



Published in final edited form as:

Lancet Respir Med. 2015 March ; 3(3): 201–209. doi:10.1016/S2213-2600(15)00036-3.

Evaluation of sputum culture conversion as a prognostic marker for end-of-treatment outcome inpatients with multidrug-resistant tuberculosis

Ekaterina V. Kurbatova, MD¹, J. Peter Cegielski, MD¹, Christian Lienhardt, PhD², Rattanawadee Akksilp, BSc³, Jaime Bayona, MD⁴, Mercedes C. Becerra, ScD⁴, Janice Caoili, MD⁵, Carmen Contreras, BA⁶, Tracy Dalton, PhD¹, Manfred Danilovits, MD⁷, Olga V. Demikhova, MD⁸, Julia Ershova, PhD¹, Victoria M. Gammino, PhD¹, Irina Gelmanova, MD⁴, Charles M. Heilig, PhD¹, Ruwen Jou, PhD⁹, Boris Kazenny, MD¹⁰, Salmaan Keshavjee, MD⁴, Hee Jin Kim, MD¹¹, Kai Kliiman, PhD⁷, Charlotte Kvasnovsky, MD¹, Vaira Leimane, MD¹², Carole D. Mitnick, ScD⁴, Imelda Quelapio, MD⁵, Vija Riekstina, MD¹², Sarah E. Smith, MPH¹, Thelma Tupasi, MD⁵, Martie van der Walt, PhD¹³, Irina A. Vasilyeva, MD⁸, Laura E. Via, PhD¹⁴, Piret Viikklepp, MD¹⁵, Grigory Volchenkov, MD¹⁶, Allison Taylor Walker, PhD¹, Melanie Wolfgang, MPH¹, Martin Yagui, MD¹⁷, and Matteo Zignol, MD²

¹CDC, Division of TB Elimination, Atlanta, GA, USA ²World Health Organization, Geneva, Switzerland ³Office of Disease Prevention and Control, Region 7. UbonRatchatani, Thailand ⁴Partners In Health (Lima, Peru; Tomsk, Russian Federation; Boston, MA, USA); Harvard Medical School, Boston, MA, USA ⁵Tropical Disease Foundation, Manila, Philippines ⁶Socios en Salud Sucursal, Lima, Peru ⁷Tartu University Hospital, Tartu, Estonia ⁸Central Tuberculosis Research Institute, Russian Academy of Medical Sciences, Moscow, Russia ⁹Reference laboratory of Mycobacteriology, Taiwan Centers for Disease Control, Taiwan ¹⁰Orel Oblast Tuberculosis

© 2015 Elsevier Ltd. All rights reserved.

Author Contributions

Conceived and designed the study: EVK, JPC, CL, JE, TD; carried out the study: EVK, JPC, CL, RA, JB, MCB, JC, CC, TD, MD, OVD, JE, VMG, IG, CMH, RJ, BK, SK, HJK, KK, CK, VL, CDM, IQ, VR, SES, TT, MvdW, IAV, LEV, PV, GV, ATW, MW, MY, MZ; analyzed the data: EVK, JPC, CMH; wrote the first draft of the manuscript: EVK, JPC, CL; contributed to the writing of the manuscript: EVK, JPC, CL, RA, JB, MCB, JC, CC, TD, MD, OVD, JE, VMG, IG, CMH, RJ, BK, SK, HJK, KK, CK, VL, CDM, IQ, VR, SES, TT, MvdW, IAV, LEV, PV, GV, ATW, MW, MY, MZ; agreed with manuscript results and conclusions: EVK, JPC, CL, RA, JB, MCB, JC, CC, TD, MD, OVD, JE, VMG, IG, CMH, RJ, BK, SK, HJK, KK, CK, VL, CDM, IQ, VR, SES, TT, MvdW, IAV, LEV, PV, GV, ATW, MW, MY, MZ; enrolled patients: RA, JB, JC, CC, MD, OVD, IG, RJ, BK, HJK, KK, VL, IQ, TT, MvdW, IAV, GV, MY.

Conflicts of Interest

No conflict of interest reported for all contributors.

Ethical approvals

The Preserving Effective TB Treatment Study (PETTS) was approved by the U.S. Centers for Disease Control and Prevention (CDC) Institutional Review Board (IRB) and IRBs at all participating sites. Informed consent for participation in the study was obtained from all patients. The original DOTS-Plus Pilot Projects Case-Based Study (CBS) study was approved by CDC's IRB. Informed consent was waived by the IRB because the data were obtained through retrospective medical records review. For the present, secondary analysis, the data were stripped of identifiers. Thus, this analysis did not involve identifiable human subjects, and it was designated accordingly by CDC.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Dispensary, Orel, Russia ¹¹Korean Institute of TB, Seoul, Republic of Korea ¹²Riga East University hospital, Tuberculosis and Lung Disease Center, Riga, Latvia ¹³Medical Research Council, Pretoria, South Africa ¹⁴National Institute for Allergy and Infectious Disease, National Institutes of Health, Bethesda, MD, USA ¹⁵National TB Registry, National Institute for Health Development, Tallinn, Estonia ¹⁶Vladimir Oblast Tuberculosis Dispensary, Vladimir, Russia ¹⁷National Institute of Health, Lima, Peru

Abstract

Background—To assess the validity of sputum culture conversion (SCC) on solid media at varying time points and the time to SCC as prognostic markers for end-of-treatment outcome in multidrug-resistant (MDR) tuberculosis (TB) patients.

Methods—Data on 1,712 MDR-TB patients from two large cohort studies were analyzed. Measures of association were determined using random effects multivariable logistic regression. Predictive values were calculated using bivariate random-effects generalized linear mixed model.

Findings—Times to SCC and SCC status at 6 months were significantly associated with treatment success compared to failure or death. SCC status at 2 months was significantly associated with treatment success among patients without known HIV infection only. The overall association of SCC with a successful outcome was substantially stronger at 6 months (adjusted odds ratio [aOR]=14.07, 95% CI 10.05–19.71) than at 2 months (HIV-negative patients: aOR=4.12 [2.25–7.54]; HIV unknown: aOR=3.59 [1.96–6.58], HIV-positive: aOR=0.38 [0.12–1.18]). The 2-month SCC had low sensitivity (27%) and high specificity (90%) for predicting treatment success. Conversely, 6-month SCC status had high sensitivity (92%), but moderate specificity (58%). The maximum combined sensitivity and specificity for SCC was reached between the 6th and 10th month of treatment.

Interpretation—Time to SCC, SCC status at 6 months, and SCC status at 2 months among patients without known HIV infection can be considered proxy markers of end-of-treatment outcome in MDR-TB patients, but the overall association with treatment success is substantially stronger for 6-month compared to 2-month SCC.

Funding—USAID, the US CDC, the Division of Intramural Research of NIAID/NIH, and the Republic of Korea's CDC.

Keywords

multidrug-resistant tuberculosis; prognostic marker; outcome; sputum culture conversion; treatment

Background

Sputum culture conversion (SCC) is commonly used as an early microbiological endpoint in Phase II clinical trials of the treatment of tuberculosis (TB) based on its assumed predictive value for end-of-treatment outcome, particularly in drug-susceptible TB. In December 2012, the U.S. Food and Drug Administration (FDA) approved bedaquiline, the first anti-tuberculosis drug developed in more than 40 years, for the treatment of multidrug-resistant

(MDR) TB, based on an accelerated procedure using time to SCC and SCC status at 6 months as surrogate microbiological markers of end-of-treatment outcomes.¹ In November 2013, the European Medicines Agency (EMA) recommended conditional marketing authorization for another new therapeutic agent, delamanid, for treatment of MDR-TB, based on a dossier using the surrogate microbiological marker of SCC at 2 months. While there is ample literature on 2-month SCC as a proxy marker for treatment outcome in drug susceptible TB^{2,3}, the validity of 2-month and 6-month SCC status as proxy markers for treatment outcome in MDR-TB patients has not been clearly demonstrated.⁴ Yet these were recognized by the FDA and EMA as suitable indicators, and marketing authorization followed.⁵

The objective of our study was to evaluate the validity of time-to-culture conversion, and 2-month and 6-month SCC as prognostic markers for end-of-treatment outcome in MDR-TB, and to identify optimum time points for SCC as a marker for final treatment outcome using data from two large cohort studies of MDR-TB. The results of this evaluation may guide the choice of relevant surrogate end-points in future Phase II/III clinical trials on MDR-TB.

Methods

Patient population and study design

The Preserving Effective TB Treatment Study (PETTS) and the DOTS-Plus Pilot Projects Case-Based Study (CBS) designs and patient populations were reported previously.⁶⁻⁸ Briefly, PETTS was a prospective cohort study of consecutively enrolled adults with pulmonary MDR-TB who started treatment with second-line drugs (SLDs), 01/01/2005–12/31/2008 in nine countries (Estonia, Latvia, Peru, the Philippines, Russia, South Africa, South Korea, Taiwan, and Thailand). CBS was a retrospective cohort study that included adult patients with MDR-TB who started treatment between 01/01/2000–12/31/2003 in four of the first five DOTS-plus projects approved by the Green Light Committee (Latvia, Peru, the Philippines, and Russia). In both studies, patients were followed until the end of treatment or for at least two years from MDR-TB treatment start, monthly sputum microscopy and cultures on solid medium were conducted as part of routine care.

Definitions

A positive culture was defined as 1 colony of *Mycobacterium tuberculosis*. SCC was defined as 2 consecutive negative cultures from sputum samples collected at least 30 days apart.⁹ Two negative culture results in sequence counted towards this definition even if there were missing culture(s) between them. Time to initial SCC was defined as the time in months from the date of starting MDR-TB treatment to the specimen collection date for the first of these two consecutive negative cultures.⁹ Sustained SCC was defined as the absence of any subsequent positive cultures after SCC. Sputum culture reversion to positive was defined as at least one subsequent positive culture result after initial SCC. Persistent culture positivity was defined as no culture conversion in patients with a baseline positive culture.

End-of-treatment outcomes were assigned by the clinical sites according to WHO definitions at the time, namely, cure: at least five consecutive negative cultures in the final 12 months of

treatment (one scant positive is allowed as long as the three last cultures are negative); treatment completed: successful completion of treatment with fewer than five cultures performed in the final 12 months of treatment; death from any cause; failure: two or more (of the five) cultures positive in the final 12 months or any one of the final three cultures positive; loss to follow-up: interruption for two or more months.⁹ Treatment success was defined as cure or completion of treatment. Poor outcome was defined as failure or death.⁹

Data analysis

Statistical analyses were performed using SAS software, version 9.3 (SAS Institute Inc., Cary, NC). A *P* value < 0.05 was considered statistically significant.

Time to initial SCC was analyzed by the Kaplan-Meier method and differences between groups were assessed with the log-rank test. For patients who never converted time to initial SCC was censored one month before their last sputum specimen date.

To estimate the association of 2-month and 6-month SCC with successful treatment outcome, odds ratios (OR) and 95% confidence intervals (CIs) were calculated using random effects multivariable logistic regression (SAS ProcGlimmix). SCC was included in the model as the main predictor. Two separate regression models were fitted for 2-month and 6-month SCC, respectively. Interactions of SCC with each of the covariates of interest were assessed. Significant interactions (if detected) were retained, and final multivariable models were adjusted for covariates that had plausible associations with treatment outcome based on the published literature, regardless of the statistical significance.

We evaluated the sensitivity and specificity of initial culture conversion at 2 months and 6 months in predicting treatment outcome. Sensitivity was defined as the proportion of patients with SCC by month 2 and month 6 among those with successful treatment outcome. Specificity was defined as the proportion of patients without SCC by month 2 and month 6 among those with poor treatment outcome. The positive predictive value (PPV) for each time-point of SCC was defined as the proportion of patients in whom treatment was successful among all those with initial culture conversion. Because PPV depends on the prevalence of the condition of interest, we assumed 60% prevalence of treatment success based on the average proportion reported in cohorts of MDR-TB patients worldwide,¹⁰⁻¹² and sensitivity analyses were carried out for alternative prevalence rates of 50%, 70% and 80% (Online supplement Table S2). Receiver operating characteristic (ROC) curves were plotted to visualize the effect of using different time-points for SCC on the balance between sensitivity and specificity.

To jointly model the sensitivities and specificities while accommodating heterogeneity between countries and studies, we used a bivariate random-effects generalized linear mixed model (SAS ProcGlimmix).^{13,14} The unit of analysis was the country in the specific study which yielded 13 strata. To assess the heterogeneity between studies and countries on the one hand, and other factors associated with the prognostic performance of SCC on the other hand, we analyzed study-level and individual-level predictors. We calculated ORs that reflected the magnitude and significance of association between each factor and the probability of correctly predicting treatment success (sensitivity) and poor outcomes

(specificity) by SCC status. Accuracy of the results by study and country were summarized in forest plots.

Role of the funding source

The sponsors had no role in the study design; in the collection, analysis, or interpretation of the data; and in the writing of the report.

EVK, JPC, TD, JE, SES, MW had full access to all of the raw PETTS data from all sites. EVK, JPC, VMG, ATW had full access to all of the raw CBS data from all sites. Co-authors from each site had full access to all of the raw data from their own site. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

Results

Characteristics of the study population and initial SCC

The two datasets included 3,529 patients. Of these, 2,043 (58%) had positive sputum cultures at the start of treatment, documented MDR-TB without additional resistance to both fluoroquinolones and injectable agents and had one of the treatment outcomes specified by WHO (Figure 1). Characteristics of the study population are shown in Online supplement Table S1.

Sputum culture conversion was initially observed in 1,738 (85.1%) of the 2,043 patients, but 340 (19.6%) had at least one subsequent positive culture after converting. Ultimately, of the 2,043 patients, 1,603 (78.5%) had sustained conversion, conversion was not sustained in 135 (6.6%), and 305 (14.9%) never converted. The median time to initial culture conversion was 3 months, with an interquartile range (IQR) of 2-3 months. HIV coinfecting patients had significantly slower time to SCC (6 months, IQR 3- 24) compared to HIV-negative (3 months, IQR 2-5) and HIV status unknown patients (2 months, IQR 1-3) ($P<.001$, Log-rank test). Among those who reverted to positive, the median time after conversion to the first subsequent positive culture was 10 months (IQR 7-14). Overall, 1,344 (65.8%) subjects had treatment success, treatment failed in 122 (6.0%), 246 (12.0%) died, and 331 (16.2%) were lost to follow up.

All further analyses focused on 1,712 of the 2,043 patients, excluding 331 patients who were lost to follow up.

Time to SCC and treatment success

Among patients with treatment success, the median time to SCC was significantly shorter, 2 months (IQR 1-3), compared with patients who had poor outcomes, 7 months (IQR 3- 24) ($P<.001$, Log-rank test) (Figure 2).

Association of SCC at 2 and 6 month with treatment success

Patients who converted sputum culture at 2 months had a higher chance of treatment success than those without SCC at 2 months (unadjusted OR=3.60, 95% CI 2.49–5.19) (Table 1). In multivariable analysis, however, 2-month SCC had a significant interaction with HIV

infection (Table 2): 2-month SCC was significantly associated with treatment success among HIV-negative patients (adjusted OR [aOR]=4.12, 95% CI 2.25-7.54) and patients with HIV status unknown (aOR=3.59, 95% CI 1.96-6.58), but not among HIV-positive patients (aOR=0.38, 95% CI 0.12-1.18) adjusted for covariates.

The odds of treatment success were higher among patients with 6-month SCC, compared to patients without 6-month SCC (unadjusted OR=15.18, 95% CI 11.23-20.53) (Table 1). In multivariable analysis, adjusting for the same covariates, the association of SCC with treatment success was substantially stronger at 6 months (aOR=14.07, 95% CI 10.05-19.71) than at 2-months (Table 2). There was no significant interaction with HIV status.

Diagnostic performance of 2-and 6-month SCC in predicting treatment success

When comparing successful with poor treatment outcomes, SCC by the end of 2 months of treatment had a sensitivity of 27.3% (95% CI 16.6%-41.4%) and specificity of 89.8% (95% CI 82.3%-94.4%) (Table 3). PPV of 2-month SCC was 80.2% (95% CI 79.3-81.0). The sensitivity of 2-month SCC varied from 14.6% to 49.0% across strata of study-level and individual-level predictors (Table 3). The sensitivity was significantly lower in patients with a history of anti-tuberculosis treatment, cavitory disease, or positive AFB smear than in the complementary groups. Specificity of 2-month SCC was 83.3%-94.8% across strata of predictors, except for HIV-positive patients (73.8%), the only category with significantly lower specificity (Table 3). When the data were stratified by country, high variability was observed for the empiric sensitivity of 2-month SCC (8.9%-69.8%), while the range of specificity (60.0%-100%) was narrower (Figure 3).

For SCC by the end of 6 months of treatment, the sensitivity and specificity were 91.8% (95% CI 85.9%-95.4%) and 57.8% (95% CI 42.5%-71.6%), respectively. PPV was 76.5% (95% CI 75.6%-77.4%). Across strata of predictors, sensitivity of SCC at 6 months was 85.7%-96.6% (Table 3). The sensitivity was significantly lower in patients with a history of anti-tuberculosis treatment, positive AFB smears, or baseline fluoroquinolone resistance compared to the complementary groups. Specificity of 6-month SCC varied across strata of the covariates (46.3%-70.8%), and was significantly higher among patients with positive AFB smear at baseline (Table 3). When the data were stratified by country, high variability was observed for specificity of 6-month SCC (10.0%-100%), while the range of sensitivity (65.5%-100%) was narrower (Figure 3).

The effect of using different time points of SCC on the balance between sensitivity and specificity is shown on the ROC curve (Figure 4). A high gain in sensitivity was observed from the 2nd to the 4th month of treatment (from 27.3% to 79.2%) with a relatively smaller decrease in specificity (from 89.9% to 68.6%). The combined sensitivity and specificity in predicting successful treatment outcome reached a maximum at the 6th to 10th months of treatment.

Discussion

Among MDR-TB patients, time-to-culture conversion and SCC at 6 months were significantly associated with end-of-treatment outcomes. SCC at 2 months was significantly

associated with end-of-treatment outcomes among patients without known HIV infection, but not among HIV-positive patients. The overall association of SCC with outcome was substantially stronger at 6 months than 2 months. The 2-month definition of culture conversion had low sensitivity (27%) and high specificity (90%) for predicting treatment success compared to failure or death, while the sensitivity of 6-month SCC definition was high (92%), but specificity was moderate (58%). The maximum combined sensitivity and specificity was reached for SCC status assessed between the 6th and 10th month of treatment. Given the average 60% prevalence of treatment success reported in MDR TB cohorts worldwide¹⁰⁻¹², PPVs were relatively high: 80% for 2-month SCC and 77% for 6-month SCC. These results are important because of growing interest in conducting Phase II trials of potential new treatments for MDR-TB to determine the best regimens to be evaluated in much larger, longer, and more expensive Phase III trials (see panel, “Research in context”).

Surrogate endpoints should satisfy the three criteria of 1) correlation with a definitive clinical endpoint, 2) reproducibility and 3) clinical/biological plausibility.¹⁵ In the context of clinical trials, a low sensitivity of the prognostic marker would bias the results toward the null, as a high proportion of patients with eventual successful outcome would be classified as unsuccessful, thus carrying the risk of dropping early a successful drug or regimen. However, low specificity of the prognostic marker is of serious concern as it would overestimate the efficacy of the drug or regimen. Thus, in the context of a clinical trial, moderate specificity of 58%, as observed for 6-month SCC in this study, would mean that 42% of patients in whom treatment fails or die would have culture conversion by 6 months and would have been falsely classified as having a successful outcome, thus overestimating the efficacy of a new drug or treatment shortening regimen. In addition, predictive values of 2-month and 6-month SCC had high variability across levels of certain individual patient characteristics, so SCC as a predictor of outcome may not be equally useful in all patient groups. For these reasons, both 2- and 6-month SCC appear imperfect prognostic markers for MDR-TB treatment outcome. However in terms of the balance of prognostic accuracy, 6-month SCC clearly outperformed 2-month SCC. In the present analysis, the PPV of 2-month and 6-month SCC supports the use of these proxy markers for end-of-treatment outcomes, but PPV depends on the prevalence of the condition of interest – treatment success in this analysis. The strength of association, represented by the calculation of the OR is more valid as it does not depend on prevalence and allows assessing the impact of individual patient characteristics. Furthermore, effect modification of the association of 2-month SCC with treatment outcome by HIV infection hampers its use as a marker of treatment outcome in HIV-infected patients, although this result should be interpreted with caution as only half of the patients received antiretroviral therapy.

The observed geographic variability in sensitivity and specificity of both 2- and 6-month SCC may reflect unmeasured or uncontrolled confounders. For example, countries had very different prevalence of baseline drug resistance.^{6,16} We cannot exclude variations in laboratory procedures for culture in different countries, although all sites had quality assured laboratories and used solid culture for microbiological monitoring of treatment. Differences in time-to-culture conversion and SCC rates by geographic location have been reported

previously in the context of a randomized, controlled clinical trial.¹⁷ Such high degree of geographic variability would thus indicate limited reproducibility of these markers.

Two-month SCC has been used as an early microbiological endpoint in clinical trials of drug-susceptible TB for a long time, based on an observed trial level correlation with relapse in the series of British Medical Research Council trials.² A recent re-analysis of these data indicate that month 3 culture outperforms month 2 culture as a surrogate marker, but both are imperfect, with evidence of geographical variation.¹⁸ The results of three large RCTs of treatment-shortening fluoroquinolone-containing regimens in drug-susceptible TB patients suggest that 2-month culture conversion does not translate directly to inferences about the overall duration of treatment without more careful modelling.¹⁹⁻²² For these reasons, and the longer time-to-culture conversion observed in cohorts of MDR-TB patients, as well as the long duration and variability of the continuation phase after culture conversion,^{9,23} extrapolating 2-month SCC as a surrogate marker of treatment outcome to MDR-TB drug trials is questionable. In a literature review of 19 studies of MDR-TB patients prepared by the FDA, the majority of studies reported SCC at a median of 2 months.⁵ Studies reporting the lowest SCC rates in general had the highest rates of relapse and highest mortality, however this association was not universal, and the FDA concluded that “SCC can be used to predict the long-term clinical outcome of cure/relapse with a reasonable degree of confidence in MDR-TB patients”.⁵ The high specificity of 2-month SCC suggests that treatment resulting in SCC by 2 months will probably result in a long-term benefit, but the low sensitivity means that many treatments leading to long term cure may not meet the criterion of SCC by 2 months.

Data on predictive value of 6-month SCC are also limited. Our literature search identified only two publications reporting data that would allow calculation of predictive values and association of 6-month SCC with treatment outcome.^{24,25} Comparing successful treatment outcome with failure/death, 6-month SCC has a sensitivity of 89% and specificity of 88% in Holtz et al²⁴ and 92% and 25%, respectively, in Joseph et al.²⁵ SCC by the end of the 6th month of treatment was significantly associated with treatment success in Holtz et al (RR=3.80), but did not reach statistical significance in Joseph et al. The high sensitivity of 6-month SCC suggests that treatments not resulting in SCC by 6-months are unlikely to have lasting benefit, while the lower specificity means that treatments meeting this criterion may not be ultimately efficacious.

This evaluation has important limitations. Results are based on two observational cohort studies (prospective and retrospective), not randomized clinical trials. Early surrogate markers of the efficacy of a specific treatment remain valid only if the treatment continues according to plan. In our studies, treatment was adjusted as required by the clinical circumstances. To the extent that substantial changes in treatment depended on serial sputum culture results, the sensitivity and specificity would have decreased. We were unable to take relapses into account because such data were not available. We did not have data on the cause of death and could not ascertain that all deaths were TB-related. Further, data on treatment regimens were not available in the Case-Based Study, thus we were unable to include specific characteristics of the treatment regimen in this assessment. HIV-status, which was an important effect modifier, was unknown for 30% of patients, but this was

primarily due to one country that did not routinely test TB patients for HIV infection because the prevalence of HIV infection was known to be low in that country at the time of these studies. A relatively large proportion of patients were excluded due to unknown susceptibility to FQ or SLI drugs (27%, 839/3,059 culture-positive patients). It should also be noted that 331 out of the 2043 (16%) patients included in the study were lost to follow up during or after treatment. These patients were excluded from analysis, but it is possible that they represented poor treatment outcomes compared to those fully retained in the treatment cohorts. The number of patients in certain predictors' strata was low limiting the power of the multivariable analysis. Lastly, cultures were performed on solid media, and the results may differ when using liquid media.

Our analysis focused on the utility of SSC as a potential surrogate endpoint for controlled clinical trials on MDR-TB. These findings do not necessarily apply to clinical practice. SCC is only one part of assessing the effect of treatment, and patients are re-evaluated every time they see their physician, which may be daily for inpatients or monthly for ambulatory patients. If a patient has SCC at 2 months, it would give some assurance that treatment is effective. However, the absence of SCC at 2 months may be too early to change treatment, unless the patient's condition is deteriorating, because the median time to SCC in our study was 3 months. Conversely, 6 months may be too long to wait for culture conversion, again, depending on the overall clinical picture. In most patients, physicians would not wait for SCC at 6 months before re-evaluating the patient and adjusting treatment.

Confirmatory Phase III trials are expensive and time-consuming, especially for MDR-TB,⁴ so the question of which are the shortest, simplest, most effective, and safest regimens to be tested in confirmatory Phase III trials becomes a major public health challenge. As new drugs and new combinations of drugs are being investigated for the treatment of MDR-TB,²⁶ and while novel adaptive trial designs are being proposed to identify rapidly the best drugs/combinations of drugs to be advanced to Phase III trials,²⁷ the definition and relevance of the optimal surrogate marker to be used to measure treatment efficacy is of paramount importance. In this study, we show that time to SCC on solid media, SCC at 6 months, and SCC at 2 months among patients without known HIV infection could be considered as surrogate markers of treatment outcome in MDR-TB trials. However, the overall association with outcome was substantially stronger for 6-month compared to 2-month SCC, and 6-month SCC was valid among HIV-infected patients as well as those without known HIV infection. Our results, however, show a series of major limitations that have to be taken into consideration when using these markers as surrogates of treatment efficacy in MDR-TB.¹⁵

Text for "Research in Context" panel

Systematic Review

In 2009 the US FDA carried out an exhaustive search of the published literature on early endpoints as surrogates for cure in the treatment of MDR-TB (5). The search strategy, resulting bibliography, and summaries of results are available on the FDA's website (5). To update the FDA's bibliography, we searched MEDLINE and EMBASE using the terms "tuberculosis", "multidrug resistant", "MDR TB", "treatment outcomes", "relapse",

“recurrence” and “culture conversion” for publications in the past five years. In addition, the trials register <http://www.clinicaltrials.gov> was reviewed for applicable studies not found in the previous search. After discarding duplicate citations, a total of 112 publications were obtained through the literature search for which full-text was reviewed. Six of these were review articles, one was a set of guidelines, and three were reports of individual cases. Excluding two of our own previous publications on different aspects of these two cohorts, none of the remaining articles focused on culture conversion as a predictor of treatment outcome.

Interpretation

Two new drugs were provisionally approved for the treatment of MDR-TB, bedaquiline in 2012 and delamanid in 2013—the first anti-TB drugs to be approved since the 1960s. The approvals were provisional because they were based on sputum culture conversion as a surrogate for efficacy in relatively small numbers of patients in Phase IIB trials. The surrogate endpoint and Phase II trials were accepted as the basis for provisional approval because treatment of MDR-TB succeeds in only 65% of cases in published reports and 48% of cases under program conditions. These new drugs, together with other potential new therapies (e.g., linezolid, sutezolid, pretonamid), have created rapidly increasing interest in and support for clinical trials of MDR-TB treatment. Because MDR-TB treatment currently takes 18-24 months, Phase III trials of new treatment regimens will take many years to complete, entailing great expense. Therefore, Phase II trials with endpoints based on early, accurate surrogates of efficacy will be crucial in choosing regimens wisely for future Phase III trials. Historically, 2-month SCC has been used as a surrogate endpoint in clinical trials of treatment for drug-susceptible TB. Treatment of drug-susceptible TB, however, takes as little as six months, and two months is the duration of the initial intensive phase of treatment. Because MDR-TB requires 18-24 months treatment, and because the duration of the initial intensive phase typically is 6 months (4 to 8 months), 2-month SCC may not be as closely correlated with treatment outcome in MDR-TB as it is in drug-susceptible TB. Therefore, it is critical to evaluate the validity of SCC at various time points as an early predictor of treatment efficacy before investing considerable resources in large, longer-term Phase III trials of new MDR-TB treatment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

The United States Agency for International Development (USAID) and CDC provided funding for both the Preserving Effective TB Treatment Study (PETTS) study and the DOTS-Plus Pilot Projects Case-Based Study. The Division of Intramural Research of NIAID/NIH and the Republic of Korea's CDC supported in part the Korean sites.

Disclaimer

The conclusions and interpretations of data presented in this report are solely those of the authors and do not necessarily represent an official position of CDC, the U.S. government, World Health Organization (WHO) or the governments of the nine participating countries.

References

1. Food and Drug Administration (U.S.). FDA News Release. Dec 31.2012
2. Mitchison DA. Assessment of new sterilizing drugs for treating pulmonary tuberculosis by culture at 2 months. *The American review of respiratory disease*. Apr; 1993 147(4):1062–1063. [PubMed: 8466107]
3. Wallis RS, Doherty TM, Onyebujoh P, et al. Biomarkers for tuberculosis disease activity, cure, and relapse. *Lancet Infect Dis*. Mar; 2009 9(3):162–172. [PubMed: 19246020]
4. Lienhardt C, Davies G. Methodological issues in the design of clinical trials for the treatment of multidrug-resistant tuberculosis: challenges and opportunities. *Int J Tuberc Lung Dis*. May; 2010 14(5):528–537. [PubMed: 20392344]
5. US Food and Drug Administration. Development of Drugs to Treat Multi-Drug Resistant Tuberculosis (MDR-TB). AIDAC Advisory Committee Meeting; June 3, 2009; <http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Anti-InfectiveDrugsAdvisoryCommittee/ucm161856.htm>
6. Dalton T, Cegielski P, Akksilp S, et al. Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study. *Lancet*. Oct 20; 2012 380(9851):1406–1417. [PubMed: 22938757]
7. Gammino V, Taylor A, Rich M, et al. Bacteriologic monitoring of multidrug-resistant tuberculosis patients in five DOTS-Plus pilot projects. *Int J Tuberc Lung Dis*. 2011; 15(10):1315–1322. [PubMed: 22283887]
8. Kurbatova EV, Gammino VM, Bayona J, et al. Predictors of sputum culture conversion among patients treated for multidrug-resistant tuberculosis. *Int J Tuberc Lung Dis*. Oct; 2012 16(10):1335–1343. [PubMed: 23107633]
9. World Health Organization. Guidelines for the programmatic management of drug-resistant tuberculosis. Emergency Update. Geneva, Switzerland: 2008.
10. Ahuja SD, Ashkin D, Avendano M, et al. Multidrug resistant pulmonary tuberculosis treatment regimens and patient outcomes: an individual patient data meta-analysis of 9,153 patients. *PLoS medicine*. 2012; 9(8):e1001300. [PubMed: 22952439]
11. Johnston JC, Shahidi NC, Sadatsafavi M, Fitzgerald JM. Treatment outcomes of multidrug-resistant tuberculosis: a systematic review and meta-analysis. *PLoS One*. 2009; 4(9):e6914. [PubMed: 19742330]
12. Orenstein EW, Basu S, Shah NS, et al. Treatment outcomes among patients with multidrug-resistant tuberculosis: systematic review and meta-analysis. *Lancet Infect Dis*. Mar; 2009 9(3): 153–161. [PubMed: 19246019]
13. Harbord R, Deeks J, Egger M, Whiting P, Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 2007; 8(2):239–251. [PubMed: 16698768]
14. Reitsma J, Glas A, Rutjes AWS, Scholten RJ, Bossuyt PM, Zwinderman A. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of clinical epidemiology*. 2005; 58(10):982–990. [PubMed: 16168343]
15. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in medicine*. 1989; 8(4):431–440. [PubMed: 2727467]
16. Nathanson E, Lambregts-van Weezenbeek C, Rich ML, et al. Multidrug-resistant tuberculosis management in resource-limited settings. *Emerg Infect Dis*. Sep; 2006 12(9):1389–1397. [PubMed: 17073088]
17. Mac Kenzie WR, Heilig CM, Bozeman L, et al. Geographic differences in time to culture conversion in liquid media: Tuberculosis Trials Consortium study 28. Culture conversion is delayed in Africa. *PLoS One*. 2011; 6(4):e18358. [PubMed: 21494548]

18. Phillips PPJ, Fielding K, Nunn A. An evaluation of culture results during treatment for tuberculosis as surrogate endpoints for treatment failure and relapse. *PLoS ONE*. 2013; 8(5):e63840–e63840. [PubMed: 23667677]
19. Gillespie SH, Crook AM, McHugh TD, et al. Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. *N Engl J Med*. Oct 23; 2014 371(17):1577–1587. [PubMed: 25196020]
20. Jindani A, Harrison TS, Nunn AJ, et al. High-dose rifapentine with moxifloxacin for pulmonary tuberculosis. *N Engl J Med*. Oct 23; 2014 371(17):1599–1608. [PubMed: 25337749]
21. Merle CS, Fielding K, Sow OB, et al. A four-month gatifloxacin-containing regimen for treating tuberculosis. *N Engl J Med*. Oct 23; 2014 371(17):1588–1598. [PubMed: 25337748]
22. Wallis RS, Wang C, Meyer D, Thomas N. Month 2 Culture Status and Treatment Duration as Predictors of Tuberculosis Relapse Risk in a Meta-Regression Model. *PLoS ONE*. 2013; 8(8):e71116.10.1371/journal.pone.0071116 [PubMed: 23940699]
23. World Health Organization. Guidelines for the programmatic management of drug-resistant tuberculosis. 2011 update. Geneva, Switzerland: 2011.
24. Holtz TH, Sternberg M, Kammerer S, et al. Time to sputum culture conversion in multidrug-resistant tuberculosis: predictors and relationship to treatment outcome. *Ann Intern Med*. May 2; 2006 144(9):650–659. [PubMed: 16670134]
25. Joseph P, Desai VB, Mohan NS, et al. Outcome of standardized treatment for patients with MDR-TB from Tamil Nadu, India. *The Indian journal of medical research*. May.2011 133:529–534. [PubMed: 21623039]
26. Resist-TB Progress Report: Clinical Trials for DR-TB.
27. Phillips PPJ, Gillespie S, Boeree M, et al. Innovative trial designs are practical solutions for improving the treatment of tuberculosis. *The Journal of infectious diseases*. 2012; 205(Suppl 2):S250–S257. [PubMed: 22448027]

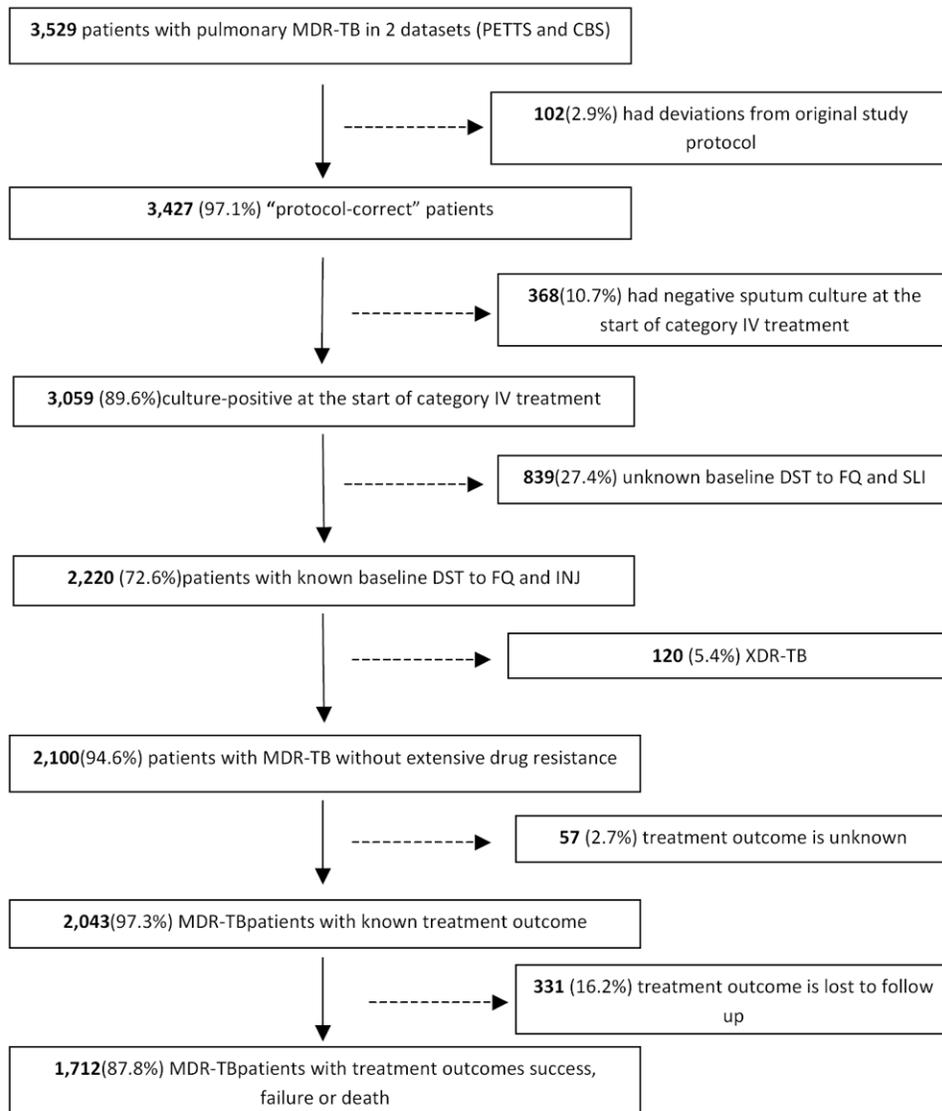


Figure 1. Diagram of patient population and derivation of study participants included in analysis

Note. TB=tuberculosis. MDR=multidrug-resistant. XDR=extensively drug-resistant.

PETTS=Preserving Effective Tuberculosis Treatment Study. CBS=Case-based study.

DST=drug-susceptibility testing. FQ=fluoroquinolone. SLI=second-line injectable drug.

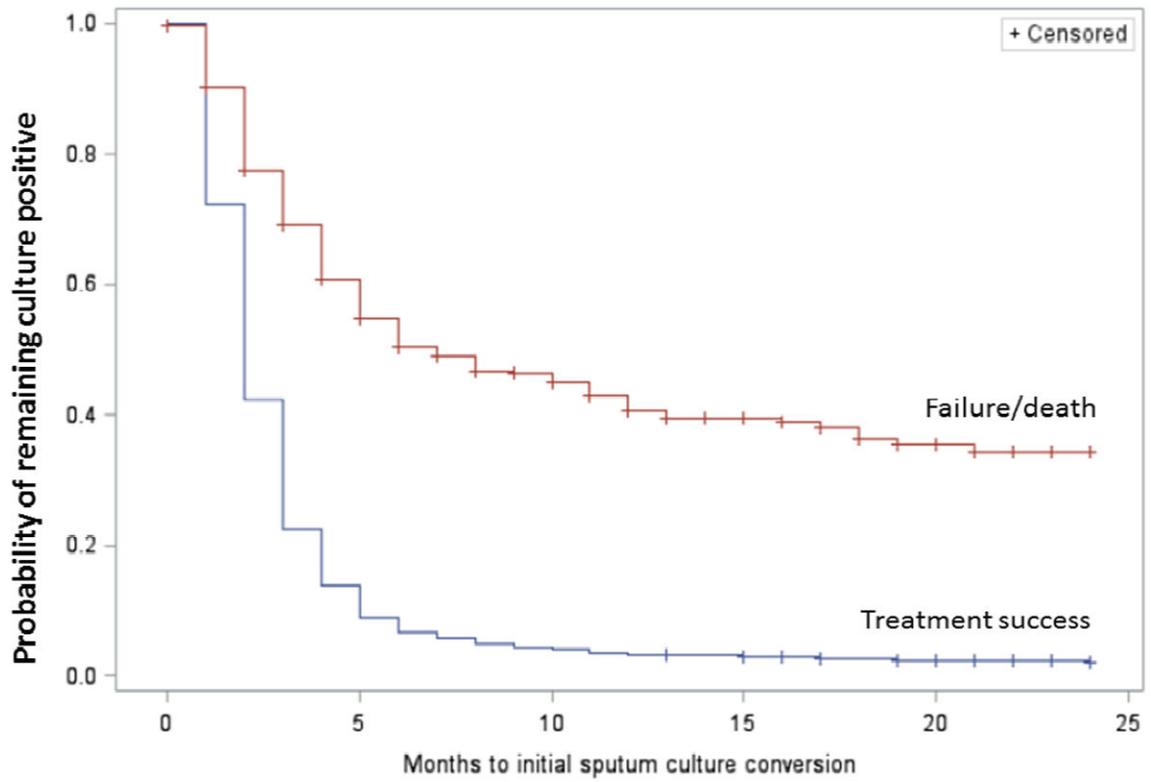
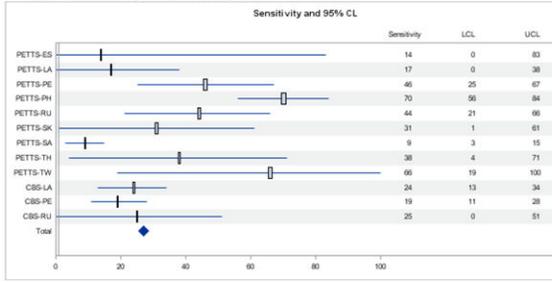


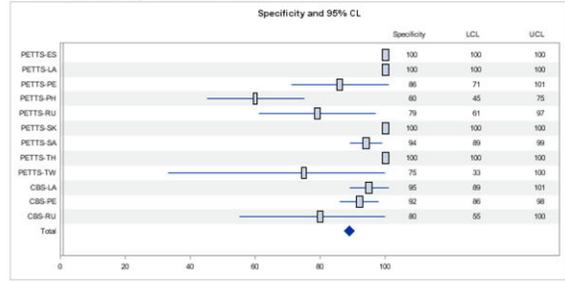
Figure 2. Time to sputum culture conversion among MDR-TB patients by treatment outcome (success versus failure or death), N=1,712

Time to initial sputum culture conversion was truncated at 24 months.

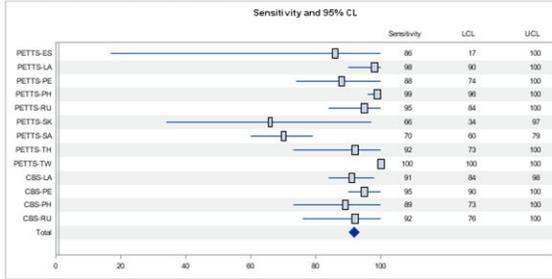
A. Sensitivity of 2-month SCC



B. Specificity of 2-month SCC



C. Sensitivity of 6-month SCC



D. Specificity of 6-month SCC

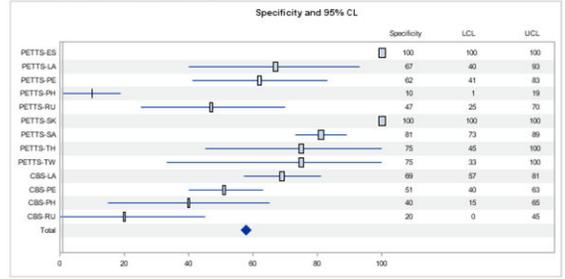


Figure 3. Forest plots for empiric (crude) sensitivity and specificity of initial sputum culture conversion (SCC) in predicting treatment outcome, by country/study. N=1,712
 Note. CL=confidence limits. LCL=lower confidence limit. UCL=upper confidence limit. PETTS=Preserving Effective Tuberculosis Treatment Study. CBS=Case-based study. ES=Estonia. LA=Latvia. PE=Peru. PH=Philippines. RU=Russia. SK=Republic of South Korea. SA=South Africa. TH=Thailand. TW=Taiwan.

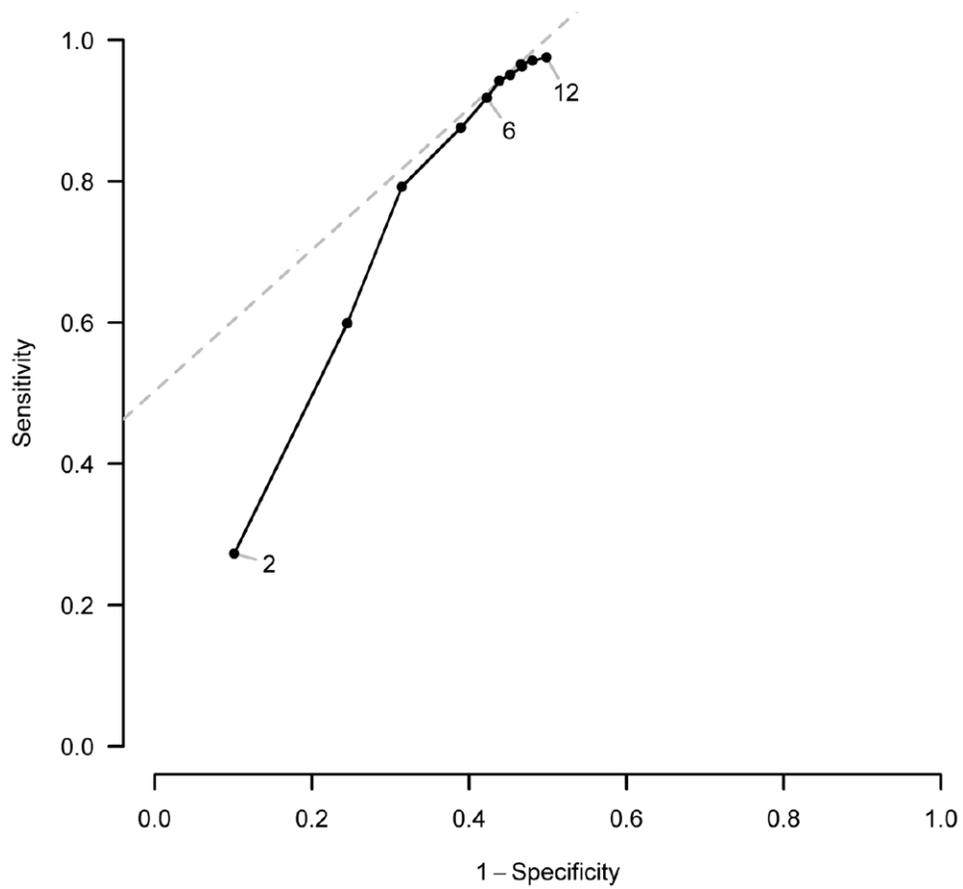


Figure 4. Diagnostic performance of timing of initial sputum culture conversion in predicting treatment outcome (success versus failure or death), N=1,712

Note. Estimates of sensitivity and specificity are based on the results of bivariate generalized linear mixed model unadjusted for covariates. Numbers on the ROC curve represent month of SCC status assessment.

Table 1

Association of sputum culture conversion status with treatment outcome (success versus failure or death), unadjusted for covariates. N=1,712

Month of treatment	Sputum culture conversion status	Success n (%) N=1,344	Failure or death n (%) N=368	OR* (95% CI)
2 months	Converted	458 (91.8)	41 (8.2)	3.60 (2.49-5.19)
	Did not convert	886 (73.0)	327 (27.0)	1.00
6 months	Converted	1,231 (89.1)	151 (10.9)	15.18 (11.23-20.53)
	Did not convert	113 (34.2)	217 (65.8)	1.00

OR=odds ratio. CI=confidence interval.

* Estimated using univariate random effects logistic regression.

Table 2

Multivariable analysis of association of sputum culture conversion and covariates with treatment outcome (success versus failure or death). N=1,712

Covariates	Model with 2-month SCC as a main predictor of outcome	Model with 6-month SCC as a main predictor of outcome
	aOR (95% CI)	aOR (95% CI)
SCC 2-month		
among HIV positive	0.38 (0.12-1.18)	-
among HIV negative	4.12 (2.25-7.54)	-
among HIV unknown	3.59 (1.96-6.58)	-
SCC 6-month	-	14.07 (10.05-19.71)
HIV infection		
Positive	-	0.64 (0.36-1.14)
Negative	-	1.00
Unknown	-	0.60 (0.39-0.91)
Prospective study	0.53 (0.27-1.06)	0.89 (0.52-1.52)
Male gender	0.88 (0.66-1.17)	0.91 (0.67-1.25)
Age (per year)	1.00 (0.99-1.01)	0.99 (0.98-1.01)
History of previous treatment		
Treated with first-line drugs	0.59 (0.36-0.96)	0.86 (0.51-1.44)
Treated with second-line drugs	0.34 (0.20-0.59)	0.57 (0.32-1.01)
Unknown	0.60 (0.14-2.55)	1.19 (0.24-5.95)
New case	1.00	1.00
BMI		
<18.5	0.37 (0.28-0.50)	0.36 (0.26-0.50)
Unknown	0.61 (0.35-1.04)	0.78 (0.43-1.44)
18.5	1.00	1.00
Cavitary disease		
Yes	0.85 (0.63-1.16)	0.85 (0.61-1.19)
Unknown	0.50 (0.27-0.91)	0.73 (0.38-1.41)
No	1.00	1.00
AFB smear status at start of treatment		
Positive	0.98 (0.68-1.42)	1.12 (0.75-1.67)
Unknown	0.77 (0.44-1.37)	0.65 (0.35-1.20)
Negative	1.00	1.00
Baseline resistance to ≥ 1 FQ	0.37 (0.23-0.58)	0.46 (0.27-0.77)
Baseline resistance to ≥ 1 SLI	0.41 (0.30-0.57)	0.50 (0.35-0.71)

Note. Bold type face indicates statistically significant results at $P < 0.05$. aOR=adjusted odds ratio.

CI=confidence interval. FQ= fluoroquinolones. SLI=second-line injectable drug. Adjusted odds ratios are estimated using multivariable random effects logistic regression.

Table 3

Sensitivity and specificity of 2-month and 6-month sputum culture conversion in predicting successful (vs. failure or death) treatment outcome stratified by covariates *, N=1,712.

Predictors	Value of predictor	Timing of initial sputum culture conversion			
		2 months		6 months	
		Sensitivity, % Mean (95% CL)	Specificity, Mean (95% CL)	Sensitivity, % Mean (95% CL)	Specificity, % Mean (95% CL)
OVERALL	-	27.3 (16.6-41.4)	89.8 (82.3-94.4)	91.8 (85.9-95.4)	57.8 (42.5-71.6)
Study	Prospective	34.8 (20.6-52.2)	86.4 (75.3-93.0)	91.3 (82.9-95.8)	62.7 (42.0-79.6)
	Retrospective	14.6 (5.4-33.8)	94.8 (85.7-98.2)	92.9 (81.1-97.5)	49.7 (23.7-75.8)
Gender	Male	27.4 (16.5-41.8)	91.6 (84.3-95.7)	91.4 (85.0-95.2)	57.2 (41.4-71.6)
	Female	27.0 (16.0-41.9)	86.0 (73.9-93.1)	92.8 (86.6-96.2)	58.8 (41.1-74.4)
Age	45 years	21.8 (12.3-35.6)	92.6 (83.6-96.8)	92.4 (86.0-96.0)	55.6 (38.1-71.8)
	<45 years	29.7 (17.9-45.0)	88.4 (79.4-93.8)	91.4 (85.1-95.2)	58.5 (42.4-72.9)
HIV	Positive	18.3 (7.4-38.3)	73.8 [†] (47.6-89.8)	85.7 (72.3-93.2)	61.1 (37.4-80.5)
	Negative	27.4(16.6-41.8)	92.6 (84.3-96.7)	90.7 (84.5-94.6)	62.1 (46.0-75.9)
	Unknown	28.7 (16.2-45.5)	90.7 (79.5-96.1)	95.7 [†] (90.6-98.1)	48.1 (29.8-66.9)
Previous treatment history	FLD	26.2 [†] (15.6-40.4)	90.3 (82.1-95.0)	91.0 [†] (84.5-95.0)	55.4 (39.2-70.5)
	New	35.8 (21.6-53.0)	85.6 (63.4-95.4)	96.5 (92.0-98.5)	46.3 (24.6-69.5)
	SLD	21.4 [†] (11.9-35.5)	90.5 (80.1-95.8)	88.9 [†] (79.5-94.2)	65.2 (47.9-79.3)
BMI, kg/m²	<18.5	25.6 (15.5-39.3)	89.7 (81.1-94.6)	93.3 (86.9-96.7)	58.9 (41.2-74.7)
	18.5	29.2 (18.6-42.8)	89.8 (80.8-94.9)	91.9 (85.6-95.5)	54.0 (36.4-70.6)
Cavitary disease	Yes	25.1 [†] (15.2-38.5)	90.0 (82.1-94.6)	92.0 (85.6-95.7)	59.6 (44.5-73.0)
	No	32.7 (20.4-48.1)	89.4 (78.9-95.0)	93.2 (87.0-96.6)	49.6 (32.3-67.0)
AFB smear positive	Yes	21.8 [†] (12.7-34.8)	91.1 (84.1-95.2)	90.8 [†] (83.7-95.0)	62.2 [†] (46.2-75.9)
	No	49.0 (32.6-65.6)	83.3 (66.8-92.5)	96.6 (92.0-98.6)	46.5 (28.0-66.1)
Baseline FQ resistance[‡]	Yes	84.6 [†] (68.3-93.3)	70.8 (50.0-85.4)
	No	92.5 (86.8-95.8)	54.8 (39.2-69.6)
Baseline SLI resistance	Yes	26.0 (14.8-41.4)	88.7 (76.6-95.0)	89.9 (80.9-94.9)	61.9 (44.5-76.7)
	No	27.6 (16.8-41.8)	90.4 (82.4-95.0)	92.2 (86.4-95.7)	56.1 (40.8-70.4)

Note. CL=confidence limits.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

*Based on the results of bivariate generalized linear mixed model adjusted for one covariate at a time.

† Significant difference in predictive values across levels of predictor based on calculation of odds ratios with respective 95% confidence intervals.

‡ Model for 2-months SCC and baseline fluoroquinolone resistance did not converge.