Open Forum Infectious Diseases

MAJOR ARTICLE

A Clinical Prediction Score Including Trial of Antibiotics and C-Reactive Protein to Improve the Diagnosis of Tuberculosis in Ambulatory People With HIV

Tom H. Boyles,[1] Matilda Nduna,[2] Thalitha Pitsi,[3] Lesley Scott,[4] Matthew P. Fox,[5,6,7] and Gary Maartens[2]

[1]Department of Medicine, University of Cape Town, Cape Town, South Africa, [2]Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, South Africa, [3]Hillbrow Community Healthcare Clinic, Johannesburg, South Africa, [4]Department of Molecular Medicine and Haematology, Faculty of Health Sciences, School of Pathology, University of Witwatersrand, Johannesburg, South Africa, [5]Department of Global Health, Boston University School of Public Health, Boston, Massachusetts, USA, [6]Department of Internal Medicine, Health Economics and Epidemiology Research Office, Faculty of Health Sciences, School of Clinical Medicine, University of the Witwatersrand, Johannesburg, South Africa, and [7]Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, USA

Background. The use of a "trial of antibiotics" as empiric therapy for bacterial pneumonia as a diagnostic tool for tuberculosis in people with HIV (PWH) was removed from World Health Organization (WHO) recommendations in 2007, based on expert opinion. Current guidelines recommend antibiotics only after 2 Xpert MTB/RIF tests (if available), chest x-ray, and clinical assessment have suggested that tuberculosis is unlikely. Despite this, a "trial of antibiotics" remains common in algorithms in low-resource settings, but its value is uncertain. C-reactive protein (CRP), which has been proposed as a "rule-out" test for tuberculosis, may be an objective marker of response to antibiotics.

Methods. We performed a passive case-finding cohort study of adult PWH with a positive WHO symptom screen. All participants received antibiotics at first visit according to the local protocol and were reviewed to ascertain clinical response. Point-of-care CRP was measured at both visits. All patients had sputum tested with Xpert MTB/RIF Ultra (Ultra), and the reference standard was based on 2 sputum mycobacterial cultures. We explored multivariable prediction models (MPM) for tuberculosis based on 1 or 2 visits.

Results. Seventy-five of 207 patients (36%) had confirmed tuberculosis. Clinical response to antibiotics after 2 days was a good predictor of disease. An MPM based on 2 visits, without CRP, had acceptable discrimination (c-statistic, 0.75) and calibration (goodness-of-fit $P = .07$). Addition of CRP after antibiotics improved the model moderately (c-statistic, 0.78). CRP at first visit was not an independent predictor of tuberculosis.

Conclusions. In adult PWH seeking care for symptoms suggestive of tuberculosis, lack of response to antibiotics is a strong predictor of disease and is likely to be useful, particularly when access to Ultra is limited. CRP adds value when measured after antibiotics but is of limited value at first visit.

Keywords. C-reactive protein; diagnostic accuracy; multivariable prediction models; WHO algorithm; Xpert MTB/RIF Ultra.

Limitations of current diagnostics remain a challenge in the fight against tuberculosis. A trial of antibiotics (ToA) is defined as a course of broad-spectrum antibiotics, with negligible *Mycobacterium tuberculosis* activity, given to patients suspected of having tuberculosis [1]. Patients with a clinical response to antibiotics, which is not consistently defined, are considered unlikely to have tuberculosis and vice versa. The 2003 World Health Organization (WHO) guidelines for diagnosing tuberculosis, which did not differentiate by HIV status, included a ToA after 3 negative sputum smears, with clinical improvement ruling out tuberculosis [1]. In 2007, the revised WHO guideline for HIV-prevalent and resource-constrained settings [2] stated that the primary role of antibiotics should not be as a diagnostic aid; instead they should be used to treat concomitant bacterial infection in people with HIV/AIDS (PWH) with cough. The level of evidence for this recommendation was IV (expert opinion based on evaluation of other evidence). The 2007 algorithm was shown to expedite the diagnosis of smear-negative pulmonary tuberculosis and increase diagnostic accuracy compared with the 2003 algorithm [3], although the diagnostic accuracy remained suboptimal [4].

In 2016, the WHO published a revised algorithm for diagnosing tuberculosis in ambulatory PWH [5]. All patients should have Xpert MTB/RIF (Cepheid, Sunnyvale, CA, USA) testing of sputum, although the WHO has subsequently endorsed the more sensitive Xpert MTB/RIF Ultra (Ultra) assay.

Received 29 August 2019; editorial decision 30 December 2019; accepted 2 January 2020.

Correspondence: Tom Boyles, Division of Infectious Diseases, Helen Joseph Hospital, Auckland Park, Johannesburg, South Africa (tomboyles@yahoo.com).

Open Forum Infectious Diseases®

DOI: 10.1093/ofid/ofz543

If Xpert is negative or unavailable, the algorithm recommends further investigations (chest x-ray, repeat Xpert MTB/RIF, and mycobacterial culture if available) at the first visit, with antibiotics only advised if tuberculosis remains unlikely at this point. Patients with a response to antibiotics are presumed not to have active tuberculosis and should initiate isoniazid preventive therapy; patients with no or partial response to antibiotics should have further tuberculosis investigations, rather than empiric antituberculosis treatment. To our knowledge, this algorithm has not been formally validated.

Despite being removed from the WHO guidelines in 2007, ToA is common in diagnostic algorithms in resource-limited settings, although its value remains uncertain. A recent systematic review of 8 studies including 2586 participants found a pooled sensitivity and specificity of a ToA of 0.72 and 0.77, respectively [6].

The specificity of symptom screening for tuberculosis in PWH, which is used as the entry criterion for diagnostic algorithms, is 0.71 for patients on antiretroviral therapy (ART; sensitivity, 0.51) and 0.28 for ART-naïve patients (sensitivity, 0.89) [7], so large numbers of patients without tuberculosis enter the algorithm but are Xpert-negative. This poses a huge challenge in resource-constrained settings. Multivariable prediction models (MPM) have been proposed as a way to prescreen symptomatic patients and focus investigations on those with increased probability of disease [8–10], but none have been externally validated.

C-reactive protein (CRP), which is now available as a point-of-care (POC) test costing <$2, has been evaluated as an initial test for tuberculosis in patients with HIV. A systematic review and meta-analysis of diagnostic accuracy studies found that a value <10 mg/L has a sensitivity of 0.93 and specificity of 0.61 among outpatients with HIV [11]. Analysis of 3 studies including only patients self-reporting symptoms found a higher sensitivity (range, 0.96–0.97) and lower specificity (range, 0.33–073). CRP has not been evaluated as an additional variable to clinical prediction rules or as an objective marker of response to antibiotics for the diagnosis of HIV-associated tuberculosis.

We performed a passive case-finding cohort study of PWH meeting entry criteria for the 2016 WHO algorithm in a setting where local standard operating procedures (SOPs) are for all patients to receive antibiotics at the first visit. We aimed to develop MPMs including subjective response to antibiotics and POC measurements of CRP.

The study is reported according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [12] and the Standards for Reporting of Diagnostic Accuracy Studies (STARD) statement [13].

Objectives

To determine the independent predictors of tuberculosis at index visit and following a trial of empiric antibiotics among adult PWH seeking clinic care with compatible symptoms.

To develop multivariable prediction models (MPMs) for tuberculosis based on independent predictors at index and return visit.

METHODS

Study Participants and Data Collected

We recruited HIV-positive patients at a community health care clinic in Johannesburg, South Africa, with a positive WHO symptom screen (any 1 of current cough, fever, night sweats, or weight loss) based on research team capacity, from the queue of adults actively seeking care for their symptoms. Recruitment was therefore nonconsecutive and nonrandom. We included patients regardless of ART status, receipt of isoniazid preventive therapy, or prior tuberculosis. We excluded those who received antibiotics within 14 days to minimize the possibility of referral bias, as well as those currently taking tuberculosis treatment.

At the first visit, all patients were systematically evaluated for tuberculosis with vital signs, body mass index (BMI), sputum smear microscopy, Ultra, and liquid mycobacterial culture (mycobacterial growth indicator tube [MGIT], BACTEC MGIT 960 TB System). Sputum induction with nebulized hypertonic saline (5%) was used when participants were unable to produce sputum spontaneously. Ultra was repeated if the first result was indeterminate. CRP was measured at the point of care (Abbot Afinion AS100 analyzer, the range of values was 5–>200 mg/L), and antibiotics for respiratory pathogens (amoxicillin, amoxicillin-clavulanate, or azithromycin) were dispensed as part of usual clinical care. Serum and urine samples were stored. After a lead-in period of 20 patients, M.N., who has 20 years of nursing experience assessing patients for tuberculosis, was asked if tuberculosis was likely or unlikely after clinically assessing the patient and reviewing the CRP results.

Patients were followed up after 2–5 days when the results of the Ultra were known. Vital signs were repeated, and patients were asked if their symptoms were better, about the same, or worse; a clinical response to antibiotics was defined as symptoms being reported as better. CRP and sputum mycobacterial culture were repeated for all patients. All patients with positive Ultra or tuberculosis culture on sputum were initiated on antituberculosis therapy according to national guidelines, and those who remained symptomatic were referred to their clinic for follow-up care.

Outcomes

The primary outcome was active tuberculosis, defined as ≥1 sputum culture positive for *Mycobacterium tuberculosis*. Patients with 2 negative cultures or those with 1 negative and 1 contaminated culture and who were asymptomatic after 6 weeks without antituberculous therapy were considered not to have tuberculosis.

Predictors

There is no consensus on the best method for selecting candidate variables for MPM, but suggested approaches include using literature review, clinical knowledge, and studying the distribution of predictors in the study data. A priori, we considered predictors known to be associated with prevalent and/or incident tuberculosis among PWH at index visit: age, sex, ART status, CD4 count, duration of symptoms, temperature, number of tuberculosis symptoms, current smoking status, BMI, and CRP.

We defined ART status as no ART or on ART <3 months vs on ART for >3 months at the time of presentation. Patients who had interrupted ART for >3 months before presentation were categorized as ART <3 months.

We also considered factors that may be predictive of tuberculosis at a second visit following a course of oral antibiotics: CRP, change in CRP, and subjective improvement of symptoms. Change in CRP was measured on a logarithmic scale due to the exponential decline in CRP when patients are treated for community-acquired pneumonia [14].

Sample Size

It is recommended, to ensure predictive accuracy, that the total number of candidate predictors be limited so that there are at least 10 outcomes for each candidate predictor studied [15]. We predicted a tuberculosis prevalence rate of 35% and chose a sample size of 200 to ensure at least 70 outcomes so that up to 7 candidate predictors could be included in the final models.

Statistical Analysis Methods

Baseline characteristics were described as proportions or medians. We used simple logistic regression to assess the ability of selected variables to predict tuberculosis.

Model development used a backward stepwise approach. A multiple logistic regression model using only the a priori selected variables was created (Supplementary Appendix). Variables with the highest *P* value were removed sequentially, with reduced models being compared with the original model using the log likelihood ratio test. Model development was complete if removal of a variable led to a statistically significant difference ($P < .05$) between the reduced model and the original model or when all predictors in a reduced model had $P < .1$. We considered models based on information available at visit 1 and visit 2 (excluding the Ultra result).

Final models were visually assessed by a calibration plot and by the Hosmer-Lemeshow goodness-of-fit test (GOF) [16]. An estimate of the c-statistic was used to assess discrimination (0.7–0.8 was deemed acceptable, 0.8–0.9 good, and ≥0.9 outstanding) [17]. Internal validation used 1000 bootstrap resamples [18]. A risk score was generated for each patient, which was converted to predicted risk using the equation predicted risk = $1/(1 + e^{-\text{risk score}})$.

To externally validate the clinical prediction model developed by Hanifa et al. [8], we obtained the measured predictors and outcome values in the participants of our study. We then quantified the discrimination and calibration of the XPHACTOR rule in our participants, using a calibration plot, GOF, and an estimate of the c-statistic.

Missing Data

The full data set was used, as <5% of the values were missing. In a sensitivity analysis, we used multiple imputation with chained equations. The upper limit of detection of the CRP assay was 200 mg/L. Values >200 were imputed as 300 based on the range of CRP values >200 seen in a similar patient group [19].

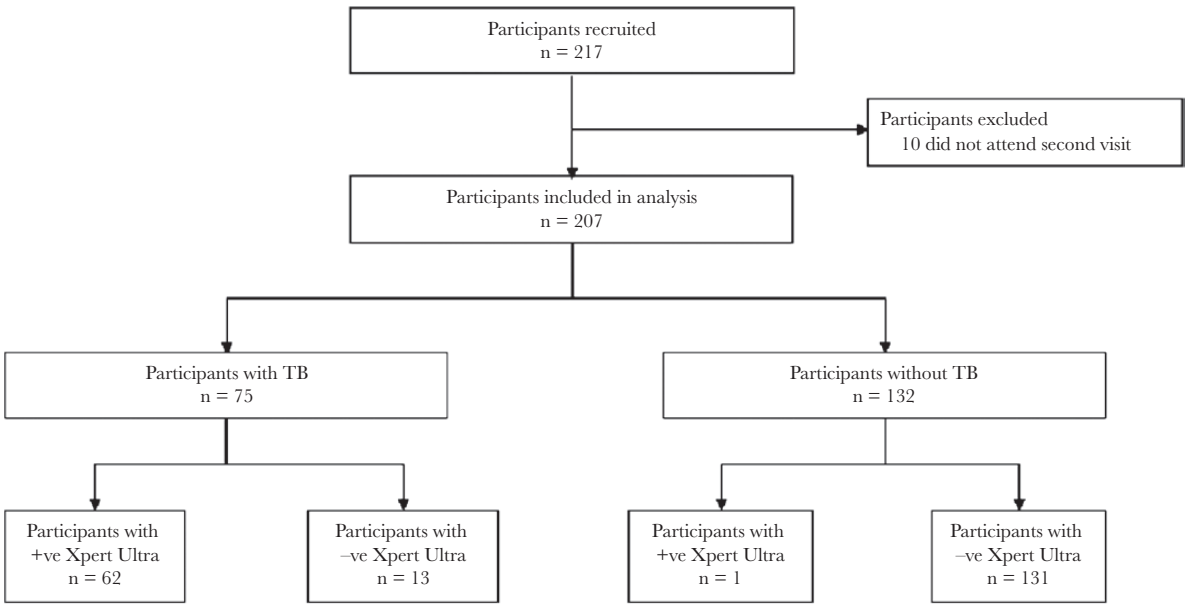

Figure 1. Participant flow diagram for 217 adults with HIV and symptoms suggestive of tuberculosis. Ultra + ve includes trace.

RESULTS

Two hundred seventeen participants were recruited between June 2018 and February 2019. Ten were excluded from the final analysis as they did not attend visit 2 and therefore had only 1 available sputum culture (Figure 1). One hundred thirty-eight participants (76%) were female, and the median age was 36 years. The median CD4 count was 185 cells/μL, 116 (56%) had been diagnosed with HIV within the past month, and most (95%) had cough (Table 1).

Seventy-five (36%) participants had confirmed tuberculosis, of whom 9 (12%) were sputum smear–positive (including scanty) and 62 (83%) were Ultra-positive (including trace) (Table 2). Two participants had confirmed rifampicin monoresistant tuberculosis, and 1 had multidrug resistance. Sensitivity and specificity of Ultra, compared with the outcome defined above, were 0.83 (95% confidence interval [CI], 0.72–0.90) and 0.99 (95% CI, 0.96–1.00), respectively; against a reference standard of 1

Table 1. Baseline Characteristics of 207 Adult Participants With HIV and Symptoms Suggestive of Tuberculosis

	All Participants (n = 207), No. (%) or Median (IQR)	Confirmed TB (n = 75), No. (%) or Median (IQR)	Not TB (n = 132), No. (%) or Median (IQR)
Female sex	138 (67)	54 (72)	84 (64)
Age, y	36 (31–41)	35 (30.5–39)	37 (31–41)
HIV/ART status			
Months since HIV diagnosis	0 (0–45)	0 (0–6.5)	0 (0–59)
Diagnosed with HIV <1 mo	116 (56)	51 (68)	65 (49)
Ever been on ART	82 (40)	22 (29)	60 (45)
Currently on ART	80 (39)	21 (28)	59 (45)
ART status <3 mo	145 (70)	61 (81)	84 (64)
Current CD4 count, cells/μL	185 (51–242)	114 (46–259)	227 (55–416)
TB history			
Previous TB	34 (17)	9 (12)	25 (19)
Previous IPT	4 (2)	2 (3)	2 (2)
Current IPT	10 (5)	2 (3)	8 (6)
WHO symptoms			
Duration of symptoms	94 (47)	44 (59)	50 (38)
≥14 d			
Cough	196 (95)	73 (97)	123 (93)
Fever	124 (60)	52 (69)	72 (55)
Night sweats	155 (75)	61 (81)	94 (71)
Weight loss	186 (90)	73 (97)	113 (86)
No. of symptoms	8 (4)	1 (1)	7 (5)
1	41 (20)	10 (13)	31 (23)
2	62 (30)	22 (29)	44 (33)
3	96 (46)	46 (61)	50 (38)
4			
Medical history			
Current smoker	60 (29)	23 (31)	37 (28)
Diabetes	1 (0.5)	1 (1)	0 (0)
Measurements			
BMI, kg/m^2	20.1 (18.3–22.3)	19.3 (18.2–21.7)	20.5 (18.7–22.8)
Pulse rate, bpm	76 (68–82)	78 (72–88)	74 (68–82)
Respiratory rate, bpm	14 (14–16)	16 (14–16)	14 (14–16)
Temperature, °C	184 (89)	62 (83)	122 (92)
36.0–37.4	19 (14)	10 (13)	9 (7)
37.5–38.9	4 (2)	3 (4)	1 (1)
39.0–40.0			
C-reactive protein, mg/L	57 (16.8–115)	74 (45–126)	38 (9–94.5)
Visit 2			
Days after visit 1	2 (2–4)	2 (2–4)	2 (2–4)
Symptoms improved	155 (75)	42 (56)	113 (86)
C-reactive protein, mg/L	34 (9–91)	81 (38–123)	16 (5–56)
Log C-reactive protein change	–0.103 (0.105 to –0.721)	–0.008 (–0.206 to 0.34)	0.22 (0 to 1.07)

Missing values: previous TB [4], duration of symptoms ≥ 14 days [6], night sweats [1], current smoker [3], diabetes [3], C-reactive protein visit 1 [3], symptoms improved [1], C-reactive protein visit 2 [4], log C-reactive protein change [6].

Abbreviations: ART, antiretroviral therapy; BMI, body mass index; IPT, isoniazid preventive therapy; IQR, interquartile range; TB, tuberculosis; WHO, World Health Organization.

Table 2. Sputum Results From 207 Participants With HIV and Symptoms Suggestive of Tuberculosis

	Confirmed TB, No.	No TB, No.
Sputum smear positive		
Negative	66	131
Scanty	1	1
1+	0	0
2+	4	0
3+	4	0
Xpert MTB/RIF Ultra		
Negative	13	131
Trace	1	1[a]
Positive	61	0
Xpert MTB/RIF Ultra		
Rifampicin resistance	4	0
Positive sputum culture		
Visit 1	65	0
Visit 2	65	0
LPA resistance		
Rifampicin monoresistance	2	0
Isoniazid monoresistance	0	0
Multidrug resistance	1	0

Abbreviations: LPA, line probe assay; TB, tuberculosis.

[a]Patient was asymptomatic without antituberculosis treatment after 6 weeks.

sputum culture, sensitivity and specificity were 0.89 (95% CI, 0.79–0.96) and 0.97 (95% CI, 0.93–0.99). The equivalent figures for CRP ≥10 mg/L were 0.95 (95% CI, 0.87–0.98) and 0.26 (95% CI, 0.19–0.34). No improvement in symptoms after antibiotics had a sensitivity and specificity of 0.43 (95% CI, 0.32–0.55) and 0.86 (95% CI, 0.78–0.91), respectively (Table 3).

Multiple variables were associated with tuberculosis on simple logistic regression. The strongest association was with the nurse's opinion of likely tuberculosis (odds ratio [OR], 6.5; 95% CI, 3.11–13.6). Other strong associations were with no improvement in symptoms after antibiotics (OR, 4.53; 95% CI, 2.32–8.85) and CRP at visit 2 (OR, 1.11 per 10-unit increase; 95% CI, 1.06–1.16) (Table 4).

The final model using information available at visit 1 included ART status, number of symptoms, duration of symptoms, and temperature (Table 5). The model's c-statistic was 0.70; the equivalent in bootstrap validation was 0.68, with an optimism estimate of 0.042, indicating good stability of the model in internal validation (Figure 2). There was no evidence of poor calibration based on the GOF test (P = .63). However, visual inspection of the calibration plot shows underestimation of risk in the midrange of values (Figure 3).

The final model based on information available at visit 2 included change in symptoms after antibiotics, CRP at visit 2, number of symptoms, duration of symptoms, and ART status (Table 6). The model's c-statistic was 0.78; the equivalent in bootstrap validation was 0.76, with an optimism estimate of 0.046, indicating good stability of the model in internal validation (Figure 4). There was no evidence of poor calibration based on the GOF test (P = .07). Visual inspection of the calibration plot indicates good agreement between the rate of tuberculosis estimated by the model and the tuberculosis frequency observed in the study population (Figure 5).

We explored a model with data available at visit 2 without CRP, which included change in symptoms after antibiotics, number of symptoms, duration of symptoms, and ART status (Table 7). The model's c-statistic was 0.75; the equivalent in bootstrap validation was 0.73, with an optimism estimate of 0.037, indicating good stability of the model in internal validation (Figure 6). There was no evidence of poor calibration based on the GOF test (P = .60). Visual inspection of the calibration plot indicates reasonable agreement between the rate of tuberculosis estimated by the model and the tuberculosis frequency observed in the study population (Figure 7).

The XPHACTOR clinical prediction rule was not adequately validated in our cohort. Discrimination was poor (c-statistic, 0.65), although calibration was good (GOF test P = .96) (Figure 8).

DISCUSSION

Since 2007, WHO algorithms have discouraged the use of a ToA as a diagnostic strategy for ambulatory PWH suspected of having tuberculosis. Our study suggests that lack of clinical

Table 3. Diagnostic Accuracy of Sputum Tests and C-Reactive Protein in 207 Adult Participants With HIV and Symptoms Suggestive of Tuberculosis

Index Test	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Likelihood Ratio +ve (95% CI)	Likelihood Ratio –ve (95% CI)
Smear (including scanty)	0.12 (0.06–0.22)	0.99 (0.96–1.00)	0.90 (0.55–1.00)	0.66 (0.59–0.73)	15.84 (2.05– 122.6)	0.89 (0.81–0.97)
Xpert MTB/RIF Ultra (including trace)[a]	0.89 (0.79–0.96)	0.97 (0.93–0.99)	0.94 (0.84–0.98)	0.95 (0.90–0.98)	29.89 (11.3–78.8)	0.11 (0.06–0.22)
Xpert MTB/RIF Ultra (including trace)	0.83 (0.72–0.90)	0.99 (0.96–1.00)	0.98 (0.91–1.00)	0.91 (0.85–0.95)	109.12 (15.44–771.03)	0.17 (0.11–0.29)
CRP ≥10 mg/L visit 1	0.95 (0.87–0.98)	0.26 (0.19–0.34)	0.42 (0.34–0.49)	0.89 (0.75–0.97)	1.28 (1.14–1.43)	0.21 (0.08–0.57)
Symptoms not improved after antibiotics	0.43 (0.32–0.55)	0.86 (0.78–0.91)	0.63 (0.48– 0.76)	0.73 (0.65– 0.80)	3.00 (1.84– 4.91)	0.66 (0.54– 0.82)
CRP ≥10 mg/L after antibiotics	0.94 (0.86–0.98)	0.37 (0.28–0.46)	0.45 (0.37–0.53)	0.92 (0.81–0.98)	1.49 (1.29–1.72)	0.15 (0.06–0.40)

Reference standard of 2 sputum cultures unless stated.

Abbreviations: CI, confidence interval; CRP, C-reactive protein.

[a]Reference standard culture at visit 1 only (8 participants excluded due to contaminated cultures).

Table 4. Simple Logistic Regression Showing Factors Associated With Tuberculosis in 207 Adult Participants With HIV

	Unadjusted Odds Ratio	95% CI
Visit 1		
Age	0.97	0.93–1.01
Sex		
Female	Referent group	
Male	1.21	0.66–2.22
ART status		
<3 mo	Referent group	
>3 mo	0.40	0.20–0.79
Current CD4 count[a]	0.0998	0.0997–0.0999
No. of symptoms	1.80	1.25–2.58
Duration of symptoms		
<14 d	Referent group	
≥14 d	2.26	1.26–4.05
Smoking status		
Nonsmoker	Referent group	
Current smoker	1.21	0.65–2.25
BMI	0.90	0.83–0.98
Temperature, °C		
≤37.4	Referent group	
≥37.5	2.32	0.95–5.67
C-reactive protein[a]	1.04	1.00–1.07
Nurse opinion		
Unlikely TB	Referent group	
Likely TB	6.51	3.11–13.6
Visit 2		
Symptoms change		
Improved	Referent group	
Not improved	4.53	2.32–8.85
C-reactive protein[a]	1.11	1.06–1.16
C-reactive protein change (log scale)	0.46	0.29–0.73

CD4 count, number of symptoms, BMI, and C-reactive protein were modeled linearly.

Abbreviations: ART, antiretroviral therapy; BMI, body mass index; CI, confidence interval; TB, tuberculosis.

[a]Per 10-unit increase.

Table 5. Multiple Logistic Regression Model Predicting Tuberculosis Based on Data Available at Presentation in 207 Adults With HIV and a Positive Symptom Screen

	Adjusted Odds Ratio (95% CI)	Beta Coefficient, Log (Adjusted OR)
ART status		
<3 mo	1	0
>3 mo	0.49 (0.24–0.10)	–0.72
No. of symptoms	1.62 (1.09–2.41)	0.48
Duration of symptoms		
<14 d	1	0
≥14 d	2.03 (1.10–3.76)	0.71
Temperature, °C	1.61 (0.97–2.66)	0.48

Number of symptoms and temperature were modeled linearly. Risk score = –19.8 + 0.72 (if ART status >3 months) + (0.48 × No. of symptoms) + 0.71 (if cough >14 days) + (0.48 × temperature).

Abbreviations: ART, antiretroviral therapy; CI, confidence interval; OR, odds ratio.

response to antibiotics, as assessed by improvement in symptoms, is the strongest independent predictor of tuberculosis in patients seeking care for their symptoms (passive case-finding). The best clinical prediction model based on 2 visits (excluding CRP) included change in symptoms after antibiotics, number and duration of symptoms, and ART status. It showed acceptable discrimination (c-statistic, 0.75) with good calibration (GOF $P = .60$). A systematic review of 2586 patients found the sensitivity and specificity of ToA to be 0.72 and 0.77, respectively [6], and noted substantial heterogeneity due to inconsistent methodologies across studies. In our study, ToA had a lower sensitivity (0.43) but a higher specificity (0.86). This suggests that response to antibiotics may be an important part of diagnostic algorithms, particularly when access to Ultra is limited. Further information regarding the diagnostic accuracy and clinical impact of ToA as a diagnostic aid will come from an ongoing randomized controlled trial (https://clinicaltrials.gov/ct2/show/NCT03545373).

This is the first study to evaluate CRP as an objective marker of response to antibiotics in PWH who are suspected of having tuberculosis. Change in CRP and absolute CRP were both predictors of tuberculosis in simple logistic regression analysis, but only the absolute value after antibiotics was retained in the final predictive models. Including CRP in the final model based on 2 visits slightly improved discrimination (c-statistic 0.78) with good calibration (GOF $P = .07$).

CRP has been proposed as a useful initial test in the evaluation of PWH suspected of having tuberculosis. In our study, the sensitivity and specificity of CRP ≥10 were 0.95 (95% CI, 0.87–0.98) and 0.26 (95% CI, 0.19–0.34), respectively, which is compatible with the findings of a meta-analysis that showed a range of sensitivity and specificity of 0.96–0.97 and 0.33–0.73, respectively, in patients self-reporting symptoms compatible with tuberculosis [11]. CRP <10 mg/L is most likely to be useful as a "rule-out" test, allowing resources such as Ultra to be allocated to patients with a higher probability of disease. However, in our passive case-finding cohort, the negative predictive value (NPV) of 0.89 (95% CI, 0.75–0.97), with a prevalence of 36%, is unlikely to be high enough to confidently exclude disease. When analyzed as a continuous variable, CRP at presentation was a significant predictor of tuberculosis on simple logistic regression (OR, 1.04 per 10-unit increase; 95% CI, 1.00–1.07) but was not retained in a diagnostic model, also suggesting that it may not be useful in the initial evaluation of passive case-finding patients.

There have been 3 attempts to develop MPMs for adult patients presenting with symptoms of tuberculosis. The studies by Balcha et al. [9] and Rudolf et al. [10], from Ethiopia and Bissau, respectively, included Karnofsky status, mid–upper arm circumference (MUAC), peripheral lymphadenopathy, anemia, cough, dyspnea, chest pain, and BMI as predictors. Each showed reasonable discrimination, but their

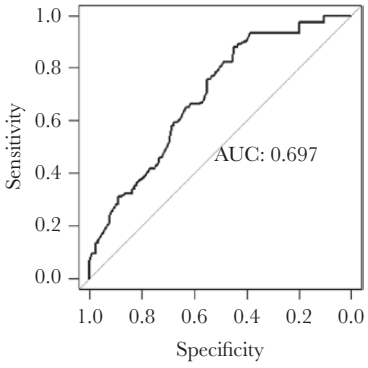

Figure 2. Receiver operating characteristics curve for a clinical prediction model for tuberculosis based on data available at index visit in 207 adults with HIV and a positive symptom screen.

Table 6. Multiple Logistic Regression Model Predicting Tuberculosis Based on Data Available at a Second Visit Following Antibiotics in 207 Adults With HIV and a Positive Symptom Screen

	Adjusted Odds Ratio (95% CI)	Beta Coefficient, Log (Adjusted OR)
Symptom change		
Improved	1	0
Not improved	3.24 (1.51–6.94)	1.17
CRP at visit 2[a]	1.07 (1.02–1.12)	0.07
No. of symptoms	1.46 (0.95–2.23)	0.38
Duration of symptoms		
<14 d	1	0
≥14 d	2.18 (1.12–4.23)	0.78
ART status		
<3 mo	1	0
>3 mo	0.50 (0.22–1.09)	–0.70

Number of symptoms was modeled linearly. Risk score = –2.81 + 1.17 (if symptoms not improved) + (0.007 × CRP at visit 2) + (0.38 × No. of symptoms) + 0.78 (if symptom duration ≥14 days) – 0.70 (if ART status >3 months).

Abbreviations: ART, antiretroviral therapy; CI, confidence interval; CRP, C-reactive protein; OR, odds ratio.

[a]Per 10-unit increase.

sample sizes were relatively small, and neither has been externally validated. The XPHACTOR rule is based on ART status, CD4 count, BMI, and number of WHO symptoms and shows acceptable discrimination (c-statistic, 0.75) and calibration (GOF $P = .31$). Our study was unable to validate the XPHACTOR rule, showing poor discrimination (AUC, 0.65). In keeping with standard validation procedures, our study population was similar but different. Both studies were carried out in the same city but differed with respect to the prevalence of tuberculosis, the number of symptoms, and whether presentation was self-directed. Our results suggest that different models may be necessary for active and passive case-finding approaches. We were also unable to develop a model with good discrimination and calibration based on data available at visit 1 only. The best model, which did not include CRP, showed only moderate discrimination (c-statistic, 0.70), with no evidence of poor calibration (GOF $P = .63$).

The sensitivity of Ultra is insufficiently high to exclude tuberculosis when negative, particularly in high-prevalence settings. Current WHO guidance when Xpert is negative is to investigate further with chest x-ray, clinical assessment, and a repeat Xpert MTB/RIF using a fresh specimen. The prevalence of culture-confirmed tuberculosis in participants with negative Ultra in our study was 9%, suggesting that further investigation is likely to be needed, although an algorithm for ruling out disease in a proportion of these patients would be useful to save resources.

Our study has a number of limitations. The sample was taken from a single clinic. All patients self-presented with symptoms suggestive of tuberculosis (passive case-finding), and the tuberculosis prevalence was high (36%); therefore, our findings may not be generalizable to active case-finding populations or where the prevalence of tuberculosis is lower. In addition, the WHO symptom screen is known to have imperfect sensitivity, so some patients with active tuberculosis will not be captured by this passive case-finding methodology. The time between index and return visit was relatively short (median,

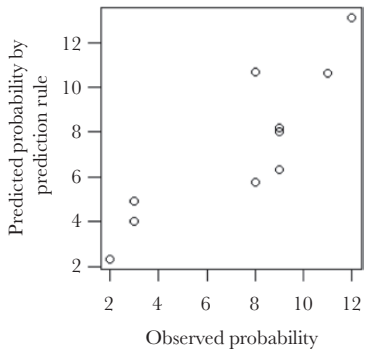

Figure 3. Calibration plot of the multivariate logistic regression model aimed at establishing a clinical prediction rule for the diagnosis of tuberculosis at index visit among 207 adults with HIV and a positive symptom screen.

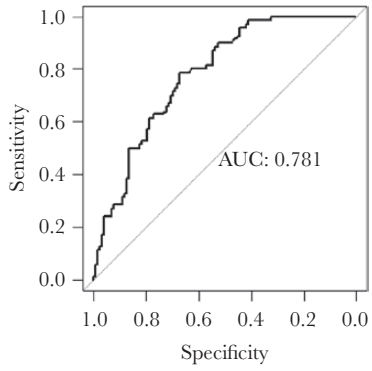

Figure 4. Receiver operating characteristics curve for a clinical prediction model for tuberculosis based on data available at a second visit following antibiotics index visit in 207 adults with HIV and a positive symptom screen.

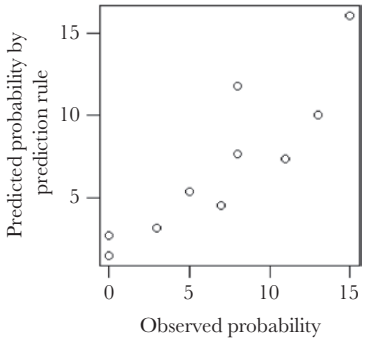

Figure 5. Calibration plot of the multivariate logistic regression model aimed at establishing a clinical prediction rule for the diagnosis of tuberculosis at a second visit following antibiotics among 207 adults with HIV and a positive symptom screen.

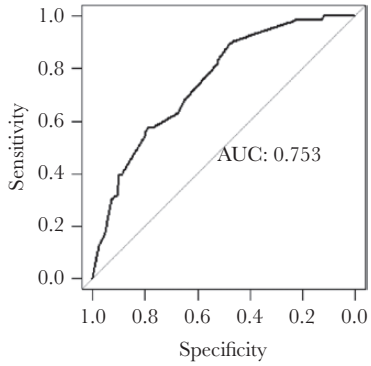

Figure 6. Receiver operating characteristics curve for a clinical prediction model for tuberculosis based on data available at a second visit following antibiotics index visit in 207 adults with HIV and a positive symptom screen (excluding C-reactive protein).

2 days) due to the SOPs of this urban clinic. It may not be practical to receive Ultra results and arrange return visits within 2 days in more rural clinics. The choice of antibiotic prescribed for each patient was made by independently of the study and not recorded. Therefore, a proportion of participants received azithromycin, which is known to have anti-inflammatory properties, which may have reduced CRP though a mechanism unrelated to treatment of pneumonia. A strength of the study was that we obtained 2 sputum mycobacterial cultures from each patient, ensuring a highly sensitive and specific reference standard.

Future work should be to validate our findings in different populations. In settings where access to Ultra is limited, symptomatic improvement to a ToA, possibly including repeat CRP, should be evaluated. Ultra could be targeted to patients with higher probability of tuberculosis at visit 2. Our study lacked sufficient power to determine predictors of tuberculosis when Ultra was negative (13 cases). Future work should investigate predictors of tuberculosis and rules for prioritizing further testing with repeat Ultra, chest x-ray, and culture in this group, particularly given that the prevalence of tuberculosis remains high (9%). It would also be important to test whether similar strategies apply to patients at index visit and whether they have received antibiotics before presentation, a population that was excluded from our study. Practical considerations for the implementation of ToA as a diagnostic strategy include the need for an uninterrupted supply of antibiotics and the possible increase in clinic workload. However, this may be minimized if patients are asked to return for a second visit only if their symptoms persist.

In conclusion, a clinical response to antibiotics is useful for diagnosing tuberculosis in PWH who are identified by passive case-finding and is a strategy that could be implemented immediately. CRP after antibiotics adds value but is not essential. CRP at index presentation is of limited value in this population.

Table 7. Multiple Logistic Regression Model Predicting Tuberculosis Based on Data Available at a Second Visit Following Antibiotics in 207 Adults With HIV and a Positive Symptom Screen (Excluding C-Reactive Protein)

	Adjusted Odds Ratio (95% CI)	Beta Coefficient, Log (Adjusted OR)
Symptom change		
Improved	1	0
Not improved	4.50 (2.19–9.23)	1.50
No. of symptoms	1.70 (1.12–2.57)	0.53
Duration of symptoms		
<14 d	1	0
≥14 d	2.09 (1.10–3.97)	0.74
ART status		
<3 mo	1	0
>3 mo	0.54 (0.25–1.14)	–0.62

Risk score = –2.90 + 1.50 (if symptoms not improved) + (0.53 × No. of symptoms) + 0.74 (if symptom duration ≥14 days) – 0.62 (if ART status >3 months).

Abbreviations: ART, antiretroviral therapy; CI, confidence interval; OR, odds ratio.

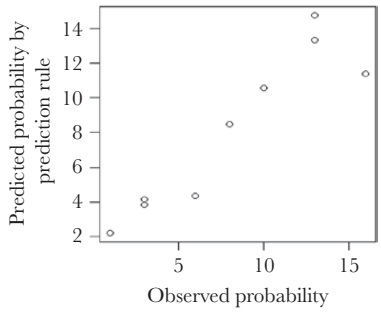

Figure 7. Calibration plot of the multivariate logistic regression model aimed at establishing a clinical prediction rule for the diagnosis of tuberculosis at a second visit following antibiotics among 207 adults with HIV and a positive symptom screen (excluding C-reactive protein).

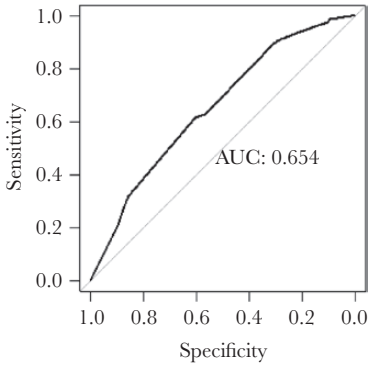

Figure 8. Receiver operating characteristics curve for validation of the XPHACTOR clinical prediction model for tuberculosis in 207 adults with HIV and a positive symptom screen.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Acknowledgments

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Alere (now Abbott) loaned the Afinion machine.

Ethical approval. All participants signed informed consent. The study was approved by the Human Research Ethics Committee of the University of the Witwatersrand, protocol number M180401.

Financial support. Research reported in this publication was supported by the Fogarty International Center and the National Institute of Mental Health, of the National Institutes of Health, under award number D43 TW010543.

Potential conflicts of interest. The authors declare no conflicts of interest.

References

1. World Health Organization. Treatment of Tuberculosis: Guidelines for National Programmes. Geneva: World Health Organization; **2003**.
2. World Health Organization. Improving the Diagnosis and Treatment of Smear-Negative Pulmonary and Extrapulmonary Tuberculosis Among Adults and Adolescents: Recommendations for HIV-Prevalent and Resource-Constrained Settings. Geneva: World Health Organization; **2007**.
3. Alamo ST, Kunutsor S, Walley J, et al. Performance of the new WHO diagnostic algorithm for smear-negative pulmonary tuberculosis in HIV prevalent settings: a multisite study in Uganda. Trop Med Int Health **2012**; 17:884–95.
4. Huerga H, Varaine F, Okwaro E, et al. Performance of the 2007 WHO algorithm to diagnose smear-negative pulmonary tuberculosis in a HIV prevalent setting. PLoS One **2012**; 7:e51336.
5. World Health Organization. Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection—Recommendations for a Public Health Approach. Geneva: World Health Organization; **2016**.
6. Divala TNM, Gupta-Wrigh A, et al. Is the current practice of using response to empirical broad-spectrum antibiotic treatment as an exclusion diagnostic for tuberculosis supported by evidence? Systematic review and meta-analysis. Paper presented at: 49th World Conference on Lung Health of the International Union Against Tuberculosis and Lung Disease; 24-27 October **2018**; The Hague, The Netherlands.
7. Hamada Y, Lujan J, Schenkel K, et al. Sensitivity and specificity of WHO's recommended four-symptom screening rule for tuberculosis in people living with HIV: a systematic review and meta-analysis. Lancet HIV **2018**; 5:e515–23.
8. Hanifa Y, Fielding KL, Chihota VN, et al. A clinical scoring system to prioritise investigation for tuberculosis among adults attending HIV clinics in South Africa. PLoS One **2017**; 12:e0181519.
9. Balcha TT, Skogmar S, Sturegård E, et al. A clinical scoring algorithm for determination of the risk of tuberculosis in HIV-infected adults: a cohort study performed at Ethiopian health centers. Open Forum Infect Dis **2014**; 1(X):XXX–XX.
10. Rudolf F, Haraldsdottir TL, Mendes MS, et al. Can tuberculosis case finding among health-care seeking adults be improved? Observations from Bissau. Int J Tuberc Lung Dis **2014**; 18:277–85.
11. Yoon C, Chaisson LH, Patel SM, et al. Diagnostic accuracy of C-reactive protein for active pulmonary tuberculosis: a meta-analysis. Int J Tuberc Lung Dis **2017**; 21:1013–9.
12. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). Ann Intern Med **2015**; 162:735–6.
13. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open **2016**; 6:e012799.
14. Takata S, Wada H, Tamura M, et al. Kinetics of C-reactive protein (CRP) and serum amyloid A protein (SAA) in patients with community-acquired pneumonia (CAP), as presented with biologic half-life times. Biomarkers **2011**; 16:530–5.
15. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol **1996**; 49:1373–9.
16. Hosmer DLS. Applied Logistic Regression. New York: Wiley; **2013**.
17. Hosmer DLS. *Applied Logistic Regression*. New York: Wiley; **1989**.
18. Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; **2001**.
19. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Diagnostic and prognostic value of serum C-reactive protein for screening for HIV-associated tuberculosis. Int J Tuberc Lung Dis **2013**; 17:636–43.