Impact of Targeted Testing for Latent Tuberculosis Infection Using Commercially Available Diagnostics

James D. Mancuso,[1,2] David Tribble,[2] Gerald H. Mazurek,[5] Yuanzhang Li,[1] Cara Olsen,[2] Naomi E. Aronson,[3] Lawrence Geiter,[4] Donald Goodwin,[6] and Lisa W. Keep[2]

[1]Division of Preventive Medicine, Walter Reed Army Institute of Research, Silver Spring, Maryland; Departments of [2]Preventive Medicine and Biometrics and [3]Department of Medicine, Uniformed Services University of the Health Sciences, Bethesda, Maryland; and [4]TB Products Unit, Otsuka Pharmaceuticals, Rockville, Maryland; [5]Division of TB Elimination, Centers for Disease Control and Prevention, Atlanta, Georgia; and [6]US Air Force School of Aerospace Medicine, Brooks City-Base, Texas

Background. The interferon-γ release assays (IGRAs) are increasingly being used as an alternative to the tuberculin skin test (TST). Although IGRAs may have better specificity and certain logistic advantages to the TST, their use may contribute to overtesting of low-prevalence populations if testing is not targeted. The objective of this study was to evaluate the accuracy of a risk factor questionnaire in predicting a positive test result for latent tuberculosis infection using the 3 commercially available diagnostics.

Methods. A cross-sectional comparison study was performed among recruits undergoing Army basic training at Fort Jackson, South Carolina, from April through June 2009. The tests performed included: (1) a risk factor questionnaire; (2) the QuantiFERON Gold In-Tube test (Cellestis Limited, Carnegie, Victoria, Australia); (3) the T-SPOT.TB test (Oxford Immunotec Limited, Abingdon, United Kingdom); and (4) the TST (Sanofi Pasteur Ltd., Toronto, Ontario, Canada). Prediction models used logistic regression to identify factors associated with positive test results. RFQ prediction models were developed independently for each test.

Results. Use of a 4-variable model resulted in 79% sensitivity, 92% specificity, and a c statistic of 0.871 in predicting a positive TST result. Targeted testing using these risk factors would reduce testing by >90%. Models predicting IGRA outcomes had similar specificities as the skin test but had lower sensitivities and c statistics.

Conclusions. As with the TST, testing with IGRAs will result in false-positive results if the IGRAs are used in low-prevalence populations. Regardless of the test used, targeted testing is critical in reducing unnecessary testing and treatment.

Clinical Trial Registration. NCT00804713.

Universal screening for latent tuberculosis infection (LTBI) in the United States is no longer recommended; current practice favors a targeted approach. Centers for Disease Control and Prevention (CDC) guidelines recommend targeted testing of only persons with known

risk factors for TB, specifically stating that "targeted tuberculin testing programs should be conducted only among groups at high risk and discouraged in those at low risk" [1 p. 1]. Similarly, the Institute of Medicine has called for the development of targeted TB screening programs based on epidemiologic risk analysis [2]. Targeted testing offers logistic and efficiency advantages over universal screening and increases the positive predictive value (PPV) of a positive test result by selecting a higher-prevalence population for testing [3]. In relatively homogeneous exposure situations, such as among immigrants, in prisons, and in hospitals, universal testing is still performed on the basis of the association with a high-risk setting, although health care workers are increasingly being recognized as having heterogeneous

exposures [4]. Targeted testing has been implemented and evaluated using predictive models in several such heterogeneous contexts, including in contact investigations [5, 6], among pediatric populations [7, 8], and among university entrants [9]. These have all found that targeted testing may dramatically reduce the amount of testing without negative effects on disease control efforts.

The US military has performed universal testing of recruits entering military service since the 1960s [10]. However, the US military has a low risk for active TB, with a rate of 0.65 cases of confirmed pulmonary TB per 100 000 person-years from 1998 to 2007, a rate 84% lower than the age-adjusted US rate (J. D. Mancuso; Walter Reed Army Institute of Research; unpublished data). Because it is a population with heterogeneous exposures to TB prior to accession (ie, entry into military service), the US military is challenged with mitigating the risk of TB among higher-risk recruits without exposing low-risk recruits to unnecessary therapy. Targeted testing relies on identifiable risk factors and an assessment tool that can accurately predict LTBI. Interferon-γ release assays (IGRAs), including the QuantiFERON® Gold In-Tube test (QFT) (Cellestis Limited, Carnegie, Victoria, Australia) and the T-SPOT®.TB test (T-Spot) (Oxford Immunotec Limited, Abingdon, United Kingdom), are thought to be more specific and to have other logistic advantages over the tuberculin skin test (TST) (Tubersol®, Sanofi Pasteur Ltd., Toronto, Ontario, Canada) [11]. The purpose of this study is to provide an evaluation of the accuracy of a risk factor questionnaire (RFQ) in predicting positive TST, QFT, and T-Spot results. This is used to compare the effectiveness of targeted testing strategies using the 3 commercially available diagnostic tests for LTBI.

METHODS

Study Procedures

The study was approved in 2009 by the Uniformed Service University Institutional Review Board and was conducted in the same year. Study procedures included obtaining informed consent of all subjects. This cross-sectional comparison study among Army recruits at Fort Jackson, South Carolina, consisted of (1) the RFQ, (2) QFT, (3) the T-Spot, and (4) the TST. A total of 2017 subjects were enrolled in the study from 1 April to 11 June 2009.

The RFQ was developed from previous literature and validation studies [5, 6, 8–10, 12]. The RFQ was performed before all other TB testing, and subjects were encouraged to complete all fields. The RFQ took 3–5 minutes for participants to complete. The primary variables of interest included foreign birth, race and ethnic background, and contact with a case of active TB [10]. Other factors included demographic characteristics, foreign residence or travel, exposure assessment, bacille Calmette-

Guérin (BCG) vaccination, history of prior TB diagnosis or treatment, and prior positive skin test result, as shown in Table 1. Human immunodeficiency cirus (HIV) and other immuno-suppressive conditions are disqualifying for entry into military service and therefore are not reported here. The TB prevalence reported by the World Health Organization in 1990 was used to estimate exposure risk in country of birth and during overseas travel or residence using groups of (1) <20 cases per 100 000 persons, (2) 20–100 cases per 100 000 persons, and (3) >100 cases per 100 000 persons [13, 14].

Blood samples were obtained for QFT and T-Spot at the time of routine phlebotomy for recruit inprocessing. Personnel performing IGRA assays were blinded to all patient data. QFT was performed in accordance with package insert instructions, including incubation and centrifugation within the prescribed times [15]. Testing was completed with the aid of a Triturus automated enzyme-linked immunosorbent assay workstation (Grifols USA). Blood samples were obtained for T-Spot into sodium heparin tubes and shipped overnight at room temperature to the Oxford Immunotec laboratory (Marlborough, MA). T-Spot was performed in accordance with the package insert instructions [16], except for the addition of 25 μl/mL T cell Xtend (Oxford Immunotec) immediately before peripheral blood mononuclear cell recovery, to increase the processing window from 8 hours up to 32 hours.

All personnel involved in placement and reading of the skin test findings were trained and monitored to strictly adhere to standard operating procedures on the basis of the published methods [17, 18]. The Mantoux technique was used to administer 0.1 mL (5 TU) of Tubersol tuberculin PPD (Sanofi Pasteur). The transverse diameter of induration at each skin test site was measured 2 days after administration.

Statistical and Data Analysis

SAS software, version 9.2 (SAS Institute), was used for all analyses. A positive TST result was defined using risk-stratified interpretation (RSI) criteria from published CDC guidelines [1]. IGRA end points were defined by using established cutoffs from the manufacturer [15, 16]. The RFQ was used to develop predictive models for positive responses to TST and IGRA [5–9]. RFQ prediction models were developed independently for each of the three diagnostic tests. Factors associated with a positive result were evaluated using the Pearson χ^2 test, the Fisher exact test, and unconditional logistic regression. An alpha level of .05 was used to identify significance in all statistical tests.

Sensitivity, specificity, PPV, and negative predictive value (NPV) were calculated for each RFQ variable response and for different combinations of variables. Missing data were rare (generally <1%), so no imputation techniques were necessary. Receiver operator curves were constructed by plotting sensitivity versus 1 - specificity for each probability level. Variables were

Table 1. Association of Selected Factors With Positive Tuberculin Skin Test (TST), QuantiFERON Gold In-Tube (QFT), and T-SPOT.TB (T-Spot) Among US Army Recruits at Entry Into Military Service

Factor	No. of subjects	TST		QFT		T-Spot	
		No. (%) of persons with positive results	OR (95% CI)	No. (%) of persons with positive results	OR (95% CI)	No. (%) of persons with positive results	OR (95% CI)
Age, years							
18–24	1438	20 (1.4)	1 (Ref)	28 (2.0)	1 (Ref)	22 (1.5)	1 (Ref)
25–29	193	4 (2.1)	1.5 (0.4–4.5)	2 (1.0)	0.6 (0.06–2.1)	4 (2.1)	1.4 (0.3–4.1)
≥30	149	14 (9.4)	7.4 (3.2–15.7)	6 (4.0)	2.1 (0.7–5.3)	8 (5.4)	3.7 (1.4–8.7)
Sex							
Male	1160	23 (2.0)	1.0 (Ref)	23 (2.0)	1 (Ref)	27 (2.3)	1 (Ref)
Female	621	15 (2.4)	1.2 (0.6–2.4)	13 (2.1)	1.1 (0.5–2.1)	7 (1.1)	0.5 (0.2–1.1)
Race/ethnic group[a]							
White	1175	8 (0.7)	1 (Ref)	15 (1.3)	1 (Ref)	16 (1.4)	1 (Ref)
Black	408	12 (2.9)	4.4 (1.6–12.5)	10 (2.5)	1.2 (0.8–4.7)	7 (1.7)	1.3 (0.5–3.5)
Hispanic	203	7 (3.5)	5.2 (1.6–16.6)	10 (4.9)	4.0 (1.6–9.7)	7 (3.5)	2.6 (0.9–6.8)
Asian/Pacific Islander	106	13 (12.3)	20.4 (7.6–57.9)	5 (4.7)	3.8 (1.1–11.4)	6 (5.7)	4.3 (1.4–12.0)
Other	137	6 (4.4)	6.7 (1.9–22.3)	7 (5.1)	4.2 (1.4–11.1)	5 (3.7)	2.7 (0.8–8.0)
Prevalence of TB in country of birth							
<20 cases per 100 000 persons	1687	19 (1.1)	1 (Ref)	27 (1.6)	1 (Ref)	24 (1.4)	1 (Ref)
20–100 cases per 100 000 persons	33	4 (12.1)	12.1 (2.8–39.5)	3 (9.1)	6.1 (1.1–21.7)	3 (9.1)	6.9 (1.3–24.7)
>100 cases per 100 000 persons	63	15 (23.8)	27.4 (12.1–60.6)	6 (9.5)	6.5 (2.1–16.8)	7 (11.1)	8.7 (3.0–21.8)
Lived with family member who was not born in US							
No	1701	25 (1.5)	1 (Ref)	26 (1.5)	1 (Ref)	27 (1.6)	1 (Ref)
Yes	79	13 (16.5)	13.2 (6.5–27.0)	10 (12.7)	9.3 (4.3–20.1)	7 (8.9)	6.0 (2.1–14.8)
Prevalence of TB in country the subject lived in or traveled to for >1 mo							
<20 cases per 100 000 persons	1717	31 (1.8)	1 (Ref)	33 (1.9)	1 (Ref)	31 (1.8)	1 (Ref)
20–100 cases per 100 000 persons	25	1 (4.0)	2.3 (0.1–14.8)	0 (0)	0 (0–7.9)	0 (0)	0 (0–8.5)
>100 cases per 100 000 persons	41	6 (14.6)	9.3 (3.0–24.6)	3 (7.3)	4.0 (0.8–13.7)	3 (7.3)	4.3 (0.8–14.7)
Contact with TB case							
None	1697	30 (1.8)	1 (Ref)	27 (1.6)	1 (Ref)	31 (1.8)	1 (Ref)
Casual	65	5 (7.7)	4.6 (1.4–12.6)	6 (9.2)	6.3 (2.0–16.3)	2 (3.1)	1.7 (0.2–7.0)
In same household	20	3 (15.0)	9.8 (1.7–36.5)	3 (15.0)	10.9 (1.9–40.9)	1 (5.0)	2.8 (0.1–18.9)
Health care work							
No	1573	9 (4.3)	1 (Ref)	32 (2.0)	1 (Ref)	30 (1.9)	1 (Ref)
Yes	209	29 (1.8)	2.4 (1.1–5.1)	4 (1.9)	0.9 (0.2–2.7)	4 (1.9)	1.0 (0.4–2.9)
Lived or worked in congregate setting							
No	1671	33 (2.0)	1 (Ref)	31 (1.9)	1 (Ref)	30 (1.8)	1 (Ref)
Yes	112	5 (4.5)	2.3 (0.9–6.1)	5 (4.5)	2.5 (0.9–6.5)	4 (3.6)	2.0 (0.7–5.9)
Bacille Calmette-Guérin vaccination							
No	1722	22 (1.3)	1 (Ref)	30 (1.7)	1 (Ref)	25 (1.5)	1 (Ref)
Yes	61	16 (26.2)	27.5 (13.5–55.8)	6 (9.8)	6.2 (2.5–15.4)	9 (14.8)	11.7 (5.2–26.4)
Treated for TB							
No	1739	33 (1.9)	1 (Ref)	34 (2.0)	1 (Ref)	31 (1.8)	1 (Ref)
Yes	42	5 (11.9)	7.0 (2.6–18.9)	2 (4.8)	2.5 (0.3–10.4)	3 (7.1)	4.2 (1.2–14.5)
Prior positive TST result							
No	1762	26 (1.5)	1 (Ref)	34 (1.9)	1 (Ref)	29 (1.7)	1 (Ref)
Yes	21	12 (57.1)	89.0 (34.5–229)	34 (1.9)	5.3 (0.6–23.6)	5 (23.8)	18.7 (6.4–54.4)

Table 1. (continued)

Factor	No. of subjects	TST No. (%) of persons with positive results	TST OR (95% CI)	QFT No. (%) of persons with positive results	QFT OR (95% CI)	T-Spot No. (%) of persons with positive results	T-Spot OR (95% CI)
Region of US							
Northeast	303	8 (2.6)	1 (Ref)	7 (2.3)	1 (Ref)	5 (1.7)	1 (Ref)
Southeast	601	9 (1.5)	0.6 (0.2–1.7)	9 (1.5)	0.6 (0.2–2.1)	10 (1.7)	1.0 (0.3–3.8)
West	627	11 (1.8)	0.7 (0.2–1.9)	13 (2.1)	0.9 (0.3–2.7)	12 (1.9)	1.2 (0.4–4.3)
Other	252	10 (4.0)	1.5 (0.5–4.5)	7 (2.8)	1.2 (0.4–4.1)	7 (2.8)	1.7 (0.5–6.9)
Education							
<12 years	234	0 (0)	0 (0–1.0)	3 (1.3)	0.6 (0.1–1.9)	6 (2.6)	1 (Ref)
12 years	989	16 (1.6)	1 (Ref)	22 (2.2)	1 (Ref)	16 (1.6)	0.6 (0.2–2.0)
13-15 years	421	13 (3.6)	1.9 (0.8–4.3)	10 (2.4)	1.1 (0.4–2.4)	7 (1.7)	0.6 (0.2–2.3)
≥16 years	138	7 (5.1)	3.2 (1.1–8.5)	1 (0.7)	0.3 (0.01–2.0)	5 (3.6)	1.4 (0.3–5.7)
Smoking							
None	1322	34 (2.6)	1 (Ref)	24 (1.8)	1 (Ref)	30 (2.3)	1 (Ref)
<1 pack per day	365	3 (0.8)	0.3 (0.1–1.0)	10 (2.7)	1.5 (0.6–3.3)	4 (1.1)	0.5 (0.1–1.4)
≥1 pack per day	90	1 (1.1)	0.4 (0.01–2.6)	2 (2.2)	1.2 (0.1–5.1)	0 (0)	0 (0–1.8)

NOTE. CI, confidence interval; OR, prevalence odds ratio; Ref, referent; TB, tuberculosis; TST, Tuberculin Skin Test; QFT, Quantiferon® Gold in-tube; T-Spot, T-SPOT®.TB.

[a] Note: Soldiers were allowed to choose >1 race/ethnic group.

selected for inclusion into the models on the basis of prior knowledge of risk factors for TB exposure and contribution to the predictive ability of the model. The contribution to the model was assessed by change in area under the curve (AUC), or c statistic, when adding each predictor variable to the model. To validate the prediction model, the analysis was performed on a second set of samples obtained by bootstrap methods [19, 20]. For this analysis, 1000 bootstrap samples of the same size as the original model were taken from the original data set with replacement. The AUC and estimates of the odds ratios were reported for the bootstrap validation data set estimates and compared with the original data set.

RESULTS

Of the 3095 recruits approached from 1 April to 11 June 2009, 2697 were eligible to participate in the study, and 2017 subjects (75%) enrolled (Figure 1). Thirty-eight recruits withdrew before blood collection or completion of skin testing; 30 of these withdrawals were for administrative reasons unrelated to the study. TST results were available for 1978 (99.9%) of the remaining 1979 participants, and T-Spot and QFT results were available for 1888 (95.4%) and 1835 (92.7%), respectively. For comparability between the prediction models, this analysis was limited to subjects who had positive or negative results for all 3 tests (*n* = 1783) and excluded subjects with an indeterminate or invalid result by any test.

Characteristics of study participants are shown in Table 1 and were similar to the overall recruit population. TST induration

was detected in 105 participants (5.9%) and ranged from 2 to 80 mm. TST induration size of ≥10 mm was seen in 58 subjects (3.3%), but only 38 were positive by RSI criteria (2.1%). Of note, 15 of the 38 (39%) did not have any of the traditional risks used to stratify TST interpretation as defined by CDC, but had induration of 15 mm or greater. Similar proportions of positive results were seen for the IGRAs, with 34 positive T-Spot results (1.9%) and 36 positive QFT results (2.0%).

The unadjusted association of demographic and exposure risk factors with test results is shown in Table 1. Birth in a TB-endemic country had a particularly strong association with a positive test result, as did age, race, contact with a TB case, and other established risk factors [6, 8, 9, 21, 22]. The multivariate associations of selected factors with positive TST or IGRA results are shown in Table 2. After adjusting for the other variables in the model, significant associations were found between a positive test result and exposure to a TB case, TB prevalence in the country of birth, residence with a family member btorn outside of the United States, positive prior TST result, and residence in a congregate setting, such as homeless shelter, prison, or drug treatment facility. To account for possible overfitting, validation of the models with bootstrap estimates was performed. The estimates obtained via bootstrap were similar to those obtained in the original data set, although some predictors were no longer statistically significant, and the AUCs were moderately lower.

The characteristics of the RFQ models in predicting positive tests are shown in Tables 3, 4, and 5. As expected, the sensitivity, NPV, and AUC of the RFQ were seen to improve with increasing numbers of predictors, with corresponding decreases in RFQ
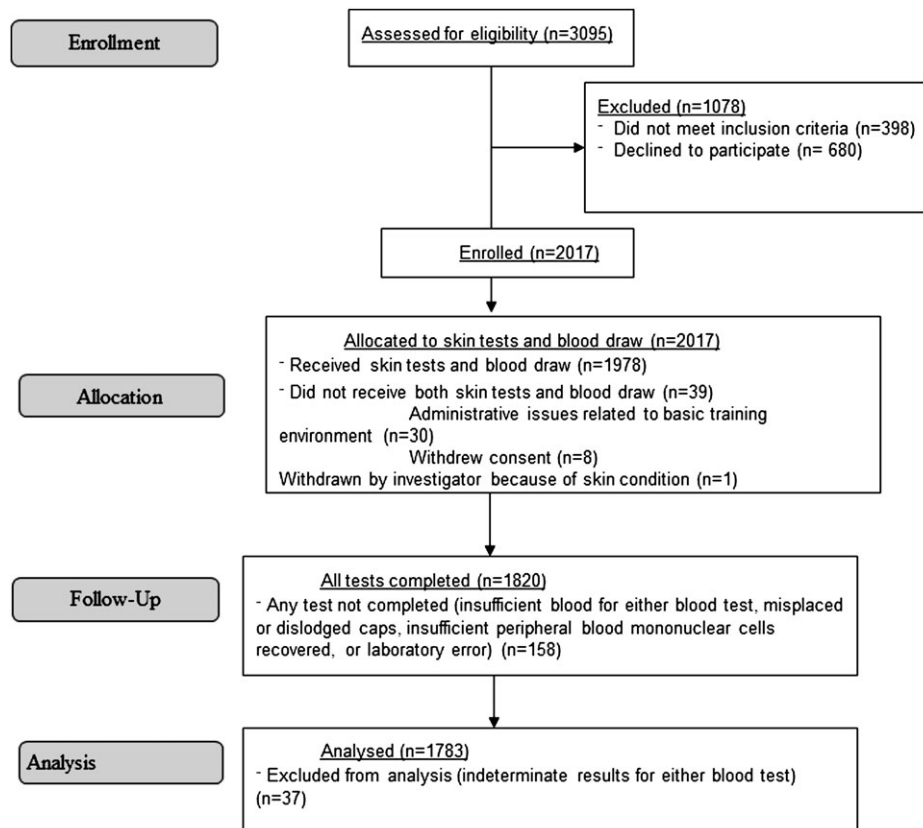
Figure 1. Flow chart of study comparing the tuberculin skin test with 2 interferon-γ release assays in 2009 at Fort Jackson, South Carolina.

specificity and PPV. A 4-variable model for RFQ prediction of a positive TST result was selected as having the best bias-variance tradeoff, with a sensitivity of 79% (95% confidence interval [CI], 63%–90%), specificity of 92% (95% CI, 91%–93%), and AUC of 0.871. Because only 9.3% of subjects had a positive response to 1 of these 4 variables, targeted testing of only these positive results would be expected to reduce testing by 90.7% (95% CI, 89%–92%). In contrast, when all potential risk factors were included, 32.5% had at least 1 "positive" response, but this increased the sensitivity only slightly while dramatically lowering specificity.

Figure 2 compares the receiver operator curves for the performance of the full RFQ model for predicting positive results for the TST, QFT, and T-Spot. This graphically demonstrates the lower performance characteristics of the RFQ in predicting an IGRA outcome, compared with the TST. Tables 4 and 5 show the characteristics of the independently created prediction models using the same combinations of variables for the QFT and T-Spot as used in Table 3 for the TST. Although the specificities and NPVs were similar, the sensitivity, PPV, and AUC of the RFQ were all considerably lower in predicting a positive IGRA than a positive TST. For the 4-variable model, the RFQ had a sensitivity of 44% (95% CI, 28%–62%), specificity of 91% (95% CI, 90%–93%), and AUC of 0.684 in predicting a positive QFT. The RFQ had very similar estimates in predicting

a positive T-Spot, with a sensitivity of 44% (95% CI, 27%–62%), specificity of 91% (95% CI, 90%–93%), and AUC of 0.688.

DISCUSSION

Risk factors for LTBI among US Army recruits were similar whether measured by the TST or 1 of the 2 commercially available IGRAs. RFQ prediction models were constructed using variables including birth in a country with a high prevalence of TB, close contact with an active TB case, history of living with a family member born outside the United States, and history of a prior positive TB skin test result. Use of these 4 variables resulted in 79% sensitivity, 92% specificity, and an AUC of 0.871 in predicting a positive TB skin test. Targeted testing of only those with a positive response to 1 of these 4 questions would reduce testing by >90%, increasing the efficiency of the testing program. Prediction models for the IGRAs had similar specificities and reductions in testing but had lower sensitivities and AUCs.

This is the first study to compare the effectiveness of targeted testing using either IGRA as an end point in any population, as well as the first to compare targeted testing as a predictive tool using IGRA and TST criterion standards. As in previous studies that used TST result as the outcome, birth in a TB-endemic country was found to be a strong predictor of

Table 2.   Association of Selected Factors With Positive Tuberculin Skin Test (TST), QuantiFERON Gold In-Tube (QFT), and T-SPOT.TB (T-Spot) Among US Army Recruits

Factor	TST		QFT		T-Spot	
	Adjusted OR (95% CI)	Bootstrap adjusted OR (95% CI)	Adjusted OR (95% CI)	Bootstrap adjusted OR (95% CI)	Adjusted OR (95% CI)	Bootstrap adjusted OR (95% CI)
Close contact of active TB case	2.8 (0.4–18.0)	1.4 (<0.01 to 44.7)	4.1 (0.96–17.6)	2.4 (<0.01 to 70.1)	0.7 (0.06–7.8)	<0.01 (<0.01 to <0.01)
Prevalence of TB in country of birth						
<20 cases per 100 000 persons	1 (Ref)	1 (Ref)	1 (Ref)	1 (Ref)	1 (Ref)	1 (Ref)
20–100 cases per 100 000 persons	4.3 (1.0–18.2)	2.4 (<0.01 to 31.2)	3.5 (0.93–13.3)	1.5 (<0.01 to 20.5)	3.5 (0.9–13.8)	1.3 (<0.01 to 20.5)
>100 cases per 100 000 persons	12.6 (4.8–33.0)	13.7 (2.2–101)	3.0 (0.99–8.8)	2.7 (0.3–26.0)	3.8 (1.3–10.9)	3.3 (0.4–23.6)
Lived with family member not born in the US	4.8 (1.8–13.1)	5.2 (0.67–37.3)	5.7 (2.2–14.8)	5.6 (0.7–31.5)	3.3 (1.2–9.6)	3.1 (0.3–22.9)
Prior TST positive	46.2 (15.1–142)	125.2 (10.4 to >999)	2.2 (0.4–11.7)	0.2 (<0.01 to 13.9)	9.5 (2.9–31.3)	9.6 (1.04–70.1)
Lived in shelter or congregate setting	4.4 (1.5–13.0)	2.91 (0.23–20.9)	3.0 (1.1–8.3)	2.5 (0.3–13.9)	2.4 (0.8–7.2)	1.4 (0.1–10.1)
Age ≥30 years	3.5 (1.5–8.4)	3.42 (0.86–13.2)	1.5 (0.6–4.0)	1.2 (0.2–5.5)	2.1 (0.9–5.2)	1.8 (0.3–7.0)
Traveled ≥1 month to a country with a TB prevalence of						
<20 cases per 100 000 persons	1 (Ref)	1 (Ref)	1 (Ref)	1 (Ref)	1 (Ref)	1 (Ref)
20-100 cases per 100 000 persons	0.3 (0.02–4.3)	<0.01 (<0.01 to 2.72)	<0.01 (<0.01 to >999)	<0.01 (<0.01 to <0.01)	<0.01 (<0.01 to >999)	<0.01 (<0.01 to <0.01)
>100 cases per 100 000 persons	2.2 (0.6–8.0)	2.0 (0.12–15.6)	1.0 (0.2–4.3)	0.4 (<0.01 to 5.3)	1.5 (0.3–6.6)	0.5 (<0.01 to 8.7)
AUC (c-statistic)	0.88	0.84 (0.76–0.91)	0.72	0.69 (0.62–0.78)	0.74	0.70 (0.61–0.79)

NOTE.   Variables were selected for inclusion based on contribution to the predictive model as measured by change in area under the curve. CI, confidence interval; OR, prevalence odds ratio; Ref, referent; TB, tuberculosis; TST, tuberculin skin test; QFT, Quantiferon® Gold in-tube ; T-Spot, T-SPOT®.TB.

Table 3. Validity of Predictors of Latent Tuberculosis Infection, as Measured by the Tuberculin Skin Test (TST), Among US Army Recruits

Factor	No. (%) of positive RFQ responses (n = 1783)	No. of persons with positive TST result (n = 38)	RFQ sensitivity, %	RFQ specificity, %	RFQ PPV, %	RFQ NPV, %	RFQ AUC, %
1. Close contact of TB case	20 (1.1)	3	7.9	99.0	15.0	98.0	0.535
2. TB prevalence ≥20 cases per 100 000 persons in country of birth	96 (5.4)	19	50.0	95.6	19.8	98.9	0.730
3. Lived with parent who was born outside of the US	79 (4.4)	13	34.2	96.2	16.5	98.5	0.652
4. Prior positive TST result	21 (1.2)	12	31.6	99.5	57.1	98.5	0.655
5. Age ≥30 years	149 (8.4)	14	36.8	92.3	9.4	98.5	0.646
Combinations of factors							
1 or 2	113 (6.3)	21	55.3	94.7	18.6	99.0	0.753
1, 2, or 3	124 (7.0)	27	71.1	94.4	21.8	99.3	0.798
1, 2, 3, or 4 (selected model)	**166 (9.3)**	**30**	**79.0**	**92.2**	**18.1**	**99.5**	**0.871**
1, 2, 3, 4, or 5	282 (15.8)	31	81.6	85.6	11.0	99.5	0.876
1, 2, 3, 4, 5, or overseas travel ≥1 month	318 (17.8)	31	81.6	83.6	9.8	99.5	0.877
1, 2, 3, 4, 5, or residence in congregate setting	367 (20.6)	31	81.6	80.7	8.5	99.5	0.876
Any risk factor[a]	579 (32.5)	32	84.2	68.7	5.5	99.5	0.878

NOTE. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value; RFQ, risk factor questionnaire; TB, tuberculosis; TST, tuberculin skin test.

[a] Risk factors include all of the previously mentioned risk factors, plus health care work, casual contact with a TB case, and prior TB treatment.

a positive test result [7–9, 22]. Close contact with a TB case, foreign-born family members, and prior positive TST have also been associated with LTBI in previous studies [7–9, 23]. Other variables have also sometimes been associated with a positive TST result, including travel [23], smoking [6], male sex [5, 9], health care work [4], and education [8], but these were not found to be important predictors of LTBI in this study. Race and ethnic group did not contribute meaningfully as predictors after adjusting for other factors. The only study to assess use of a questionnaire to target testing in a similar heterogeneous adult population was among college students in Virginia [9]. That study showed that using only the 2 variables of foreign birth and close contact with a patient with TB resulted in a sensitivity of 81.6% and a specificity of 91%. Although our 2-variable model had lower sensitivity than this, we found comparable sensitivity and specificity using a 4-variable model. Two studies in pediatric populations also found that using 4 or 5 questions to identify high-risk patients with LTBI resulted in similar sensitivities and specificities as those seen in this study [7, 8]. Prediction models of LTBI among contacts of active TB cases have had more modest reductions in testing, because of a higher pretest probability of infection and less concern about false-positive results than about false-negative results [5, 6].

This study has several important strengths. Despite other differences (such as age), the population was a good geographic representation of the underlying low prevalence, heterogeneous US source population (data not shown). Also, the 3 forms of TB testing allowed direct comparisons of the effectiveness of targeted testing to predict LTBI as measured by each test, which has not been done previously. There are also several limitations to this study; the most important is the lack of a gold standard in evaluating the presence of LTBI. The potential for false-positive TST results due to receipt of BCG, cross-reactivity to nontuberculous mycobacteria, and other factors is well known [24]. The IGRAs are also known to have limitations in sensitivity and specificity [11], and it is uncertain whether the predictive capability of the IGRAs is better than that of the TST. The small number of positive test results also may have led to less power to detect small differences in the groups studied. Misclassification of exposures and outcomes was also possible in association with measurement error, although the outcomes were probably better controlled in this study than they would be in practice. Finally, this study is not expected to be generalizable to higher risk populations, including those with HIV infection or other immunosuppressive conditions, hospital workers, or prison guards.

An important implication of this study is that targeted testing of heterogeneous populations is feasible and effective using the TST or either IGRA. Validation of targeted testing had previously only been performed in a few populations, and it had only been done using the TST. In this study, targeted testing was seen to be less predictive of commercially available IGRAs than

Table 4. Validity of Predictors of Latent Tuberculosis Infection, as Measured by the QuantiFERON Gold In-Tube (QFT), Among US Army Recruits

Factor	No. (%) of positive RFQ responses (n = 1783)	No. of persons with positive QFT results (n = 36)	RFQ sensitivity, %	RFQ specificity, %	RFQ PPV, %	RFQ NPV, %	RFQ AUC
1. Close contact of TB case	20 (1.1)	3	8.3	99.0	15.0	98.1	0.537
2. TB prevalence ≥20 cases per 100 000 persons in country of birth	96 (5.4)	9	25.0	95.0	9.4	98.4	0.600
3. Lived with parent born outside of the US	79 (4.4)	10	27.8	96.0	12.7	98.5	0.619
4. Prior positive TST result	21 (1.2)	2	5.6	98.9	9.5	98.1	0.522
5. Age ≥30 years	149 (8.4)	6	16.7	91.8	4.0	98.2	0.543
Combinations of factors							
1 or 2	113 (6.3)	12	33.3	94.2	10.6	98.6	0.638
1, 2, or 3	124 (7.0)	12	33.3	93.6	9.7	98.6	0.685
1, 2, 3, or 4 (selected model)	**166 (9.3)**	**16**	**44.4**	**91.4**	**9.6**	**98.8**	**0.684**
1, 2, 3, 4, or 5	282 (15.8)	18	50.0	84.9	6.4	98.8	0.691
1, 2, 3, 4, 5, or overseas travel ≥1 month	318 (17.8)	18	50.0	82.8	5.7	98.8	0.697
1, 2, 3, 4, 5, or residence in congregate setting	367 (20.6)	21	58.3	80.2	5.7	98.9	0.719
Any risk factor[a]	579 (32.5)	24	66.7	68.2	4.2	99.0	0.718

NOTE. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value; RFQ, risk factor questionnaire; TB, tuberculosis; QFT, Quantiferon® Gold in-tube.

[a] Risk factors include all of the previously mentioned risk factors, plus health care work, casual contact with a TB case, and prior TB treatment.

the TST but still had effectiveness comparable to universal testing. Although the IGRAs may be more specific tests, their use in low-prevalence populations will still result in predominantly false-positive results if testing is not targeted. Therefore, all testing of low-prevalence populations should be targeted, regardless of the choice of the diagnostic test used.

Table 5. Validity of Predictors of Latent Tuberculosis Infection, as Measured by the T-SPOT.TB (T-Spot), Among US Army Recruits

Factor	No. (%) of positive RFQ responses (n = 1783)	No. of persons with positive T-Spot results (n = 34)	RFQ sensitivity, %	RFQ specificity, %	RFQ PPV, %	RFQ NPV, %	RFQ AUC
1. Close contact of TB case	20 (1.1)	3	2.9	98.9	5.0	98.1	0.509
2. TB prevalence ≥20 cases per 100 000 persons in country of birth	96 (5.4)	10	29.4	95.1	10.4	98.6	0.623
3. Lived with parent born outside of the US	79 (4.4)	7	20.6	95.9	8.9	98.4	0.582
4. Prior positive TST result	21 (1.2)	5	14.7	99.1	23.8	98.4	0.569
5. Age ≥30 years	149 (8.4)	8	23.5	91.9	5.4	98.4	0.577
Combinations of factors							
1 or 2	113 (6.3)	11	32.4	94.2	9.7	98.6	0.633
1, 2, or 3	124 (7.0)	12	35.3	93.6	9.7	98.7	0.667
1, 2, 3, or 4 (selected model)	**166 (9.3)**	**15**	**44.1**	**91.4**	**9.0**	**98.8**	**0.688**
1, 2, 3, 4, or 5	282 (15.8)	18	52.9	84.9	6.4	98.9	0.708
1, 2, 3, 4, 5, or overseas travel ≥1 month	318 (17.8)	18	52.9	82.9	5.7	98.9	0.712
1, 2, 3, 4, 5, or residence in congregate setting	367 (20.6)	19	55.9	80.1	5.2	98.9	0.711
Any risk factor[a]	579 (32.5)	21	61.8	68.1	3.6	98.9	0.744

NOTE. AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value; RFQ, risk factor questionnaire; TB, tuberculosis; T-Spot, T-SPOT®.TB.

[a] Risk factors include all of the previously mentioned risk factors, plus health care work, casual contact with a TB case, and prior TB treatment.

LTBI=Latent tuberculosis infection, TST=Tuberculin skin test, QFT=QuantiFERON® Gold-in-tube, T-Spot=T-SPOT®.TB, AUC=Area Under the Curve
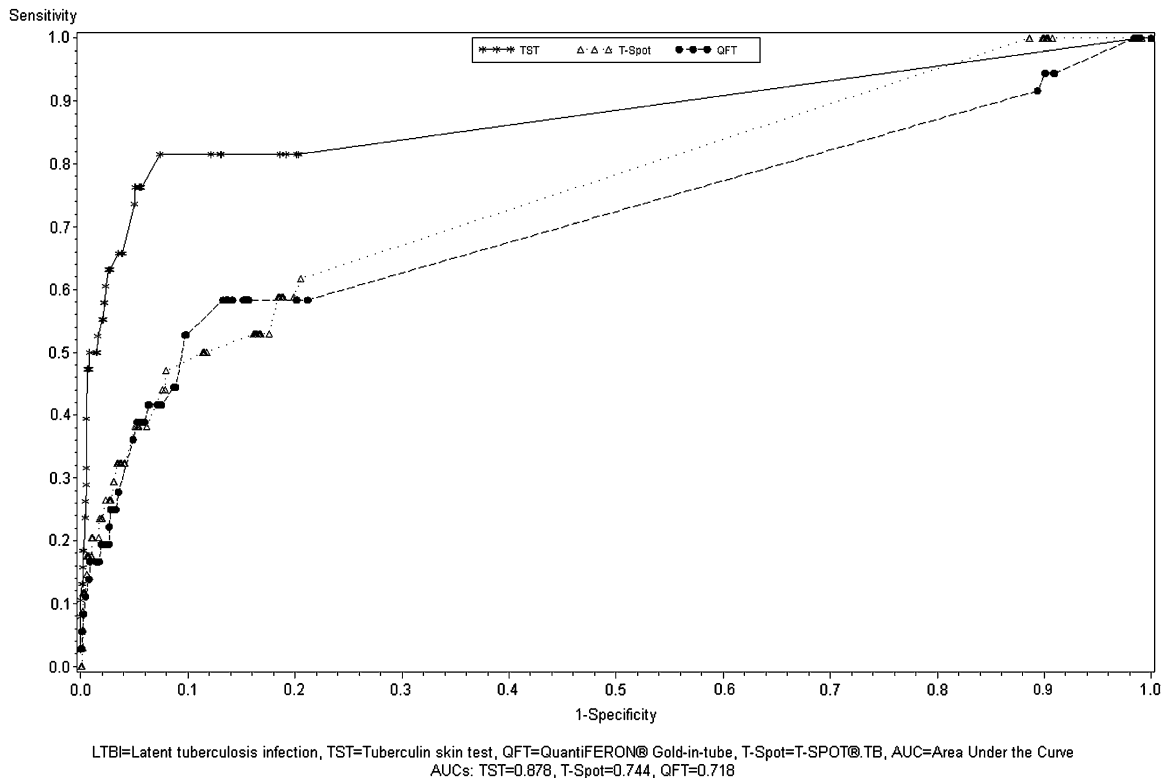AUCs: TST=0.878, T-Spot=0.744, QFT=0.718

Figure 2. Receiver operator characteristics curve for predictors of latent tuberculosis infection (LTBI), as measured by the tuberculin skin test (TST), QuantiFERON® Gold In-Tube (QFT), and T-SPOT®.TB (T-Spot) among US Army Recruits. Predictors included in the logistic regression model: close contact with a tuberculosis (TB) case, casual contact with a TB case, TB prevalence in country of birth, history of living with parent born outside the US, prior positive skin test result, prior TB treatment, history of living in a congregate setting, and health care work. AUC, area under the curve.

This study demonstrates that targeted testing using an RFQ is a useful strategy to test for LTBI and can be operationalized with acceptable performance characteristics using any of the commercially-available tests, consistent with CDC recommendations [25]. Although, the RFQ in this study was better at predicting a positive TST result than for either IGRA, it does not demonstrate that the TST is superior for use in conjunction with targeted testing for several reasons. This study may have been somewhat biased by the use of CDC RSI, because the factors under evaluation were also correlated with a positive result of both the RFQ and the TST. Similarly, the use of a history of a prior positive TST result may bias the prediction model in favor of the TST, although it is noted that the RFQ still had superior sensitivity and specificity in predicting TST, compared with the IGRAs, even when discarding this as a risk factor. This is seen in the 3 variable models in Tables 3, 4, and 5. It is concerning that 56% of positive IGRA results would be missed by the use of the RFQ as compared with 21% for the TST. However, the vast majority of these discordant positive results were positive for only 1 of the 3 tests and had no identifiable risk factors, suggesting that most were false-positive results. In addition, the known risk factors for TB had weaker associations with the IGRAs than with the TST, and no new risk factors were

identified using the IGRAs. Finally, the RSI used for the TST increased the specificity of the test, decreasing the number of false-positive results. Because this is not currently done for the IGRAs, this may bias targeted testing against them in this type of evaluation. Therefore, the most likely explanation for the lower predictive ability in a low-prevalence population such as this is false-positive IGRA results. This suggests that risk-stratified interpretation of IGRAs, as is done for the TST, may be useful. It also suggests that IGRAs should not be used to replace targeted testing, because testing in low-prevalence populations will still result in false-positive results, even if the specificity is very high.

As with the TST, testing with IGRAs will result in false-positive results if IGRAs are used in low-prevalence populations. Regardless of the test used, targeted testing is critical in reducing unnecessary testing and treatment and is consistent with CDC guidelines [25]. Targeted testing in this population would reduce testing by >90%, which would in turn reduce costs of the screening program and adverse events from therapy while still maintaining effectiveness. Some studies suggest that more than 50% of all positive results in low prevalence populations be false-positive results due to nontuberculous mycobacterium and other factors [3, 26, 27]. Targeted testing should therefore

reduce treatment for people with false-positive results who derive no benefit from LTBI therapy but still incur the risk of adverse events.

Future studies suggested by this study include further analysis to improve targeted testing in US and other populations. Analysis to determine the magnitude and relative cost-effectiveness of targeted testing programs for the IGRAs versus the TST is also warranted. Prediction models in other populations may also be considered, including health care workers, prison guards, long-term travelers, and military service members deploying to TB-endemic countries [28, 29]. Finally, studies comparing the long-term rate of progression to active TB among TST- and IGRA-positive persons will allow a more accurate determination of LTBI status and risk of progression to active TB.

References

1. Centers for Disease Control and Population. Targeted tuberculin testing and treatment of latent tuberculosis infection. American Thoracic Society. MMWR Recomm Rep **2000**; 49:1–51.
2. Geiter L, ed. Ending neglect: the elimination of tuberculosis in the United States. Washington, DC: National Academy of Sciences, 2000.
3. Rose DN, Schechter CB, Adler JJ. Interpretation of the tuberculin skin test. J Gen Intern Med **1995**; 10:635–42.
4. Jensen PA, Lambert LA, Iademarco MF, Ridzon R. Guidelines for preventing the transmission of Mycobacterium tuberculosis in health-care settings, 2005. MMWR Recomm Rep **2005**; 54:1–141.
5. Bailey WC, Gerald LB, Kimerling ME, et al. Predictive model to identify positive tuberculosis skin test results during contact investigations. JAMA **2002**; 287:996–1002.
6. Aissa K, Madhi F, Ronsin N, et al. Evaluation of a model for efficient screening of tuberculosis contact subjects. Am J Respir Crit Care Med **2008**; 177:1041–7.
7. Ozuah PO, Ozuah TP, Stein RE, Burton W, Mulvihill M. Evaluation of a risk assessment questionnaire used to target tuberculin skin testing in children. JAMA **2001**; 285:451–3.
8. Froehlich H, Ackerson LM, Morozumi PA. Targeted testing of children for tuberculosis: validation of a risk assessment questionnaire. Pediatrics **2001**; 107:E54.
9. Koppaka VR, Harvey E, Mertz B, Johnson BA. Risk factors associated with tuberculin skin test positivity among university students and the use of such factors in the development of a targeted screening program. Clin Infect Dis **2003**; 36:599–607.
10. Edwards LB, Palmer CE. Part II. tuberculous infection. In Lowell AM, ed. Tuberculosis. Cambridge, Massachusetts: Harvard University Press, 1969; 123–204.
11. Pai M, Zwerling A, Menzies D. Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. Ann Intern Med **2008**; 149:177–84.
12. Bennett DE, Courval JM, Onorato I, et al. Prevalence of tuberculosis infection in the United States population: the national health and nutrition examination survey, 1999–2000. Am J Respir Crit Care Med **2008**; 177:348–55.
13. Mazurek GH, Zajdowicz MJ, Hankinson AL, et al. Detection of *Mycobacterium tuberculosis* infection in United States Navy recruits using the tuberculin skin test or whole-blood interferon-gamma release assays. Clin Infect Dis **2007**; 45:826–36.
14. World Health Organization. Global tuberculosis database. March 24 2009. Available at: http://www.who.int/tb/country/global_tb_database/en/. Accessed 27 December 2009.
15. Cellestis. Quantiferon-TB gold (in tube method) package insert. Valencia, CA: Cellestis, Inc., 2007.
16. Oxford Immunotec. T-SPOT.TB package insert. Marlborough, MA: Oxford Immunotec Limited, 2008.
17. Centers for Disease Control and Population. Mantoux tuberculin skin test: facilitator guide. 2008. Available at: http://www.cdc.gov/TB/pubs/Mantoux/images/Mantoux.pdf. Accessed 4 November 2008.
18. Centers for Disease Control and Population. National health and nutrition examination survey: tuberculosis skin test procedures manual. 2000. Available at: http://www.cdc.gov/nchs/data/nhanes/tb.pdf. Accessed 4 November 2008.
19. Cassell DL. Don't be loopy: re-sampling and simulation the SAS(R) way. SAS Global Form 2007. Cary, NC, 2007.
20. Efron B, Tibshirani R. An introduction to the bootstrap. Boca Raton, FL: Chapman and Hall/CRC, 1998.
21. Scholten JN, Fujiwara PI, Frieden TR. Prevalence and factors associated with tuberculosis infection among new school entrants, New York City, 1991–1993. Int J Tuberc Lung Dis **1999**; 3:31–41.
22. Gounder CR, Driver CR, Scholten JN, Shen H, Munsiff SS. Tuberculin testing and risk of tuberculosis infection among New York City schoolchildren. Pediatrics **2003**; 111:e309–15.
23. Lobato MN, Hopewell PC. *Mycobacterium tuberculosis* infection after travel to or contact with visitors from countries with a high prevalence of tuberculosis. Am J Respir Crit Care Med **1998**; 158:1871–5.
24. Mancuso JD, Tobler SK, Keep LW. Pseudoepidemics of tuberculin skin test conversions in the U.S. Army after recent deployments. Am J Respir Crit Care Med **2008**; 177:1285–9.
25. Mazurek GH, Jereb J, Vernon A, LoBue P, Goldberg S, Castro K. Updated guidelines for using Interferon Gamma Release Assays to

detect *Mycobacterium tuberculosis* infection - United States, 2010. MMWR Recomm Rep **2010**; 59:1–25.

26. Huebner RE, Schein MF, Bass JB Jr. The tuberculin skin test. Clin Infect Dis **1993**; 17:968–75.

27. Rust P, Thomas J. A method for estimating the prevalence of tuberculosis infection. Am J Epidemiol **1975**; 101:311–22.

28. Cobelens FG, van Deutekom H, Draayer-Jansen IW, et al. Risk of infection with *Mycobacterium tuberculosis* in travellers to areas of high tuberculosis endemicity. Lancet **2000**; 356:461–5.

29. US Army. Supplemental guidance for the Army Latent Tuberculosis Infection (LTBI) Surveillance and Control Program. Office of the Surgeon General, 2008. (25 September 2008).