Online Data Supplement

Examining test cutoffs to optimize diagnosis of latent tuberculosis infection in non–US-born people

Authors:

Sofia Zavala, MD, MSc[1], Kathryn Winglee, PhD[2], Christine S. Ho, MD, MPH[2], April C. Pettit, MD, MPH[3], Amina Ahmed, MD[4], Dolly J. Katz, PhD[2], Robert W. Belknap, MD[5], Jason E. Stout, MD, MHS[1] for the Tuberculosis Epidemiologic Studies Consortium

1) Division of Infectious Diseases and International Health, Department of Medicine, Duke University Medical Center, Durham, NC, USA
2) Division of Tuberculosis Elimination, Centers for Disease Control and Prevention, Atlanta, GA, USA
3) Division of Infectious Diseases, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
4) Department of Pediatrics, Atrium Health, Charlotte, NC, USA
5) Denver Health and Hospital Authority, Denver, CO, USA

Supplemental material.

I. Latent class model fit

The latent class model estimated the prevalence of latent TB infection (LTBI) as well as the sensitivity and specificity of the three tests used to diagnose LTBI. Our model incorporated a random individual-participant "intensity" factor (per the method of Dendukuri et al.) that increased or decreased the sensitivity of all tests (in the same direction) for each patient, as well as a separate random "intensity" factor that increased or decreased the specificity of all tests (in the same direction) for each patient. The table below provides observed test result patterns as well as model-predicted patterns with and without these "intensity" factors. The "intensity" factors essentially modeled conditional dependence of the tests, which is biologically plausible given that they are all immunologic tests based on responses to *Mycobacterium tuberculosis* antigens. The model-predicted N without intensity factors was obtained from using a standard latent class model that assumed conditional independence of the tests; this model was obtained using the poLCA package (1)(verson 1.6.0.1) in R. The model used each test as a binary indicator (positive/negative) and fitting a two-class model in which (similar to the Bayesian latent class model described in the main manuscript) the two classes are assumed to represent no LTBI and LTBI, respectively. The default parameters for maximum iterations and convergence for poLCA were used, with random starting values for the expectation-maximization algorithm (also the default).

Test pattern (TST/QFT/TSPOT)	Observed	Model-predicted N (without intensity factor)	Model-predicted N (with intensity factor)
-/-/-	7048	6980	6956
-/-/+	107	175	83
-/+/-	321	274	237
-/+/+	256	303	311
+/-/-	2949	2839	2878
+/-/+	249	359	288
+/+/-	656	488	760
+/+/+	2581	2749	2654

The goodness-of-fit chi-square for the simple model without intensity factor is 148.5, while for the model with intensity factor is 70.9 (both with 7 degrees of freedom). While it is challenging to statistically test this difference given non-nested models, these results suggest that the Bayesian model with the intensity factor provides a better fit to the data than a simple conditional independence model.

II. Sample calculations to evaluate test characteristics at each cutoff.

The table below demonstrates the prevalence of each test combination and the estimated positive predictive value (for LTBI) of each test combination:

Test combination (TST/QFT/T-SPOT)*	Number of participants	Positive predictive value	Estimated number with LTBI
-/-/-	7048	0.0326	230
-/-/+	107	0.7366	79
-/+/-	321	0.660	212
-/+/+	256	0.9962	255
+/-/-	2949	0.0668	197
+/-/+	249	0.9344	233
+/+/-	656	0.9127	599
+/+/+	2581	0.9996	2580
Total	14167		4384

* The test combinations represent positive/negative tests for standard US cutoffs: TST ≥ 10 mm per CDC guidelines (2); QFT ≥ 0.35 IU/mL, and TSPOT ≥ 8 spots per FDA-approved labeling (3, 4) (e.g. for non-US-born people eligible for this analysis, ≥10 mm for the TST, and for all patients ≥0.35 IU/mil TB1/2 antigen-nil for QFT and ≥8 spots difference between panel A/B and nil for TSPOT)

The last column (estimated number with LTBI) is the product of the two middle columns (i.e. for the first row, 7048 x 0.0326=230).

To calculate the sensitivity of a given test at a given cutoff, the assumption is made that within each test combination cell, the probability of having LTBI is homogeneous across test results within that cell. The table below uses the example of an arbitrarily chosen TST cutoff of 4 mm:

Test combination (TST/QFT/T-SPOT)*	Number of participants with TST≥ 4 mm	Estimated number with LTBI and TST ≥4 mm
-/-/-	1067	34.8
-/-/+	32	23.6
-/+/-	67	44.2
-/+/+	95	94.6
+/-/-	2949	197
+/-/+	249	233
+/+/-	656	599
+/+/+	2581	2580
Total	7696	3806

*Test cutoffs are the same as in the previous table. The last four rows agree with those in the previous table since TST≥10 mm ensures TST≥4 mm, whereas the lower numbers in the first four rows correspond to 4 mm ≤ TST < 10 mm.

In this case, the sensitivity of the TST at a cutoff of 4mm or greater would be calculated as the estimated number of individuals with LTBI who had a TST of 4mm or greater divided by the total estimated number of individuals with LTBI in the study, which equals 3806/4384=86.8%. Similar calculations were performed at each designated cutoff for each test to generate the ROC curves.

III. Sensitivity Analysis

Our analysis assumed that the probability of an individual having LTBI was uniform across all the numerical values of a negative or positive test, given that the other two tests results remain unchanged. In other words, we assumed that the probability that an individual had latent tuberculosis with a negative TST of 4mm, negative QFT and negative TSPOT was the same as that of an individual with a negative TST of 9mm, negative QFT and negative TSPOT. One could argue that the risk or probability of LTBI for the individual with a TST of 9mm is higher, because an immunologic process is responsible for the larger skin reaction. We performed a sensitivity analysis to evaluate the impact of violating our assumption of uniform risk regardless of the numerical value of the test. The sensitivity analysis assumed that participants at or above an arbitrary value of a negative test (5 mm for TST, 0.1 IU/mL for QFT, and 3 spots for TSPOT) had a higher prevalence of LTBI than participants below this value. Numerically, we increased the prevalence of LTBI within each test pattern by 25% for participants above the cutoff and correspondingly reduced the prevalence for participants below the cutoff. For example, of the 7048 participants with three negative tests, 6118 had a TST result <5 mm and 930 had a TST result of 5-9 mm. Using the baseline assumption we would estimate that 200 (6118 x 0.0326) of the participants with three negative tests and TST

<5 mm had LTBI, and 30 (930 x 0.0326) with TST 5-9 mm had LTBI. The sensitivity analysis assumed that participants with TST 5-9 mm had a 25% greater risk of LTBI, so about 38 (30 x 1.25) in that group would have LTBI and 192 (200-(30 x 1.25)) with TST <5 mm would have LTBI. The ROC curves did not substantially change in the sensitivity analysis compared with the primary analysis.

IV. Simulating sequential testing

Using the same data from the first table above, one can estimate the sensitivity and specificity of pairs of test combinations for LTBI. For example, the sensitivity of the combination of a positive TST and positive QFT would be calculated as follows:

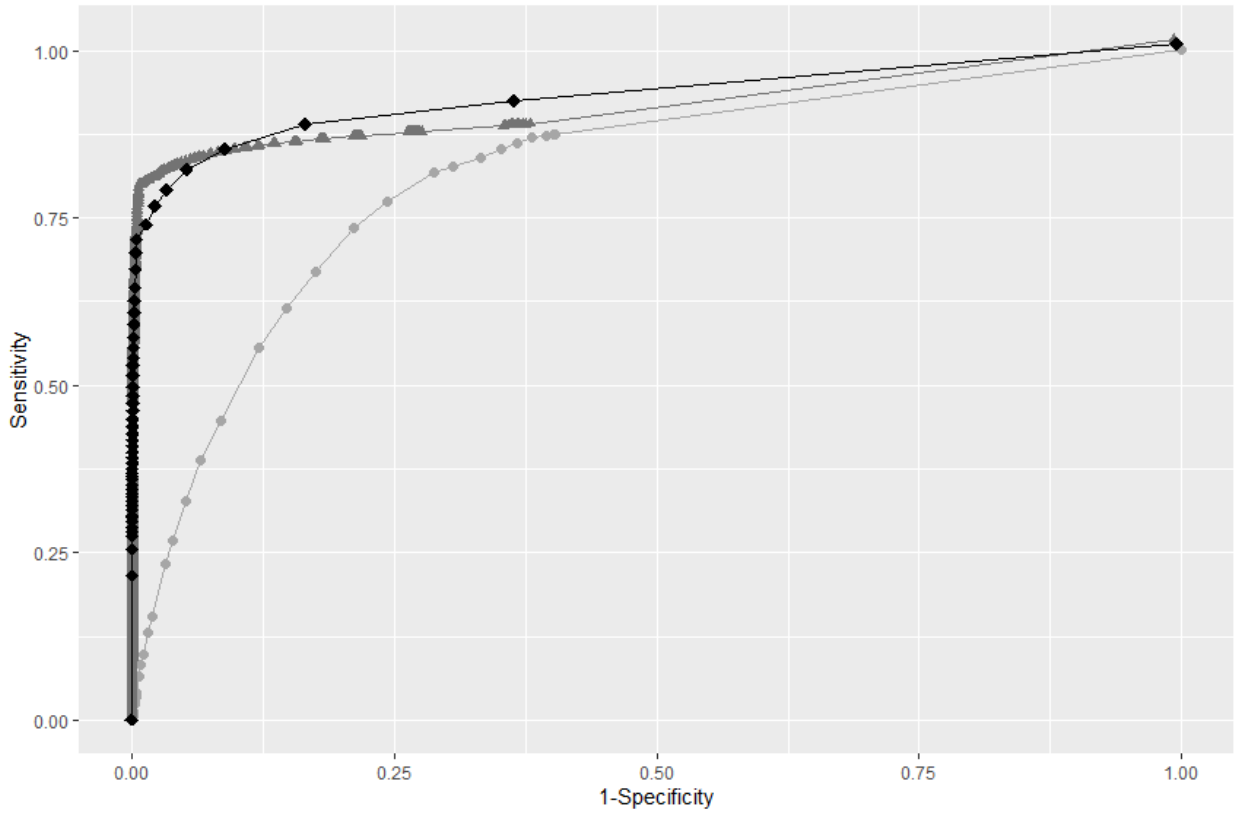
$$\frac{\text{Estimated number of participants with LTBI who have +TST and +QFT}}{\text{Estimated total number of participants with LTBI}}$$

From the table above, 599 + 2580=3179 participants with +TST/+QFT are estimated to have LTBI, out of a total of 4384 total participants estimated to have LTBI, so the sensitivity of the combination is 3179/4384=72.5%. Similarly, the specificity of the combination would be calculated as follows:

$$\frac{\text{Estimated number of participants without LTBI who do not have +TST and +QFT}}{\text{Estimated total number of participants without LTBI}}$$

From the table above, (7048-230) + (107-79) + (321-212) + (256-255) + (2949-197) + (249-233) = 9724 participants who did not have both a positive TST and a positive QFT and who did not have LTBI, while 14167-4384=9783 total participants estimated not to have LTBI, for a calculated specificity of 9724/9783=99.4%.

Supplemental Figure E1. Superimposed receiver operating characteristic curves for the TST, QFT, and TSPOT.



Legend: Receiver operating characteristic curves for the TST (gray circles), QFT (gray triangles), and TSPOT (black diamonds). Note that the TST curve does not intersect the other two curves except at the (0,1) and (1,0) points, while the QFT and TSPOT curves do intersect at other points.

REFERENCES

1. Linzer DA, Lewis JB. poLCA: An R Package for Polytomous Variable Latent Class Analysis. Journal of Statistical Software. 2011;42(10):1 - 29.
2. Targeted tuberculin testing and treatment of latent tuberculosis infection. American Thoracic Society. MMWR Recomm Rep. 2000;49(RR-6):1-51.
3. Food and Drug Administration. QuantiFERON-TB Gold In-Tube - P010033/S011 [Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf/P010033B.pdf.
4. Food and Drug Administration. T-SPOT-TB - P070006 2008 [Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf7/p070006b.pdf.