Full Study Protocol and Statistical Plan

Study Title: A Clinical Risk Score for Early Management of TB in Uganda
(NIH/NIAID R21AI161301, PI: Dr. David Dowdy)

NCT Number: NCT05122624

Date of the document: August 03, 2021

1.SIGNIFICANCE

Tuberculosis (TB) remains the leading cause of mortality, causing over 1.4 million deaths in 2018.[1] A major contributor to TB mortality is underdiagnosis and undertreatment[3]– particularly in resource-limited primary health clinics. In these settings, up to 40% of patients diagnosed with TB are either lost to care before starting treatment or have treatment substantially delayed.[3] Data from our study site in Kampala, Uganda, confirm this finding:[13] 45% of patients who test positive for TB by Xpert MTB/RIF ("Xpert", Cepheid, Inc., Sunnyvale, CA, USA) remain untreated one week later. Patients who are diagnosed with Xpert-positive TB and yet remain untreated are at great risk of both transmission and mortality.[14] Many of these infections[15] and deaths[16] could be averted if TB treatment were initiated on the same day as initial presentation.

The most straightforward way to provide same-day TB treatment initiation is to offer same-day microbiological diagnosis, as recommended by the World Health Organization (WHO).[17] Point-of-care Xpert testing shortens time to treatment[18,19] and in rural Malawi,[20] reduced all-cause mortality among patients newly diagnosed with advanced HIV. Similarly, same-day testing and treatment initiation for HIV has been demonstrated to improve patient outcomes, both short-term (e.g., >25% increase in 3-month linkage-to-care)[21] and longer-term (e.g., 18% increase in 12-month viral suppression)[4], relative to testing with delayed treatment initiation. However, implementation of same-day Xpert testing is not feasible in many highly resource-limited settings, which often lack the necessary infrastructure and financial resources.[14,22–24] In these settings, confirmatory TB testing may take days to weeks, and the risk of pre-treatment loss-to-follow-up is high.[3,25] If global targets (e.g., 90% TB mortality reduction by 2035)[26] are to be met, there is therefore an urgent need to improve TB management in these most-resource-constrained clinics where same-day microbiological testing is infeasible.

When same-day diagnosis cannot be made, clinicians must balance the risks and benefits of same-day empiric treatment initiation. For many diseases – including gonorrhea[27], chlamydia[28], and urinary tract infections[29] – same-day empiric treatment (after an initial risk assessment) is commonly prescribed, as the harms of treatment delay outweigh the harms of overtreatment. Same-day empiric treatment for TB, however, is uncommon – as the duration of treatment is at least six months, and the perceived harms of overtreatment are high.[6] As a result of these perceptions, one common practice is to prescribe broad-spectrum antibiotics while awaiting the results of sputum TB testing – a practice that lacks a strong evidence base.[30] An alternative approach would be to perform a rapid TB risk assessment and prescribe same-day empiric treatment for TB to those at high risk. The harm of a missed TB diagnosis has been estimated at more than 30 times that of a false-positive diagnosis.[5] Thus, same-day empiric treatment of TB could provide clinical benefit by preventing progression of symptoms (and perhaps mortality) and epidemiological benefit by preventing transmission (as effective treatment rapidly renders most patients non-infectious[31,32]). Despite its potential to greatly improve patient and population outcomes, same-day empiric treatment of TB in settings where same-day confirmatory diagnosis is unavailable would represent a major shift in the clinical management of TB – a shift that would require a large cluster-randomized trial for confirmation. Prior to conducting a major clinical trial, it is essential to test the hypothesis that same-day empiric treatment improves clinical outcomes for patients with TB in the most resource-constrained settings.

A necessary precondition to implementing same-day empiric treatment for TB is the availability of a risk assessment that is sufficiently easy-to-use that it could be implemented in peripheral clinics in high-TB-burden settings.[15,33] Unfortunately, most existing models for predicting active TB require equipment or infrastructure (e.g., radiology[34], laboratory testing[34–37], stadiometer measurements[36,37], and calculations requiring a computer or smartphone[38]) that are generally unavailable in clinical settings that lack same-day microbiological testing for TB. Other prediction models incorporate data only available in specific contexts (for example, household contact investigations[39,40] or HIV clinics[34,38]) and thus cannot be applied more broadly. Our team recently developed and validated a clinical risk score for pulmonary TB that easily calculable using pen and paper, employing only data that would be readily accessible to mid-level clinicians after a brief patient interview. This risk score has sufficiently high accuracy to improve clinical decision-making (see Preliminary Data below) but has not yet been demonstrated to impact patient outcomes. The proposed study will generate these data, which are necessary before deciding whether to proceed with a larger cluster randomized clinical trial.

2.INNOVATION

This proposed study is innovative in terms of its intervention, setting, design, and clinical relevance:

- **Innovative tool:** We propose to evaluate a clinical risk score for pulmonary TB that was developed in rural South Africa and subsequently validated, in terms of accuracy and potential clinical utility, in urban

Uganda. This score has not yet been assessed in terms of implementation or effectiveness. This study will therefore be one of the first assessments of the clinical and epidemiological impact of a simple prediction rule designed to improve patient-important outcomes for TB in high-burden settings.

- **Innovative setting:** This clinical risk score has been explicitly designed for implementation in peripheral health clinics with minimal infrastructure. Very few clinical innovations have been created with the most-resource-constrained settings in mind.
- **Innovative study design:** We propose a highly pragmatic, hybrid implementation-effectiveness evaluation that will both provide an evidence-based justification for, and inform the implementation of, a subsequent cluster randomized trial of the impact of a clinical risk score on patient-important outcomes.
- **Potential to change clinical management:** This study is supported by the Ugandan National TB and Leprosy Program (see Letter of Support) and if successful (and confirmed in a subsequent trial) could effect a major change in the clinical management of pulmonary TB among the patients at greatest risk for catastrophic consequences, including long-term sequelae, mortality, and community transmission.

3.APPROACH

3.1. Investigative Team

The principal investigators for this proposed work have a strong track record of successful collaboration, spanning over a decade. **Dr. David Dowdy**, PI, is an infectious disease epidemiologist who leads several active R01-funded projects on TB, including a trial of TB preventive therapy (R01HL144406) and an investigation of community-level TB transmission (R01HL138728) in Uganda. He has interdisciplinary expertise in TB epidemiology (including clinical trial management), economic evaluation, and mathematical modeling. **Dr. Achilles Katamba,** site PI, is a renowned Ugandan epidemiologist and clinician with expertise in interventions for TB control. He has extensive experience in TB case detection, diagnostic evaluation, and implementation research. He also has well-established relationships with laboratory managers (e.g., Moses Joloba), the Ugandan National TB and Leprosy Program, and local leaders of multiple affected communities. Drs. Dowdy and Katamba collaborate on multiple studies and have published extensively together,[9,13,15,41–57] thereby ensuring the likely success of this project (and any subsequent trial).

3.2. Preliminary Studies and Experience

3.2.1. The PredicTB Risk Score

Our team has developed a simple clinical risk score ("PredicTB") for predicting active pulmonary TB in settings where same-day microbiological testing is not available (Figure 1). This score was developed using data from 1,387 adult patients who presented with classical TB symptoms (cough, fever, weight loss, night sweats) at 28 primary health clinics in rural Limpopo district, South Africa, and for whom a sputum Xpert test for TB was ordered.[12] We then externally validated this score using data from 387 patients who presented with TB symptoms and were tested using Xpert in four primary health clinics in Kampala, Uganda. This clinical risk score is designed for ease of use, being scored on a scale from 1-10 and using only readily accessible demographic and clinical data (sex, age, self-reported HIV status, self-reported diabetes status, number of symptoms, and duration of symptoms >14 vs ≤14 days; see Figure 1). The c-statistic for this easy-to-use risk score in the validation population was 0.75, consistent with other scores using more detailed clinical data,[34–38] and decision curve analysis indicated that the net utility of using this score (at a cutoff of ≥5/10) would be significantly higher than no empiric treatment. As an example, a person with a PredicTB score of 7/10 in a population where 10% of all diagnostic Xpert tests are positive (i.e., pre-test probability of 0.1) would have a calculated 43% chance of having TB – a post-test probability that likely is sufficiently high to merit same-day empiric treatment. *De novo* development of a corresponding score in the Ugandan population resulted in very similar attributes being selected (diabetes was dropped and antiretroviral status plus observed cough were added), and very similar performance (validation c-statistic of 0.78).[12] This easy-to-use risk score is therefore sufficiently accurate to improve clinical decision-making in the most-resource-constrained settings; the proposed study will evaluate whether real-world implementation of this score can improve patient outcomes as well.

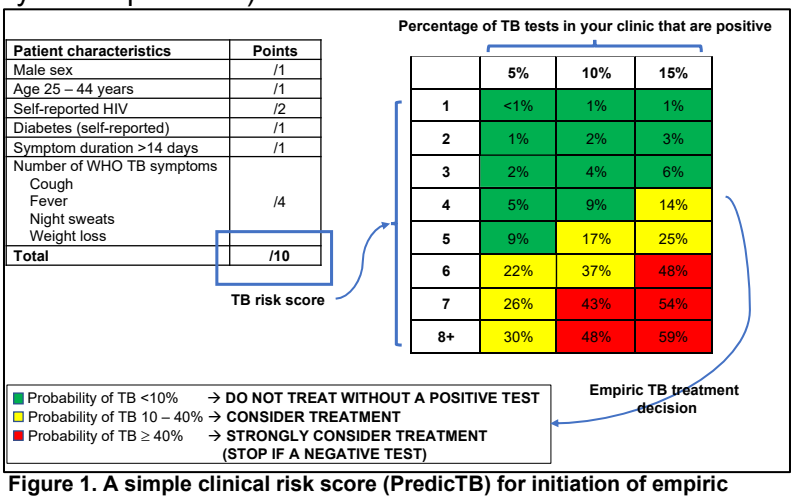


Figure 1. A simple clinical risk score (PredicTB) for initiation of empiric TB treatment when same-day testing is not available

3.2.2. Impact of same-day diagnostic testing for pulmonary TB on patient outcomes in Uganda

Our team performed a comparison of patient outcomes comparing eight clinics with available Xpert testing on-site against 10 clinics using Xpert as an offsite referral test[54] (R21AI106031, PI: Dowdy). We found that the proportion of patients starting TB treatment was not significantly higher in clinics with on-site Xpert testing (11% vs 9%), and that a positive Xpert test changed management in fewer than 1% of all patients tested. We subsequently collaborated as part of a clinical trial (XPEL-TB, R01HL130192, PI: Cattamanchi) in which peripheral health centers were randomized to on-site Xpert testing versus off-site referral testing[44] for a period of 18 months. The design of this trial was highly pragmatic,[58] with minimal involvement of study staff in routine clinic activities, and the primary outcome assessed by way of ongoing record review (as also proposed for the present study). Preliminary results from this trial indicate that the number of individuals diagnosed and treated for microbiologically confirmed pulmonary TB within 14 days was 14.0 per clinic in the offsite testing arm versus 27.7 per clinic in the on-site testing arm (relative risk 2.0, 95% confidence interval 1.5-2.7) – a doubling of TB diagnoses made. Thus, in the off-site testing arm, more than half of all individuals with microbiologically confirmable TB were either never diagnosed or started on treatment with a delay of >14 days. Unfortunately, despite being a leader in Xpert procurement – with an estimated 526 Xpert modules procured by the end of 2016[46] – Uganda does not have capacity to offer on-site testing to the vast majority of its >3,000 peripheral level clinics.[59] These results indicate the urgent need for substantial improvement in TB management in the most-resource-constrained peripheral health clinics of Uganda and other high-TB-burden countries. They also demonstrate our team's ability to effectively measure patient-important TB outcomes using highly pragmatic study designs in this context.

3.2.3. Population-level impact of same-day TB treatment in Africa

In addition to the conduct of pragmatic clinical studies, our team also has expertise in modeling of long-term clinical and epidemiological impact of same-day TB treatment. For example, we used a compartmental mathematical model to estimate the impact of rapid TB treatment initiation in the WHO African Region. We projected that same-day treatment initiation could reduce annual TB mortality in sub-Saharan Africa by 44% within ten years.[15] These projections highlight the substantial epidemiological gains that could be achieved by effective implementation of same-day TB treatment initiation in Uganda and other high-burden settings.

3.3 Research Design and Methods

3.3.1. Study setting and design

We will conduct this study in four primary health clinics in peri-urban Uganda (with four additional clinics serving as a comparison). Uganda is one of 30 high TB burden countries recognized by World Health Organization (WHO), with an estimated 2018 TB incidence of 200 per 100,000; 41% of those cases were co-infected with HIV.[1] Like many countries in East Africa, Uganda is among the world's poorest countries with an estimated per-capita gross domestic product (GDP) of $777 2019;[60] the drivers of its TB epidemic are multifaceted,[61] and there is relatively little drug resistance. Most TB in Uganda is diagnosed in public sector health facilities, and despite national guidelines promoting the use of Xpert, sputum smear microscopy remains the primary diagnostic assay used.[54] Diagnostic and treatment delays are common,[62] resulting in a large number of untreated prevalent cases; the 2014-15 Ugandan National TB Prevalence Survey[63] estimated the prevalence of bacteriologically confirmed TB at 401 per 100,000 adults. These conditions are typical of sub-Saharan Africa (outside of South Africa); thus, findings from our study are likely to be broadly relevant.

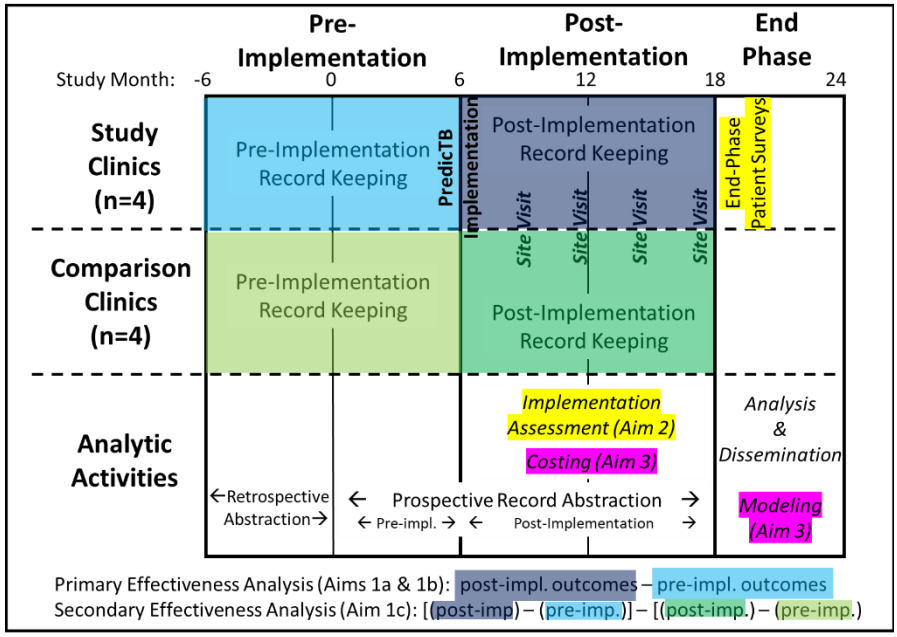
We have selected four study clinics based on the following criteria: (a) within 150km of the capital city, Kampala; (b) no access to same-day Xpert testing (a criterion that effectively excludes clinics in Kampala proper); (c) high-quality TB record-keeping; (d) at least 100 bacteriologically confirmed TB cases diagnosed in 2019 (which effectively excludes private-sector clinics); and (e) clinic in-charge willing to have clinic staff trained in use of the PredicTB tool. We have also selected four comparison clinics that are similar in location and patient population to the four primary study clinics, but for which the latter two criteria were relaxed.

3.3.2. Study overview

The overall goals of this proposal are to estimate the effectiveness of the PredicTB clinical risk score in improving clinical outcomes for patients with pulmonary TB and to generate pragmatic data regarding the implementation of this score in highly resource-constrained primary health clinics. If successful, this study will provide an evidence base to justify the design and conduct of a larger cluster-randomized trial. This study will be conducted as a type 2 hybrid effectiveness-implementation evaluation. Our primary effectiveness outcome will be the difference in the proportion of adults with bacteriologically confirmed pulmonary TB who initiate treatment within 7 days of submitting sputum for TB diagnosis, comparing the 12 months before implementing

the PredicTB score to the 12 months after implementation. Secondary analyses will assess for the potential that measured differences reflect improved general practice and/or secular trends. Our primary implementation outcome will be the proportion of patients in the study clinics in whom the PredicTB score is used as indicated. Our study will consist of four phases (Figure 2): pre-implementation, implementation, post-implementation, and end phase. We will adopt a highly pragmatic study design in which we minimize involvement of study staff in the clinics under evaluation, with the exception of the intervention itself.

Figure 2. Study Overview



3.3.3. Study procedures

In the **pre-implementation phase** (month -6 to month 6), study and comparison clinics (selected as already having high-quality TB records) will continue to maintain standard TB registers. These registers include all individuals who have sputum sent for TB diagnosis with sputum smear and/or Xpert ("presumptive register") and all individuals who are started on TB treatment ("treatment register"). We will abstract key data from these registers, including age, sex, prior TB treatment, HIV status, diagnostic result, and treatment outcome. During month 1 of the study, we will initiate a retrospective abstraction of all registers from month -6. We will also begin prospective abstraction of registers by having routine clinic staff take photos of registers and securely transfer them to our study team, who will upload these into a REDCap database. Staff from the XPEL-TB trial[44] have experience in these procedures (including the development of a pre-existing database for this purpose), so we will be able to initiate this abstraction rapidly and seamlessly.

In the **implementation phase** (month 7), we will train clinic staff in all eight clinics (study clinics and comparison clinics) on the Ugandan standard of care for diagnosis and treatment of TB. In addition, in the four study clinics, we will provide training on the PredicTB score. This training will include the following elements: (a) background on the PredicTB score and its aim of informing empiric treatment decisions for adult pulmonary TB; (b) provision of materials (paper score sheets and laminated instruction sheets) to enable clinicians to use the PredicTB score; (c) reinforcement that all treatment decisions are ultimately at the discretion of the treating physician; (d) reinforcement that PredicTB is not meant to be a replacement for microbiological testing – that sputum should be sent for microbiological testing in all patients presenting with symptoms of TB, and that any patient with a positive microbiological test should be treated for TB. Prior to this training, we will speak with the head (in-charge) physician at each of the four study clinics to describe the PredicTB score and elicit the physician's recommendations as to the percentage of TB tests that are positive in the clinic (i.e., columns in Figure 1) and the score that should be encouraged as a threshold for same-day empiric treatment initiation. Importantly, this threshold may vary from one clinic to another; our analyses will not explicitly incorporate this threshold as a covariate, but we will explore clinic-by-clinic variation by way of informing any future clinical trial. Furthermore, we will seek the in-charge physician's advice regarding an alternative approach in which patients with high (e.g., >40%) risk could be scheduled for a full treatment course, whereas those with intermediate (e.g., 10-40%) risk could be prescribed a provisional (e.g.,14-day) course, which could be discontinued if microbiological testing for TB was negative and the clinical course [e.g., full resolution in 1-2 days or continued worsening of symptoms despite TB therapy] were inconsistent with a diagnosis of TB. Training in the use of the PredicTB risk score in each clinic will therefore follow the recommendations of each clinic's in-charge physician and will only augment (not supplant) the Ugandan standard of care for TB management. We will re-evaluate these recommendations every three months during site visits (see below) and update training as needed.

In the **post-implementation phase** (months 7-18), we will encourage all eight clinics (study and comparison clinics) to continue high-quality record-keeping as during the pre-implementation phase. The four study clinics

will be asked to maintain records of the PredicTB score sheets for all patients on whom the score was calculated. We will also perform quarterly site visits that will include the following components:

- Review of the quality of record-keeping and transmittal of records to study staff [all clinics];
- Collection of cost data, including budgetary review and time-and-motion studies [study clinics];
- Assessment of PredicTB reach (proportion of patients receiving same-day empiric treatment initiation), adoption (proportion of clinicians adopting PredicTB), and implementation (proportion of patients in whom PredicTB was used as indicated) [study clinics]; and
- Performance feedback to clinic staff regarding TB diagnosis and treatment outcomes [all clinics].

During this phase, we will also perform mycobacterial culture (using both solid [Lowenstein-Jensen] and liquid [Mycobacteria Growth Indicator Tube, MGIT; BD Diagnostics, Sparks, MD, USA] media) on all sputum specimens from all eight clinics that are Xpert-negative but correspond to adult patients who were empirically initiated for TB. *These culture results will not be used for any of our primary effectiveness analyses* but instead will be used to estimate the sensitivity and specificity of empiric treatment decisions, both with and without the PredicTB score. Specimens will be sent from the central laboratories in which Xpert is performed, rather than from study clinics themselves, so as to minimize interference with study clinic procedures. Nevertheless, culture results will be transmitted back to clinicians to assist with management.

In the **end phase** (months 19-24), we will complete the statistical analysis of effectiveness (Aim 1) and implementation (Aim 2). We will also perform cost-effectiveness and impact analyses for Aim 3. In addition, in the four study clinics, we will conduct a series of 100 patient surveys (25 per clinic) in which we recruit, consent, and enroll consecutive patients in each clinic who submit sputum for TB testing. For these patients, study staff will independently administer the PredicTB score and will also assess the following: (a) treatment decisions and patient motivation to complete treatment; (b) additional TB risk factors (e.g., smoking status, observed cough, antiretroviral therapy, symptom duration/severity, body mass index); (c) patient socioeconomic status, transit time to clinic, and cost of attending clinic; and (d) patient opinions regarding the value of same-day empiric TB treatment. These additional data will enable us to consider refinements to the PredicTB model (both content and implementation) in preparation for a potential future large-scale trial.

3.3.4. Analysis Plan

Table 1. Overview of Analysis Plan

Specific Aim	**Analytic Goal**	**Primary Outcomes** (1° outcome underlined)	**Secondary Analyses**	**Analytic Strategy**
Aim 1	Evaluate PredicTB Effectiveness	Differences in proportions (pre-post): Treatment initiation (1a) Mortality (1b) Loss to follow-up (1b) Difference in differences (study – comparison) (1c)	Adjustment for age,sex, HIV status, treatment history Regression discontinuity	Mixed-effects log-binomial models (clinic as random effect) Parametric regression discontinuity analysis Difference-in-difference (“product term”) model
Aim 2	Evaluate PredicTB Implementation	Reach (2a): % same-day empiric treatment Adoption (2b): % providers using PredicTB Implementation (2c): % of encounters using PredicTB as indicated Maintenance (2d): chg in effectiveness over time	Validation with patient surveys Sensitivity/specificity of empiric tx decisions Association with patient level characteristics	Binomial distribution & confidence intervals w/ clustering by clinic Mixed-effects log-binomial models
Aim 3	Estimate PredicTB Impact & Cost-Effectiveness	Projected reduction in 5-year TB mortality (3a) Incremental cost per DALY averted (3b)	Probabilistic sensitivity analyses Uncertainty analyses Unit cost estimation	Markov state-transition model Cost-effectiveness analysis

Aim 1. Evaluate the effectiveness of the PredicTB clinical risk score on clinical outcomes. Our primary analysis will estimate the absolute difference in the proportion of patients with microbiologically confirmed TB (Xpert or smear; study-driven culture results will be excluded) who are initiated on anti-TB treatment within 7 days of submitting an initial sputum specimen for TB diagnosis, comparing the post-initiation period to the pre-initiation period (dark minus light blue in Figure 2). We will estimate this quantity using a multilevel log-binomial

model, with random-effect terms corresponding to each of the four study clinics. We will perform similar analyses for the proportion of patients with TB who complete treatment and die from any cause (Aim 1b).

This analysis is subject to bias resulting from both secular trends and non-intervention-related effects. We will therefore also perform a difference-in-differences analysis, comparing pre-post effects in the study clinics to those observed in the comparison (non-PredicTB) clinics (Aim 1c). We have not made this our primary analysis because the comparison clinics are smaller (thereby reducing power) and not randomly allocated – for example, staff in these clinics may be less motivated to change behavior. We will also perform a parametric regression discontinuity analysis to estimate the causal effect of implementing the PredicTB risk score (with the discontinuity occurring at the time of implementation). This analysis, however, has less statistical power, and the shape of the effectiveness curve over time in the post-implementation period may reflect important effects (i.e., we are interested not only in the immediate effect of the intervention but also its maintenance over time).

Sample size considerations. Based on patient volumes in 2019, we anticipate a sample size of 1100 patients with confirmed TB: 550 pre-intervention and 550 post-intervention. Using data from our STOMP-TB study,[13] we expect that 55% of patients will initiate treatment within 7 days in the pre-intervention period. Assuming a design effect of 1.67 due to clustering by clinic, we will have 80% power to detect a difference if treatment initiation increases to 66% in the post-intervention phase and 97% power if treatment initiation increases to 70% (Figure 3).

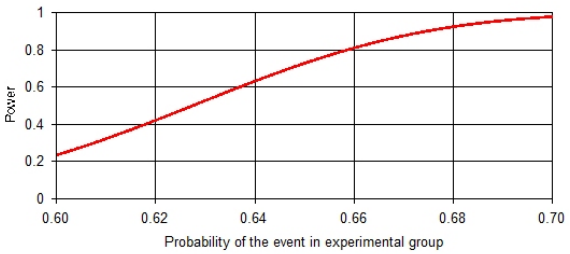


Figure 3. Study power

Aim 2. Evaluate the implementation of the PredicTB clinical risk score. Using the RE-AIM framework[64,65] as a guide, we will evaluate the implementation of PredicTB according to its reach (proportion of symptomatic patients initiating same-day empiric treatment, Aim 2a), effectiveness (Aim 1), adoption (proportion of providers managing TB in study clinics who use the PredicTB score, Aim 2b), implementation (proportion of eligible patient encounters in which PredicTB is used as indicated), and maintenance (change in effectiveness between month 7 and 18). We will estimate each of these quantities using data from treatment registers and completed PredicTB score sheets and will validate these estimates on the subset of patients who present to treatment during our quarterly site visits. All analyses will include random effect terms to account for clustering of implementation effects at the clinic level.

Aim 3. Project the long-term impact and cost-effectiveness of PredicTB implementation. Based on our experience of estimating the clinical impact of TB treatment regimens[66,67] and diagnostic tests[68,69], we will construct a Markov model of long-term patient outcomes, informed by data from this study. This model will use estimates from the scientific literature to incorporate probabilities of TB mortality, repeat diagnostic attempts, and effectiveness of partial TB treatment. Our primary impact outcome will be the projected reduction in five-year TB mortality effected by PredicTB implementation. We will expand this model to incorporate estimates of costs (which we will empirically measure during the study, using study tools developed by our team for this purpose[69,70]). We will follow standard guidance for economic evaluation, including inflation using Uganda's GDP deflator, conversion to a common currency and year (2021 US dollars), discounting of all future costs and effectiveness, and measurement of costs as economic (opportunity) costs.[71,72] We will perform a probabilistic sensitivity analysis (using beta distributions for parameters with defined upper and lower bounds and gamma distributions for those bounded at zero to infinity) to generate 95% uncertainty ranges. We will construct cost-effectiveness acceptability curves, using $500 per DALY as a conservative cost-effectiveness threshold[73,74] but also considering revealed thresholds for competing interventions.

3.3.5. Potential Challenges, Alternative Solutions, and Next Steps.

Our biggest challenges include potential reductions in patient volume (e.g., COVID-related) and the possibility that PredicTB will not improve TB treatment initiation (e.g., if uptake is low). If patient volumes are reduced, given the low resource requirement for record abstraction, we will space out our site visits and request a no-cost extension to increase our sample size. If PredicTB does not improve outcomes, this will be an important finding – and the present study will enable us to understand why this occurred, while averting the expense of a major clinical trial. If successful, however, this study will provide preliminary data and justification for such a trial – a trial with the potential to effect a major change in clinical practice and improve meaningful outcomes for patients with TB in the most resource-constrained settings. Given the tremendous burden of morbidity, mortality, and transmission risk borne by these patients – coupled with the short timelines of global End TB targets – the need for easy-to-use interventions that improve TB management could not be more urgent.