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and the "DS direction" the direction to travel from $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ to $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. See Appendix Figure 5.

In order show that one region is a subset of the other, we first find the points a'_+, a'_-, b'_+, b'_-, c'_+, c'_- relative to points a, b, and c. We will show that $c'_+ = c'_-$, so we can restrict our analysis to points a and b. We will then show that the third element (the probability the patient has DR TB) is smaller for a'_+ than a'_-, and similarly for b'_+, and b'_-,. This means the line for $S_{t-1,H_t,SS-}$ is "below" the one for $S_{t-1,H_t,SS+}$, as depicted in Appendix Figure 5, and therefore one belief set must be a subset of the other.

Step 1:

Let the point a be given by $\begin{bmatrix} h \\ 0 \\ 1-h \end{bmatrix}$, point b by $\begin{bmatrix} \tilde{h} \\ 1-\tilde{h} \\ 0 \end{bmatrix}$, and point c by $\begin{bmatrix} \bar{h} \\ 1-\bar{h} \\ 0 \end{bmatrix}$. As before, we have:

$$\left(\beta_t | z_{t-1} = +, \beta_{t-1}\right) = \overline{M}_+\left(f_{t-1}\beta_{t-1}\right) \Rightarrow \beta_{t-1} = f_{t-1}^{-1}\overline{M}_+^{-1}\left(\beta_t | z_{t-1}\right)$$

$$\text{and } \left(\beta_t | z_{t-1} = -, \beta_{t-1}\right) = \overline{M}_-\left(f_{t-1}\beta_{t-1}\right) \Rightarrow \beta_{t-1} = f_{t-1}^{-1}\overline{M}_-^{-1}\left(\beta_t | z_{t-1}\right)$$

f_{t-1}^{-1} is defined as above, and $\overline{M}_+^{-1}$ and $\overline{M}_-^{-1}$ can be found to be:

$$\overline{M}_{+}^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \frac{1}{(sens)x_1 + (1-spec)(x_2 + x_3)} \begin{bmatrix} (sens)x_1 \\ (1-spec)x_2 \\ (1-spec)x_3 \end{bmatrix}$$

$$\overline{M}_{-}^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \frac{1}{(1-sens)x_1 + (spec)(x_2 + x_3)} \begin{bmatrix} (1-sens)x_1 \\ (spec)x_2 \\ (spec)x_3 \end{bmatrix}$$

We can then find a'_+, a'_-, b'_+, and b'_- to be:

$$a'_+ = f_{t-1}^{-1}\overline{M}_{+}^{-1}\left(\begin{bmatrix} h \\ 0 \\ 1-h \end{bmatrix}\right) = \begin{bmatrix} 1-(1-h)\frac{1}{1-c_{DR}}\frac{1-spec}{h(sens)+(1-h)(1-spec)} \\ 0 \\ (1-h)\frac{1}{1-c_{DR}}\frac{1-spec}{h(sens)+(1-h)(1-spec)} \end{bmatrix}$$

$$a'_- = f_{t-1}^{-1}\overline{M}_{-}^{-1}\left(\begin{bmatrix} h \\ 0 \\ 1-h \end{bmatrix}\right) = \begin{bmatrix} 1-(1-h)\frac{1}{1-c_{DR}}\frac{spec}{h(1-sens)+(1-h)spec} \\ 0 \\ (1-h)\frac{1}{1-c_{DR}}\frac{spec}{h(1-sens)+(1-h)spec} \end{bmatrix}$$

$$b'_+ = f_{t-1}^{-1}\overline{M}_{+}^{-1}\left(\begin{bmatrix} \widetilde{h} \\ 1-\widetilde{h} \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1-\left(1-\widetilde{h}\right)\frac{1}{1-c_{DR}}\frac{1-spec}{\widetilde{h}(sens)+\left(1-\widetilde{h}\right)(1-spec)} \\ \left(1-\widetilde{h}\right)\frac{1}{1-c_{DR}}\frac{1-spec}{\widetilde{h}(sens)+\left(1-\widetilde{h}\right)(1-spec)} \\ 0 \end{bmatrix}$$

$$b'_- = f_{t-1}^{-1}\overline{M}_{-}^{-1}\left(\begin{bmatrix} \widetilde{h} \\ 1-\widetilde{h} \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1-\left(1-\widetilde{h}\right)\frac{1}{1-c_{DS}}\frac{spec}{\widetilde{h}(1-sens)+\left(1-\widetilde{h}\right)spec} \\ \left(1-\widetilde{h}\right)\frac{1}{1-c_{DS}}\frac{spec}{\widetilde{h}(1-sens)+\left(1-\widetilde{h}\right)spec} \\ 0 \end{bmatrix}$$

$$c'_{+} = f_{t-1}^{-1}\overline{M}_{+}^{-1}\left(\begin{bmatrix} 0 \\ 1-\bar{h} \\ \bar{h} \end{bmatrix}\right) = \begin{bmatrix} 1 - \frac{1-\bar{h}}{1-c_{DS}} - \frac{\bar{h}}{1-c_{DR}} \\ \frac{1-\bar{h}}{1-c_{DS}} \\ \frac{\bar{h}}{1-c_{DR}} \end{bmatrix} = f_{t-1}^{-1}\overline{M}_{-}^{-1}\left(\begin{bmatrix} 0 \\ 1-\bar{h} \\ \bar{h} \end{bmatrix}\right) = c'_{-}$$

$c'_{+} = c'_{-}$, so whether one belief set is a subset of the other is not determined by this intersection (see panel b of Appendix Figure 5). We therefore turn to analyzing a, a'_{+}, a'_{-}, b, b'_{+}, and b'_{-}.

Step 2:

We first examine a, a'_{+}, and a'_{-}. As expected, a'_{+} and a'_{-} lie on the same leg of the belief space as a – the probability the patient has DS TB, the second element of the belief vector, is 0 for all three vectors. Since each vector sums to one, we can use the third element of each vector to determine their relative ordering on the line between $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Note that:

$-\frac{1}{1-c_{DR}} \geq 1$ since $0 \leq cDR \leq 1$.

$-\ spec \geq 1 - sens$ since $0.5 < sens < 1$ and $0.5 < spec < 1$

$\Rightarrow \frac{spec}{h(1-sens)+(1-h)spec} \geq 1$ since $0 \leq h \leq 1$

$\Rightarrow (1-h) \leq (1-h)\frac{1}{(1-c_{DR})}\frac{spec}{(h(1-sens)+(1-h)spec)}$

$\Rightarrow$ a'_{-} is in the DR direction from a

$-\ sens \geq 1 - spec$

$\Rightarrow \frac{1-spec}{h(sens)+(1-h)(1-spec)} \leq 1$

$\Rightarrow (1-h)\frac{1}{1-c_{DR}}\frac{1-spec}{h(sens)+(1-h)(1-spec)} \leq (1-h)\frac{1}{1-c_{DR}}\frac{spec}{(h(1-sens)+(1-h)spec)}$

$\Rightarrow$ a'_{-} is in the DR direction from a'_{+}

– Additionally, if $\frac{1-spec}{h(sens)+(1-h)(1-spec)} \leq 1 - c_{DR}$:

$\Rightarrow (1-h)\frac{1}{1-c_{DR}}\frac{1-spec}{h(sens)+(1-h)(1-spec)} \leq (1-h)$

$\Rightarrow$ a is in the DR direction from a'_{+}

So we have shown that to travel from a'_+ and a'_- we need to move in the DR direction, and a will lie between those points if $\frac{1-spec}{h(sens)+(1-h)(1-spec)} \leq 1 - c_{DR}$.

Note that the form of the third element of a'_+ and the second element of b'_+ are the same, except c_{DS} is used instead of c_{DR}, and similarly for a'_- and b'_-. Our assumptions about c_{DR} also hold for c_{DS} – namely that $0 \leq c_{DS} \leq 1$ – so our conclusions about the relative position of a, a'_+, and a'_- also apply to b, b'_+, and b'_-, namely that to travel from b'_+ to b'_-, we need to move in the DS direction, and b will lie between those points if $\frac{1-spec}{h(sens)+(1-h)(1-spec)} \leq 1 - c_{DS}$. See Appendix Figure 5.

Since the half-space H_t must include the point $\begin{bmatrix}0\\0\\1\end{bmatrix}$ (because if the patient has DR TB the optimal action is to administer DST, by assumption), i.e., H_t is "below" the line $\bar{ab}$ in Appendix Figure 5, then $S_{t-1,H_t,SS-}$ is a subset of $S_{t-1,H_t,SS+}$. Similarly, the half-spaces that map to the complementary half-space $\sim H_t$ have the opposite relationship: $S_{t-1,\sim H_t,SS+}$ is a subset of $S_{t-1,\sim H_t,SS-}$. Then since one set is the subset of the other, the intersection between subsets will just be the smaller subset, as we desired to show.

Proof of Theorem 3—Proof by induction.

Initialization step: Let H_t be the half-space of beliefs at time t where the optimal action is to administer DST and $\sim H_t$ be the half-space of beliefs where the optimal action is to continue first-line treatment. $U_{t-1,t}$ is defined as $\left(S_{t-1,H_t,SS+} \cap S_{t-1,H_t,SS-}\right) \cup \left(S_{t-1,\sim H_t,SS+} \cap S_{t-1,\sim H_t,SS-}\right)$. Proposition 4 tells us that $\left(S_{t-1,H_t,SS+} \cap S_{t-1,H_t,SS-}\right)$ is a half-space and that $\left(S_{t-1,\sim H_t,SS+} \cap S_{t-1,\sim H_t,SS-}\right)$ is also a half-space, so $U_{t-1,t}$ is the union of two half-spaces.

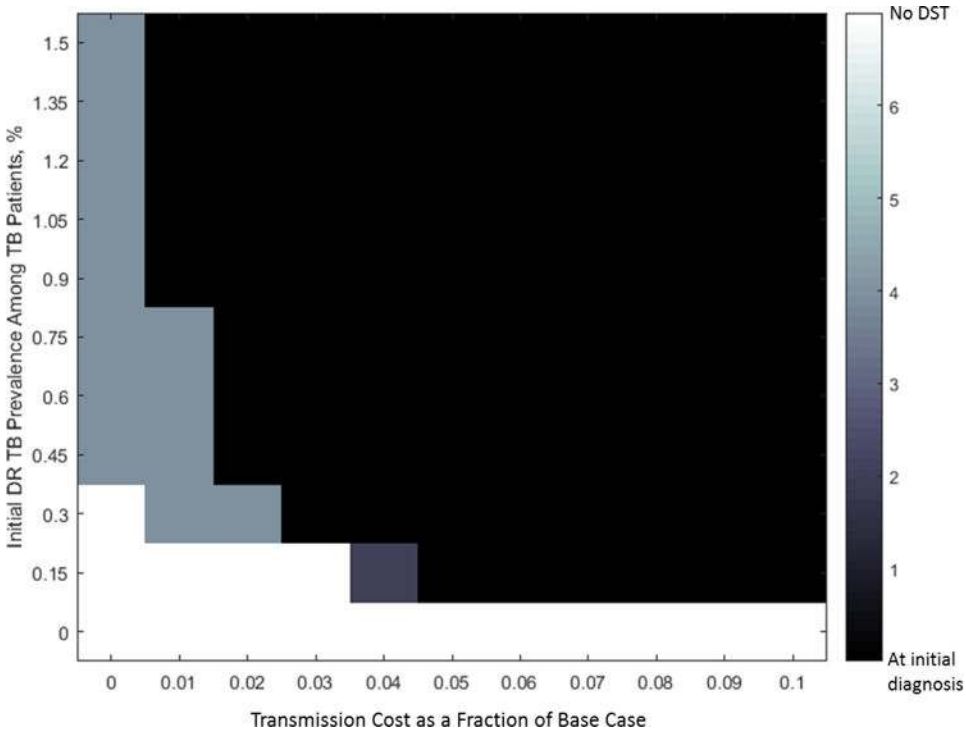
Induction step: Let the two half-spaces of $U_{k,t}$ be called L_k and M_k. Proposition 4 tells us that $\left(S_{t-1,L_t,SS+} \cap S_{t-1,L_t,SS-}\right)$ is a half-space and similarly, $\left(S_{k-1,M_k,SS+} \cap S_{k-1,H_k,SS-}\right)$ is also a half-space.

$U_{k-1},t = \left(S_{t-1,H_t,SS+} \cap S_{t-1,H_t,SS-}\right) \cup \left(S_{k-1,M_k,SS+} \cap S_{k-1,H_t,SS-}\right)$ since these are the areas that map to L_k and M_k regardless of SS test results, and in turn L_k and M_k in map to H_t and $\sim H_t$ regardless of any SS results between periods k and t. Therefore $U_{k-1,t}$ is the union of two half-spaces.

Proof of Theorem 4—Suppose the set of beliefs where DST is the optimal action is a half-space for all times $t = \{0, 1, 2, \ldots, T\}$. Beliefs at time k that lie in $U_{k,t}$ for some $t > k$ will be useless for determining the optimal decision at time t. For an observation to be useless in determining any future decisions, it must lie within $U_{k,t}$ for all $t > k$. Therefore the set of beliefs that are useless in determining any future decisions is the intersection of $U_{k,t}$ for all $t > k$. Theorem 3 tells us that $U_{k,t}$ for each t is the union of two convex regions. Since each $U_{k,t}$ can generate at most double the number of convex regions when intersected with $U_{k,\sim k}$, the maximum number of convex regions in the set of beliefs at time k for which observations do not inform future actions is 2^{T-k}.

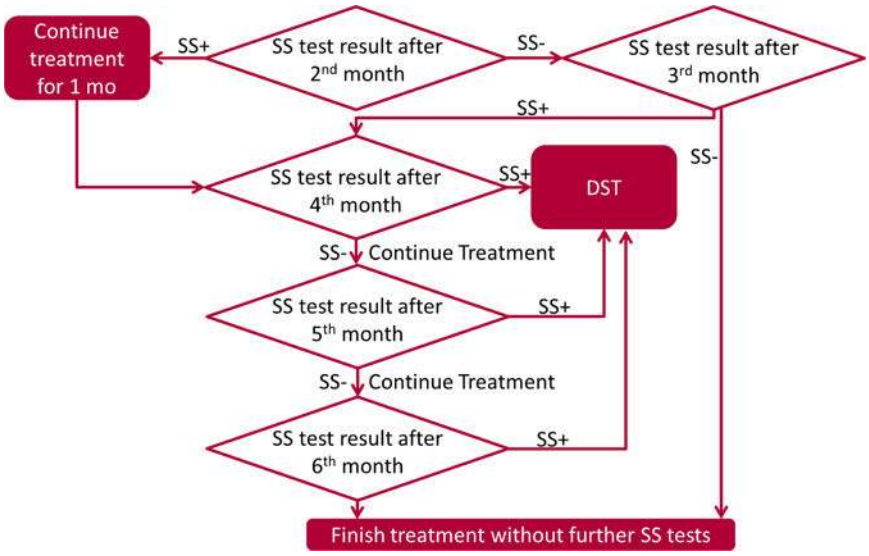
Age	Males	Females	Optimal policy in regions without transmission:
	0-9	0-9	Same as base case
	10-19	10-19	Base case but omit SS test after month 4
	20-29	20-29	Base case but omit SS tests after months 2 and 4
	30-39	30-39	No SS testing, no DST
	40-49	40-49	
	50-59	50-59	
	60-69	60-69	
	70-79	70-79	
	80-89	80-89	
	90+	90+	

Appendix Figure 1.
Optimal testing policy for regions of India without TB transmission, by age and sex



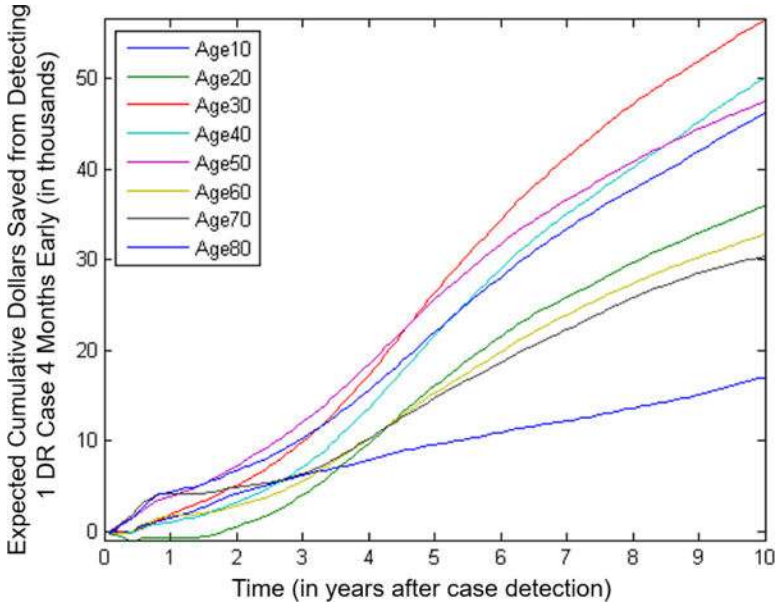
Appendix Figure 2.

Earliest DST month (at end of month) across all possible patient SS test histories for transmission costs between 0 and 10% of the base case ($0 – $1393) and initial DR TB prevalence among first-line TB patients between 0 and 50% of the base case (0 – 0.015)



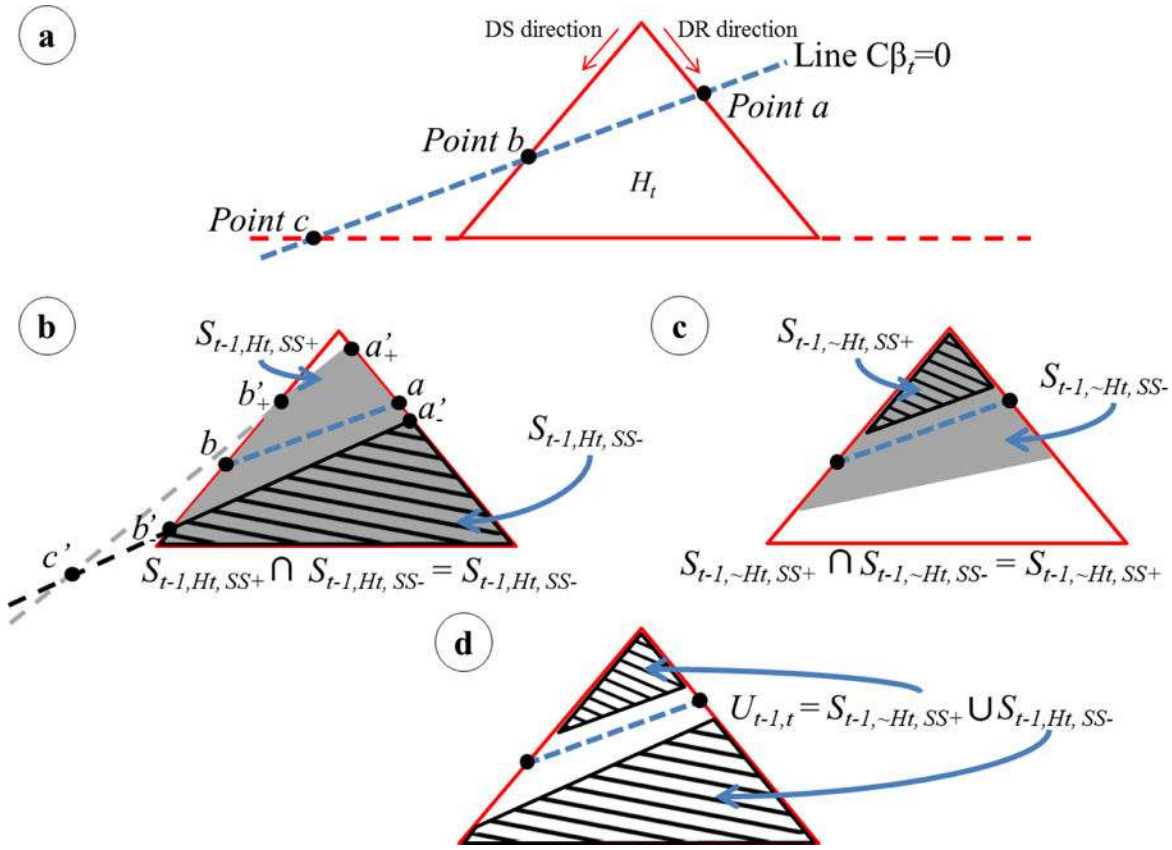
Appendix Figure 3.

Optimal testing policy for regions of India without TB transmission, with imperfect Xpert sensitivity and specificity (97% sensitivity, 95% specificity)



Appendix Figure 4.

Expected costs saved over time from early detection of a DR TB case, by age group



Appendix Figure 5.

Schematic of belief regions described in proofs of Propositions 3 and 4 and Theorem 3. Panel a: The line demarking half-space H_t, $C\beta_t = 0$, intersects the lines that form the edges of the belief space at points *a, b* and *c*. Panel b: The lines that define half-spaces $S_{t-1,H_t,SS+}$ and $S_{t-1,H_t,SS-}$ can be expressed as linear combinations of a'_+ and b'_+, and similarly for a'_- and b'_- (panel c). Panel d: We define the union of the two intersections (given in panel b and c) as $U_{t-1,t}$.

Appendix Table 1

Summary of notation

Value	Description
t	Time epochs, $t = 0, 1, 2, 3, 4, 5$, $T = 6$
s_t	Patient health state at time t. States are: s_1 = Healthy on treatment (H), s_2 = DS TB on treatment (DS), s_3 = DR TB on treatment (DR), s_4 = healthy defaulted (dH), s_5 = DS TB defaulted (dDS), s_6 = DR TB defaulted (dDR), s_7 = dead. States 4, 5, 6, and 7 are directly observable. Can be in the set $S = 1, 2, 3, 4, 5, 6, 7$
a_t	Action at time t. Can be one of *DST* (administer DST to the patient), *wait and see* (continue on first-line treatment with an SS test), *wait and not see* (continue on first-line treatment without an SS test).
δ	Discount factor
f_t	7×7 transition matrix describing treatment dynamics. Element $f_{t,ij}$ provides probability of being in state j at time $t + 1$ if in state i at time t. The terms d, c, and ν represent probabilities of default, cure, and death, respectively.

Value	Description
z_t	Observations at time t. Can be one of {$SS+$, $SS-$, *default*, *death*} if a_t = *wait*, and {*No default or death*, *default*, *death*} if a_t = *DST*.
Beliefs and Belief Updates	
Sens	SS test sensitivity
Spec	SS test specificity
β_t	Belief state at time t. A vector of probability values over each of the health states at time t.
$\overline{M}_{z_t}(f_t \beta_t)$	Function that updates the belief given observation z_t, dynamics f_t, and prior belief β_t: $\beta_t = \overline{M}_{z_t}(f_t \beta_t)$
$N_+(x)$	$\frac{1}{(1-spec)x(1)+(sens)(x(2)+x(3))}$. Used in the belief update: $M_{zt}^{-} = {}_+(x) = N_+(x)M_+x$
M_+	$\begin{bmatrix}(1-spec) & 0 & 0 & 0\,0\,0\,0\\ 0 & (sens) & 0 & 0\,0\,0\,0\\ 0 & 0 & (sens) & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\end{bmatrix}$. Used in the belief update: $M_{zt}^{-} = {}_+(x) = N_+(x)M_+x$
$N_-(x)$	$\frac{1}{(spec)x(1)+(1-sens)(x(2)+x(3))}$. Used in the belief update: $M_{zt}^{-} = {}_-(x) = N_-(x)M_-x$
M_-	$\begin{bmatrix}(spec) & 0 & 0 & 0\,0\,0\,0\\ 0 & (1-sens) & 0 & 0\,0\,0\,0\\ 0 & 0 & (1-sens) & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\\ 0 & 0 & 0 & 0\,0\,0\,0\end{bmatrix}$. Used in the belief update: $M_{zt}^{-} = {}_-(x) = N_-(x)M_-x$
$P(x)$	$\frac{1}{x(4)+(x(5)+x(6))}$. Used in the belief update: $M_{zt}^{-} = default(x) = P(x)Qx$
Q	$\begin{bmatrix}0\,0\,0\,0\,0\,0\,0\\ 0\,0\,0\,0\,0\,0\,0\\ 0\,0\,0\,0\,0\,0\,0\\ 0\,0\,0\,1\,0\,0\,0\\ 0\,0\,0\,0\,1\,0\,0\\ 0\,0\,0\,0\,0\,1\,0\\ 0\,0\,0\,0\,0\,0\,0\end{bmatrix}$. Used in the belief update: $M_{zt}^{-} = default(x) = P(x)Qx$
Rewards	
$r_{wait,a}$	Stage reward.
r_{DST}	Terminal rewards if a_t = *DST*. A 1 × 7 vector of net monetary benefits (NMB): $[NMB_{DST,H}\ NMB_{DST,DS}\ NMB_{DST,DR}\ NMB_{dH}\ NMB_{dDS}\ NMB_{dDS}\ NMB_{dead}]$.
$r_{\sim DST}$	Terminal reward if patient never undergoes DST. A 1 × 7 vector of net monetary benefits (NMB): $[NMB_{\sim DST,H}\ NMB_{\sim DST,DS}\ NMB_{\sim DST,DR}\ NMB_{dH}\ NMB_{dDS}\ NMB_{dDS}\ NMB_{dead}]$

Appendix Table 2

Transition parameters for India example[1]

Description	Value			Reference
Sputum smear test				
Sensitivity	0.60			[30]
Specificity	0.98			[30]
Cure and default in first-line treatment				
Cumulative probability of cure	After Month:	DS	DR	[37]
	1	0.35	0.00	
	2	0.88	0.00	
	3	0.97	0.00	
	4	1.00	0.30	
	5	1.00	0.30	
	6	1.00	0.30	
Monthly probability of default	Age	Males	Females	[3,26,32]
	0	0.0229	0.0229	
	20	0.0222	0.0192	
	30	0.0219	0.0173	
	40	0.0216	0.0155	
	50	0.0212	0.0136	
	60	0.0209	0.0118	
	70	0.0205	0.0099	
Mortality				
Monthly mortality in treatment, DS or DR TB	0.0101			[3,26,32]
Monthly mortality, DS or DR TB without treatment (selected ages shown)	Age	Males	Females	[35]
	0	0.0289	0.0290	
	10	0.0248	0.0248	
	20	0.0249	0.0249	
	30	0.0250	0.0249	
	40	0.0252	0.0250	
	50	0.0257	0.0252	
	60	0.0270	0.0264	
	70	0.0303	0.0292	
	80	0.0360	0.0345	
	90	0.0473	0.0462	
	99	1.0000	1.0000	
Monthly mortality, no TB and no treatment (selected ages shown)	Age	Males	Females	[48]
	0	0.0043	0.0044	
	10	0.0001	0.0001	
	20	0.0002	0.0002	
	30	0.0003	0.0002	

Description	Value		Reference
	40	0.0005	0.0003
	50	0.0010	0.0006
	60	0.0024	0.0018
	70	0.0057	0.0047
	80	0.0116	0.0100
	90	0.0231	0.0220
	99	1.0000	1.0000

[1] Transition probabilities for calculation of the post-first-line treatment dynamics, used in the simulation that generates lifetime costs and QALYs, are given in Suen et. al [14]

Appendix Table 3

QALY values and costs[2] for India example

Description	Value			Reference
Test costs ($)				
Sputum smear test	4.93			[7]
GeneXpert test	26.57			[7]
Monthly first-line treatment costs ($)				
Patient costs	6.33			[40]
Drug costs	3.60			[2,43,38]
Clinic costs	31.71			[41]
Discount factor				
Annual discount rate	3%			[46]
Societal losses due to an additional month of transmissibility				
Drug-sensitive TB transmission	Age	Costs ($)	QALYS	Calculated from simulation
	10	2,469	−9	[14]
	20	11,346	−28	
	30	11,460	−32	
	40	10,499	−41	
	50	11,539	−32	
	60	7,381	−30	
	70	10,424	−28	
	80	11,024	−16	
	90	11,024	−16	
Drug-resistant TB transmission	Age	Costs ($)	QALYS	Calculated from simulation
	10	9,400	−9	[14]
	20	14,957	−28	
	30	31,027	−31	
	40	31,867	−31	
	50	23,021	−30	
	60	23,604	−21	
	70	18,774	−17	
	80	20,036	−25	

Description	Value			Reference
	90	20,036	−24	
Quality-of-life multipliers				
Healthy, treatment naive	1.00			[47]
Healthy, with past treatment	0.94			[26]
Healthy, on first-line treatment	0.85			Assumed
DS TB, not on treatment	0.66			[42]
DS TB, on first-line treatment	0.84			[34]
DR TB, not on DR TB treatment	0.66			[42]
DR TB, on DR TB treatment	0.75			[3]
Willingness-to-pay				
GDP per capita in India ($)	1450			[44]
Monthly background healthcare costs when not on TB treatment				
	Age	Males	Females	[39]
	0	0.20	0.19	
	10	0.17	0.26	
	20	1.23	2.63	
	30	1.62	2.78	
	40	1.66	2.92	
	50	2.53	2.37	
	60	3.31	2.44	
	70	4.76	4.18	
Lifetime discounted costs and QALYs				
At the end of first-line treatment for a 30-year-old male by health and treatment state				
		Costs ($)	QALYs	Calculated from simulation
Healthy (cured)		421.79	16.530	[14]
DS TB, exiting all treatment		171.71	4.353	
DR TB, exiting all treatment		115.56	4.084	
DR TB, entering DR TB treatment		1560.00	7.079	
For a 30-year-old male DR patient successfully triaged to DR TB treatment				
	After Month	Costs ($)	QALYs	Calculated from simulation
	0	1559.10	7.117	[14]
	1	1559.30	7.111	
	2	1559.40	7.105	
	3	1559.50	7.098	
	4	1559.70	7.092	
	5	1559.80	7.086	
	6	1560.00	7.079	

[2] Costs in this table are reported in 2012 USD, following [14], as this is the simulation used to estimate lifetime costs. Final reported results are scaled to 2015 USD using data from [45] and [36].

[3] Average of no treatment and first-line treatment for patients with DS TB (effective treatment but toxic).

Appendix Table 4

Years until initial costs for Xpert implementation are recovered

	Years until initial costs are recovered, as a function of per patient initial cost		
Age of detected DR TB case	**$100**	**$500**	**$1,000**
10	1.8	9.8	37.5
20	2.9	5.2	9.2
30	1.6	4.0	5.9
40	2.1	4.4	6.8
50	0.8	3.8	6.3
60	2.3	5.3	10.3
70	0.7	5.5	12.0
80	0.8	4.3	7.0

References

1. World Health Organization. Global Tuberculosis Report 2014. 2014. URL: http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809βeng.pdf
2. India National Ministry of Health. Revised National Tuberculosis Control Programme (RNTCP) training module for medical practitioners. 2010. URL: http://tbcindia.nic.in/pdfs/Training%20Module%20for%20Medical%20Practitioners.pdf
3. India National Ministry of Health. RNTCP performance report, India: third quarter, 2010. 2010. URL: http://www.tbcindia.nic.in/WriteReadData/l892s/1778097518Perf%203Q%202010.pdf
4. World Health Organization. Country Profile: India. 2015. URL: http://www.who.int/countries/ind/en/
5. India Ministry of Health and Family Welfare. Revised National Tuberculosis Control Programme (RNTCP) national strategic plan for tuberculosis elimination 20172025. 2017. URL: http://tbcindia.gov.in/WriteReadData/NSP%20Draft%2020.02.2017%201.pdf
6. Menzies NA, Cohen T, Lin H-H, Murray M, Salomon JA. Population health impact and cost-effectiveness of tuberculosis diagnosis with Xpert MTB/RIF: a dynamic simulation and economic evaluation. PLoS Med. 2012; 9(11):e1001347. [PubMed: 23185139]
7. Vassall A, van Kampen S, Sohn H, Michael JS, John KR, den Boon S, Davis JL, Whitelaw A, Nicol MP, Gler MT, Khaliqov A, Zamudio C, Perkins MD, Boehme CC, Cobelens F. Rapid diagnosis of tuberculosis with the Xpert MTB/RIF assay in high burden countries: a cost-effectiveness analysis. PLoS Med. 2011; 8(11):e1001120. [PubMed: 22087078]
8. Choi HW, Miele K, Dowdy D, Shah M. Cost-effectiveness of Xpert MTB/RIF for diagnosing pulmonary tuberculosis in the United States. Int J Tuberc Lung Dis. 2013; 17(10):1328–1335. [PubMed: 24025386]
9. You JH, Lui G, Kam KM, Lee NL. Cost-effectiveness analysis of the Xpert MTB/RIF assay for rapid diagnosis of suspected tuberculosis in an intermediate burden area. J Infect. 2015; 70(4):409–414. [PubMed: 25573001]
10. Castillo-Chavez C, Song B. Dynamical models of tuberculosis and their applications. Math Biol Eng. 2004; 1(2):361–404.
11. Porco TC, Blower SM. Quantifying the intrinsic transmission dynamics of tuberculosis. Theor Popul Biol. 1998; 54:117–132. [PubMed: 9733654]
12. Liu L, Zhaob XQ, Zhou Y. A tuberculosis model with seasonality. Bull Math Biol (2010). 2010; 72:931–952.
13. Suen S, Bendavid E, Goldhaber-Fiebert JD. Disease control implications of Indias changing multi-drug resistant tuberculosis epidemic. PLoS One. 2014; 9(3):e89822. [PubMed: 24608234]

14. Suen SC, Bendavid E, Goldhaber-Fiebert JD. Cost-effectiveness of improvements in diagnosis and treatment accessibility for tuberculosis control in India. Int J Tuberc Lung Dis. 2015; 19(9):1115–1124. [PubMed: 26260835]
15. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton University Press; Princeton, NJ: 2008.
16. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008; 5(3):e74. [PubMed: 18366252]
17. Horby P, Pham QT, Hens N, Nguyen TT, Le QM, Dang DT, Nguyen ML, Nguyen TH, Alexander N, Edmunds WJ, Tran ND, Fox A, Nguyen TH. Social contact patterns in Vietnam and implications for the control of infectious diseases. PLoS One. 2011; 6(2):e16965. [PubMed: 21347264]
18. Ayer T, Alagoz O, Stout N. A POMDP approach to personalize mammography screening decisions. Oper Res. 2012; 60(5):1019–1034.
19. Maillart LM, Ivy JS, Ransom S, Diehl K. Assessing dynamic breast cancer screening policies. Oper Res. 2006; 56(6):1411–1427.
20. Goulionis JE, Vozikis A. Medical decision making for patients with Parkinson disease under Average Cost Criterion. Aust New Zealand Health Policy. 2009; 6:15. [PubMed: 19549341]
21. Hoey J, Poupart P, von Bertoldi A, Craig T, Boutilier C, Milhailidis A. Automated handwashing assistance for persons with dementia using video and a partially observable Markov decision process. Comput Vis Image Und. 2010; 114(5):503–519.
22. Smallwood RD, Sondik EJ. The optimal control of partially observable Markov processes over a finite horizon. Oper Res. 1973; 21(5):1071–1088.
23. World Health Organization. Cost-effectiveness and strategic planning (WHO-CHOICE): cost-effectiveness thresholds. 2014. URL: http//www.who.int/choice/costs/CER-thresholds/en/
24. India Ministry of Health and Family Welfare. 2015 Revised National TB Control Programme Annual Status Report. 2015. URL: http://www.tbcindia.nic.in/showfile.php?lid=3166
25. World Health Organization. Frequently asked questions on Xpert MTB/RIF assay. 2009. URL: http://www.who.int/tb/laboratory/xpertβfaqs.pdf
26. Cohesive India. Collaborations for Health System Improvement and Impact Evaluation in India. 2010. URL: http://www.cohesiveindia.org/research-projects.html#best
27. World Health Organization. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. 2013. Policy update. URL: http://apps.who.int/iris/bitstream/10665/112472/1/9789241506335βeng.pdf
28. World Health Organization. Use of high burden country lists for TB by WHO in the post-2015 era. 2015. URL: http://www.who.int/tb/publications/globalβreport/highβtbβburdencountrylists2016–2020.pdf
29. World Health Organization. HIV drug resistance surveillance guidance: 2015 update. 2016. URL: http://apps.who.int/iris/bitstream/10665/204471/1/9789241510097βeng.pdf?ua=1
30. Srikanth P, Kamesh S, Daley P. Bleach optimization of sputum smear microscopy for pulmonary tuberculosis. Indian J Tuberc. 2009; 56(4):174–184. [PubMed: 20469728]
31. Suen SC, Goldhaber-Fiedbert JG. An efficient, non-iterative method of identifying the cost-effectiveness frontier. Med Decis Making. 2016; 36(1):132–136. [PubMed: 25926282]
32. Babiarz KS, Suen S, Goldhaber-Fiebert JD. Tuberculosis treatment discontinuation and symptom persistence: an observational study of Bihar, Indias public care system covering 100,000,000 inhabitants. BMC Public Health. 2014; 14:1–13. [PubMed: 24383435]
33. Siegel JE, Sanders GD, Russell LB, Ganiats TG, Neumann PJ. Cost-effectiveness in health and medicine. 2ed. Oxford University Press; New York: 2016.
34. Dhingra VK, Rajpal S. Health related quality of life (HRQL) scoring in tuberculosis. Indian J Tuberc. 2003; 37(4):261–8.

35. Dye C, Scheele S, Dolin P, Pathania V, Raviglione M. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. JAMA. 1999; 282(7):677–686. [PubMed: 10517722]
36. Board of Governors of the Federal Reserve System. Historical exchange rates for the Indian rupee. 2017. URL: http://www.federalreserve.gov/releases/h10/hist/dat00βin.htm
37. Fitzwater SP, Caviedes L, Gilman RH, Coronel J, LaChira D, Salazar C, Saravia JC, Reddy K, Friedland JS, Moore DA. Prolonged infectiousness of tuberculosis patients in a directly observed therapy short-course program with standardized therapy. Clin Infect Dis. 2010; 51(4):371–378. [PubMed: 20624064]
38. Global Drug Facility, Stop TB Partnership. Buying quality, affordable tuberculosis drugs through the Global Drug Facility: a guide to the Direct Procurement Service for donors, nongovernmental organizations and programme managers. World Health Organization; Geneva, Switzerland: 2013.
39. Indian Ministry of Statistics and Programme Implementation. NSS survey reports. 2014. URL: http://mospi.nic.in/Mospi_New/site/inner.aspx?status=3&menu_id=31
40. Pantoja A, Floyd K, Unnikrishnan KP, Jitendra R, Padma MR, Lal SS, Uplekar M, Chauhan LS, Kumar P, Sahu S, Wares F, Lnnroth K. Economic evaluation of public-private mix for tuberculosis care and control, India. Part I. Socio-economic profile and costs among tuberculosis patients. Int J Tuberc Lung Dis. 2009; 13(6):698–704. [PubMed: 19460244]
41. Salje H, Andrews JR, Deo S, Satyanarayana S, Sun AY, Pai M, Dowdy DW. 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. 2014; 11(7):e1001674. [PubMed: 25025235]
42. Salomon JA, Vos T, Hogan DR, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet. 2010; 380(9859):2129–43.
43. Stop TB Partnership. Global drug facility product catalogue. World Health Organization; Geneva, Switzerland: 2013.
44. The World Bank. GDP per capita 2014. 2014. URL: http://data.worldbank.org/indicator/NY.GDP.PCAP.CD
45. The World Bank. Inflation, consumer prices (annual %). 2016. URL: http://data.worldbank.org/indicator/FP.CPI.TOTL.ZG/countries/IN?displaygraph
46. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB, for the Panel on Cost-Effectiveness in Health and Medicine. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. JAMA. 1996; 276(15):1253–8. [PubMed: 8849754]
47. Winetsky DE, Negoescu DM, DeMarchis EH, Almukhamedova O, Dooronbekova A, Pulatov D, Vezhnina N, Owens DK, Goldhaber-Fiebert JD. Screening and rapid molecular diagnosis of tuberculosis in prisons in Russia and Eastern Europe: a cost-effectiveness analysis. PLoS Med. 2012; 9(11):e1001348. [PubMed: 23209384]
48. World Health Organization. WHO Lifetables for India. 2010. URL: http://apps.who.int/ghodata/?vid=720

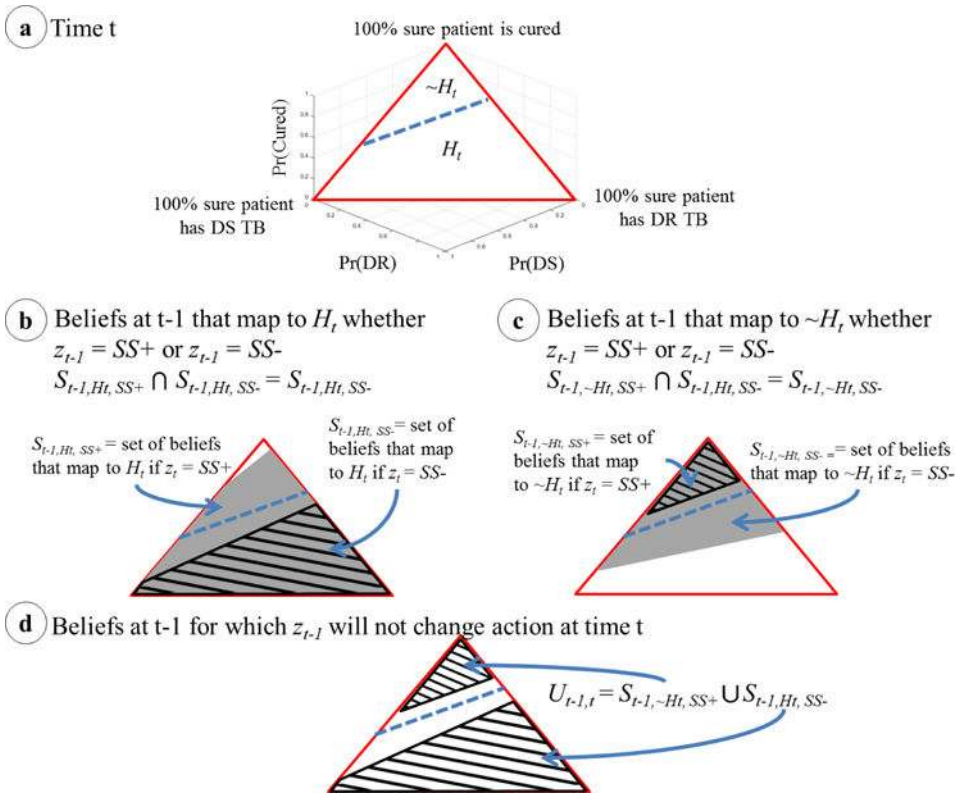


Figure 1.

Belief regions. Panel a visualizes possible values of states 1, 2, and 3 in β_t (area enclosed by red lines); these are represented as points in $\mathbf{R}^3$. We depict only the first three states for visual clarity and because we are concerned with the belief regions where the optimal action is to DST or not (no actions are possible in states 4–7). The corners of the feasible space represent beliefs that the patient is cured, has DS TB, or has DR TB with certainty. Suppose it is optimal to administer DST for all beliefs in region H_t and remain on treatment in region ~H_t (panel a). Panels b and c show that the regions that map to H_t and ~H_t at time $t-1$ regardless of what SS test result is observed are the intersection of subsets. Panel d illustrates $U_{t-1,t}$, the union of the two intersections given in panels b and c.

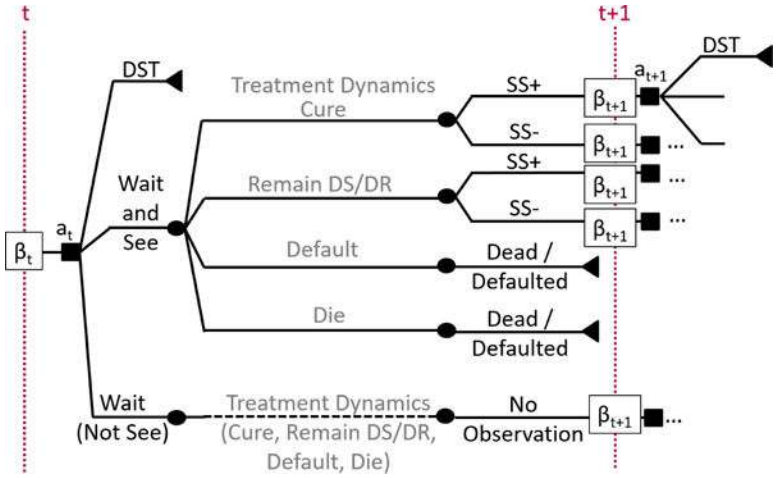


Figure 2.

Model schematic: empirical example of TB in India. This figure depicts the belief update process used in the model at every time period.

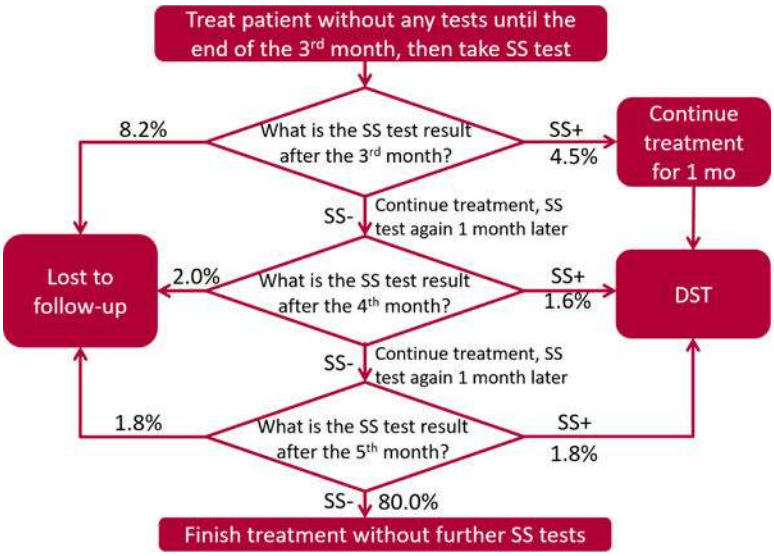


Figure 3.

Results: Optimal diagnosis algorithm in a region with no TB transmission (with expected percentage of patients shown on each path). The POMDP solution provides the optimal action for every possible combination of no observation, SS+, and SS− results. We represent these results in flowchart form.

Table 1

Selected model parameters for India (see appendix for full table and citations)

Description	Value
SS test sensitivity	0.60
SS test specificity	0.98
Cumulative probability of cure for DS TB patient after month:	
1	0.35
2	0.88
3	0.97
4	1.00
5	1.00
6	1.00
Cumulative probability of cure for DR TB patient after month:	
1	0.00
2	0.00
3	0.00
4	0.30
5	0.30
6	0.30
SS test cost	$4.93
DST test cost	$26.57
Monthly patient costs	$6.33
Monthly drug costs	$3.60
Monthly clinic costs	$31.71
Annual discount rate	3%

Table 2

Net monetary benefits compared to no treatment ($). Status quo is India's current policy.

	λ = GDP/capita		λ= 0	
	No Trans.	**With Trans.**	**No Trans.**	**With Trans.**
Status quo	6365	−1157	−1943	−5015
Optimal policy	6368	6335	−1949	−1971
Optimal policy, no DST at diagnosis	6368	4406	−1949	−2752

Table 3

Additional national number of tests, NMB, and costs (compared to the status quo) for the optimal testing policy over one year

	Values for optimal policy	
	No Trans.	**With Trans.**
# Additional SS Tests	449,000	−1,174,000
# Additional DST Tests	22,000	598,000
Additional NMB	$1.4M	$4710M
Additional Costs	$3.8M	−$1890M