The effect of new *Mycobacterium tuberculosis* infection on the sensitivity of prognostic TB signatures

T. Sumner,[1] A. Fiore-Gartland,[2] M. Hatherill,[3] R. M. G. J. Houben,[1] T. J. Scriba,[3] R. G. White[1]

[1]TB Modelling Group, TB Centre, Centre for Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK; [2]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [3]South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, Division of Immunology, Department of Pathology, University of Cape Town, Observatory, South Africa

Correspondence to: Tom Sumner, TB Modelling Group, TB Centre, Centre for Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK. email: Tom.sumner@lshtm.ac.uk

Running head: Prognostic TB signatures and new infection

SUMMARY

BACKGROUND: Tests that identify individuals at greatest risk of TB will allow more efficient targeting of preventive therapy. The WHO target product profile for such tests defines optimal sensitivity of 90% and minimum sensitivity of 75% for predicting incident TB. The CORTIS (Correlate of Risk Targeted Intervention Study) evaluated a blood transcriptomic signature (RISK11) for predicting incident TB in a high transmission setting. RISK11 is able to predict TB disease progression but optimal prognostic performance was limited to a 6-month horizon.

METHODS: Using a mathematical model, we estimated how subsequent *Mycobacterium tuberculosis* (MTB) infection may have contributed to the decline in sensitivity of RISK11. We calculated the effect at different RISK11 thresholds (60% and 26%) and for different assumptions about the risk of MTB infection.

RESULTS: Modelled sensitivity over 15 months, excluding new infection, was 28.7 (95% CI 12.3–74.1) compared to 25.0 (95% CI 12.7–45.9) observed in the trial. Modelled sensitivity exceeded the minimum criteria (>75%) over a 9-month horizon at the 60% threshold and over 12 months at the 26% threshold.

CONCLUSIONS: The effect of new infection on prognostic signature performance is likely to be small. Signatures such as RISK11 may be most useful in individuals, such as household contacts, where probable time of infection is known.

KEY WORDS: tuberculosis; biomarkers; preventive therapy; modelling

TB preventive therapy is a key component of the WHO End TB strategy;[1] however, global scale-up of preventive therapy has been slow.[2] One challenge to wider uptake of preventive therapy is the inability to identify individuals at the highest risk of developing TB disease. Current tests for *Mycobacterium tuberculosis* (MTB) infection, including interferon-gamma release assays (IGRAs) and tuberculin skin tests (TSTs), have poor positive predictive value for development of active disease.[3] Tests which could better identify those individuals at greatest risk of developing TB could reduce the TB burden, spare low-risk individuals from unnecessary treatment and ensure a more efficient utilisation of resources.[4]

The WHO and the Foundation for Innovative New Diagnostics (FIND) have developed a target product profile (TPP) for such tests, proposing an optimal sensitivity and specificity of 90% (and minimal criteria 75%) for predicting incident TB in the 2 years after testing in the absence of new or repeat infection with MTB.[5] Several blood transcriptional signatures have been identified as potential predictors of the development of incident disease.[6,7] However, the performance of these signatures has been shown to decline with time since testing in several studies.[8,9]

This decline in performance may reflect the fact that these signatures were developed to identify an early (incipient or minimal) disease state prior to the development of clinical disease, from which individuals may progress and regress at different rates. However, in high-incidence settings, where these tests could have particular value to reduce TB burden, it is possible that new or repeat infection with MTB may occur after testing, which may affect the observed test performance. The extent of this effect will depend on the time since testing, the risk of infection, and differences in the risks of infection and progression to disease among those testing positive and negative.[10]

The CORTIS (Correlate of Risk Targeted Intervention Study) study,[9] which was conducted among the non-HIV-infected adult population in five high TB burden settings in South Africa, evaluated the performance of an 11-gene blood transcriptomic signature (RISK11) for diagnosing prevalent TB and predicting incident TB. The efficacy of the 12-week 3HP (isoniazid plus rifapentine) regimen for the prevention of incident TB was evaluated in RISK11-positive (RISK11+) individuals. The RISK11 signature was able to predict incident TB, but prognostic performance was found to decline with time since screening, with optimal sensitivity (>90%) limited to a 6-month time horizon.

In this paper, we use a simple mathematical model to explore the contribution of infection after testing to TB incidence and the effect on the observed sensitivity of a biomarker test for incident TB using data from the CORTIS study as an example.

METHODS

Data

This analysis uses anonymised data from the CORTIS study.[9] The trial population consisted of HIV-negative adults aged 18–59 years, without history of TB disease in the last 3 years. Participants were recruited via word-of-mouth, house-to-house visits and liaison with non-governmental organisations. Recruitment did not target groups at high risk of TB, such as household contacts. Table 1 gives the baseline characteristics of the population.

In this analysis, we consider the cohort of 2500 individuals who did not have microbiologically confirmed prevalent TB at baseline and did not receive 3HP during the study. Participants were followed up to a maximum period of 15 months for development of incident TB with site visits at 3, 6, 12 and 15 months and telephone or site visits at 1, 2 and 9 months. At interim study visits, participants were asked about symptoms consistent with TB disease, and presence of one or more symptoms triggered TB investigation (paired sputum Xpert MTB/RIF; Cepheid, Sunnyvale, CA, USA; and MGIT™ culture; BD, Franklin Lakes, NJ, USA). At the final study visit, all participants underwent TB investigation regardless of the presence or absence of symptoms. The median duration of follow-up was 13.9 months (interquartile range 9.0–15.0), and 66% of 3HP-negative participants attended at least six scheduled visits. For each participant, baseline RISK11 score, IGRA status, duration of follow-up and time of diagnosis of incident TB was known. We assumed that participants who did not complete screening remained disease-free. As rates of study withdrawal were similar by RISK11 status, these are unlikely to significantly affect our results.

In our main analysis, we used the primary trial definition of RISK11 positivity (60% threshold RISK11 score) and primary endpoint definition of microbiologically confirmed TB on at least two separate sputum samples.

The CORTIS study was sponsored by the University of Cape Town, Cape Town, South Africa, and the study protocol was approved by the Institutional Human Ethics Committees of each participating site. All participants provided written informed consent in their language of choice.

Estimating TB due to new infection

For each participant in the trial, we simulated the probability of being infected in each month of follow-up and, if infected, the probability of developing disease due to this new infection in the time from infection to the end of their follow-up. The number of months of follow-up for each participant was extracted from the data. A random binomial number was drawn (with

probability λ_m) for each month to determine if and when an individual was infected. If infected, a random binomial number was then drawn for each remaining month (with probability d_m) to determine if and when an individual progressed to disease.

The annual risk of infection (λ) was estimated for each arm of the trial based on the prevalence of IGRA positivity (P(I)), the average age of the population (A = 28.4 years) and estimates of the sensitivity (SE) (SE = 88, 95% confidence interval [CI] 84–91) and specificity (SP) (SP = 94, 95% CI 91–96) of the IGRA assay (QuantiFERON TB Gold-Plus; Qiagen, Hilden, Germany) for MTB infection.[11,12] First, the true prevalence of infection (P(L)) was calculated as follows:

$$P(L) = \frac{P(I) - 1 + SP}{SE - 1 + SP}$$

Then, λ is given by:

$$\lambda = -\left(\frac{1}{A}\right) \ln\left(1 - P(L)\right)$$

This estimation assumed individuals had experienced a constant force of infection over their lifetime and that the force of infection was constant for each month of follow-up such that $\lambda_m = \lambda/12$.

The risk of developing disease in the first year following infection was assumed to be d = 8.7% (95% CI 8.17–9.05) based on previous modelling work.[13] As for the force of infection, the risk of disease was assumed to be constant over time such that $d_m = d/12$. We assumed that the risk of disease following infection was lower in those who were IGRA-positive, reflecting the protective effect of a prior infection. In our primary analysis, we assumed prior infection reduced the risk of disease following reinfection by 41%.[13]

Effect of TB due to new infection on test sensitivity

Sensitivity estimates from the trial data were calculated using the standard formula (where I indicates the observed cases from the trial):

$$S_{obs} = \frac{I_{RISK11+}}{I_{RISK11+} + I_{RISK11-}}$$

To calculate the effect of new exposure on the performance measured in the trial, the predicted cases due to new exposure (P) were subtracted from the observed cases (I) and sensitivity recalculated as follows:

$$S_{adj} = \frac{\max(0, I_{RISK11+} - P_{RISK11+})}{\max(0, I_{RISK11+} - P_{RISK11+}) + \max(0, I_{RISK11} - P_{RISK11-})}$$

Confidence intervals (CIs) for RISK11 sensitivity were estimated using 20,000 samples of 2,500 individuals drawn, with replacement, from the trial population. For each sample, values for the sensitivity and specificity of IGRA and risks of disease were drawn from ranges given above and the number of cases due to new exposure estimated, as described above.

Secondary analyses

We repeated the analysis at an alternative RISK11 positivity threshold of 26% (identified in exploratory analysis of the CORTIS trial as meeting the minimum TPP specificity of 75% over 15 months),[9] and using the secondary trial endpoint definition of microbiologically confirmed TB on at least one sputum sample.

We explored the effect of assumptions about the risk of infection and disease progression. We repeated the analysis assuming the annual risk of infection was half or double that estimated from the prevalence of IGRA positivity. We also repeated the analysis assuming prior infection reduced the risk of disease following reinfection by 81% (vs. 41% in primary analysis).[14]

All code used in the analysis is available from https://github.com/tomsumner/CORTIS_reinfection

RESULTS

The estimated annual risk of infection was higher in the RISK11+ population (5.2%, 95% CI 4.6–5.9) than in the RISK11– arm (4.1%, 95% CI 3.8–4.4). This reflects the higher prevalence of IGRA positivity among RISK11+ individuals than in RISK11– individuals in the CORTIS trial population (69.1%, 95% CI 65.9–72.6 vs. 62.6%, 95% CI 60.1–64.7).[9] The Figure (top row) shows the cumulative number of incident TB cases observed in the trial (solid lines) and the median number due to new infections that occurred after RISK11 testing (dashed lines) predicted by the model. The final number of cases due to new infection is small: 1 (0–4) among RISK11+ participants; 2 (0–6) among RISK11– participants.

At a RISK11 threshold of 60%, the adjusted sensitivity over 15 months was 28.7 (95% CI 12.3–74.1) compared to 25.0 (95% CI 12.7–45.9) observed in the CORTIS trial; at the lower threshold of 26%, the adjusted sensitivity was 55.7 (95% CI 22.2–100.0) compared to 47.5 (95% CI 25.9–75.0) observed in the CORTIS trial. The Figure (bottom row) shows the

sensitivity by month since RISK11 screening at 9, 12 and 15 months (observed and adjusted sensitivity was 100% to 6 months; not shown). The median adjusted sensitivity exceeds the observed values at all times, although the 95% confidence intervals overlap. Of note, the median-adjusted sensitivity exceeded the minimum sensitivity of 75% in the TPP up to 9 months at the 60% threshold and up to 12 months at the 26% threshold.

Table 2 shows estimates of the adjusted sensitivity over 15 months from the secondary analysis using both the primary and secondary trial endpoint definitions and risk thresholds, and for different assumptions about the annual risk of infection and protection against disease due to prior infection. The highest estimated sensivity is 72.1 (95% CI 27.2–1.00) based on two sample endpoints, a RISK11 threshold of 26, 81% protective effect of prior infection and an annual risk of infection of approximately 10%, double that estimated from the trial data.

DISCUSSION

Our results suggest that infection with MTB after testing may have contributed to the drop in sensitivity of the RISK11 signature over time in the CORTIS study. However, even in a high-transmission settings (with an estimated annual risk of infection of approximately 5%), the effect was relatively small, such that the RISK11 signature would still not meet the minimum TPP criteria for an incipient TB test over a 15-month horizon at either threshold considered. Based on this analysis, RISK11 would not be expected to meet the minimum TPP sensitivity over 15 months in a low-transmission setting.

This finding supports the view that the sensitivity of biomarkers such as RISK11 for predicting incident TB in HIV-negative individuals is likely to be limited by the natural progression from the pre-clinical disease state measured by the test(s) to active disease, and not by risk of re-infection. If this is the case, such tests may be most useful in individuals such as household contacts, where the probable timing of exposure relative to testing, and therefore, the period of highest risk of progression is known. The test performance may also be affected by spontaneous reversion of the pre-clinical disease state: if a significant proportion of individuals do not progress to active disease, the specificity of tests which identify pre-clinical states will be limited.[15] We are unable to address this question with the data from the CORTIS study.

Our analysis had some limitations. We assumed that the risk of infection was constant over time and age, and therefore, may have under or over-estimated the risk of exposure following RISK11 screening. However, when we assumed an annual risk of infection of approximately 10%, the median adjusted sensitivity still did not exceed the minimum TPP

criteria over 15 months. In addition, our estimates of disease risks are drawn from modelling of populations in different settings and may not be completely generalisable to the current setting.

During the CORTIS study, testing for incident TB was triggered by symptoms at interim trial visits, and all participants were evaluated for TB regardless of symptoms at the final visit. As a result, asymptomatic cases that occurred during follow-up may only have been identified at the final study visit. In contrast, our model assumes that all cases due to new exposure are identified when they occur. While this does not affect our estimates of the adjusted sensitivity at 15 months, it may affect findings at interim time points.

References

1 World Health Organization. The End TB Strategy. Geneva, Switzerland: WHO, 2015.

2 Surie D, et al. Policies, practices and barriers to implementing tuberculosis preventive treatment-35 countries, 2017. Int J Tuberc Lung Dis 2019; 23(12): 1308–1313.

3 Rangaka MX, et al. Predictive value of interferon-gamma release assays for incident active tuberculosis: a systematic review and meta-analysis. Lancet Infect Dis 2012; 12(1): 45–55.

4 Sumner T, et al. Potential population level impact on tuberculosis incidence of using an mRNA expression signature correlate-of-risk test to target tuberculosis preventive therapy. Sci Rep 2019; 9(1): 11126.

5 World Health Organization. Consensus meeting report: development of a Target Product Profile (TPP) and a framework for evaluation for a test for predicting progression from tuberculosis infection to active disease. Geneva, Switzerland: WHO, 2017.

6 Zak DE, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet 2016; 387(10035): 2312–2322.

7 Suliman S, et al. Four-gene pan-African blood signature predicts progression to tuberculosis. Am J Respir Crit Care Med 2018; 197(9): 1198-1208

8 Gupta RK, et al. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. Lancet Respir Med 2020; 8(4): 395–406.

9 Scriba TJ, et al. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. Lancet Infect Dis 2021; 21:354-365.

10 Kik SV, et al. An evaluation framework for new tests that predict progression from tuberculosis infection to clinical disease. Eur Respir J 2018; 52(4): 1800946

11 Barcellini L, et al. First independent evaluation of QuantiFERON-TB Plus performance. Eur Respir J 2016; 47(5): 1587–1590.

12 Wang M, He J. Sensitivity and specificity of QuantiFERON-TB Gold Plus to detect *Mycobacterium tuberculosis* infection. Am J Respir Crit Care Med 2020; 201:A5438

13 Vynnycky E, Fine PEM. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. Epidemiol Infect 1997; 119: 183–201.

14 Andrews JR, et al. Risk of progression to active tuberculosis following reinfection with *Mycobacterium tuberculosis*. Clin Infect Dis 2012; 54(6): 784–791.

15 Kendall EA, Shrestha S, Dowdy DW. The epidemiological importance of subclinical tuberculosis: a critical re-appraisal. Am J Respir Crit Care Med 2020; 203(2): 168-174

Table 1 Baseline characteristics of the CORTIS trial population

Sex	%
Female	54.2
Male	45.8
Age, years, mean ± SD	28.4 ± 9.0
Ethnicity	
Asian	0.1
Black African	66.6
Caucasian	0.1
Mixed race	33.1
BMI, kg/m^2, mean ± SD	24.6 ± 7.7
Prior TB	7.9
Smoking	50.6
Family TB history	15.8

CORTIS = Correlate of Risk Targeted Intervention Study; BMI = body mass index.

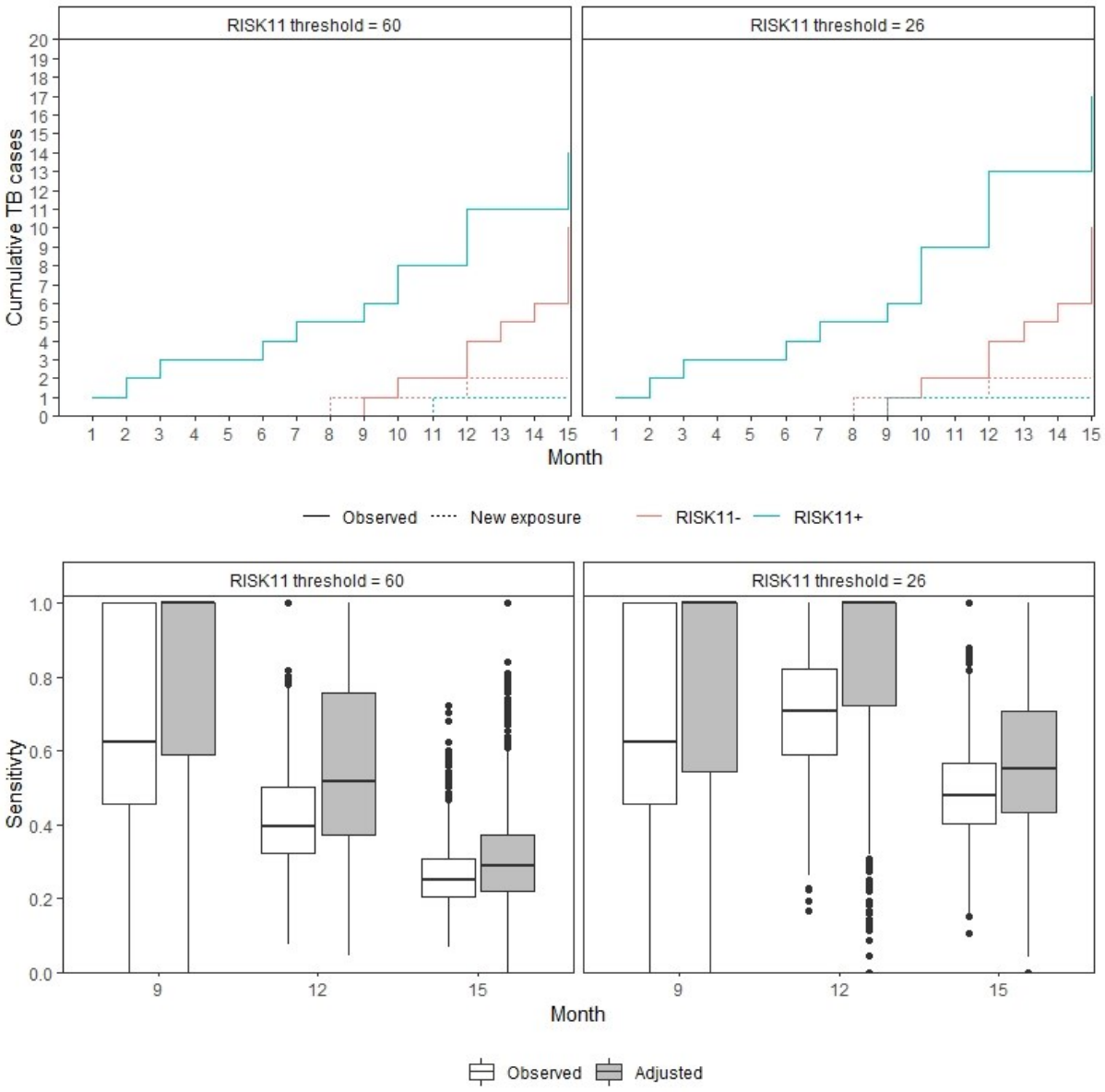
Table 2 Adjusted sensitivity estimates from secondary analyses at 15 months

RISK11 threshold	Endpoint definition	Protection	Relative annual risk of infection: 0.5 Median (95% CI)	1 Median (95% CI)	2 Median (95% CI)
60%	2 samples	0.41	26.3 (12.5–54.3)	28.7 (12.3–74.1)	34.3 (12.2–1.00)
		0.81	26.3 (12.8–52.6)	27.6 (13.1–60.7)	30.0 (12.5–84.6)
	1 or 2 samples	0.41	21.6 (12.6–34.7)	22.4 (13.7–36.6)	23.4 (12.1–41.6)
		0.81	21.6 (13.2–33.9)	22.0 (13.3–35.4)	22.7 (13.7–37.7)
26%	2 samples	0.41	51.4 (24.7–90.0)	55.7 (22.2–100.0)	67.2 (28.6–1.00)
		0.81	51.5 (25.4–1.00)	56.6 (26.7–1.00)	72.1 (27.2–1.00)
	1 or 2 samples	0.41	42.8 (26.5–62.8)	43.8 (26.8–66.1)	46.8 (25.5–76.5)
		0.81	43.3 (27.0–62.8)	44.5 (26.6–67.3)	48.6 (28.9–76.3)

CI = confidence interval.

FIGURE LEGEND

Figure Top row: cumulative incident TB cases. Red lines show RISK11– cases, blue lines RISK11+. Solid lines show the number of cases observed in the trial, dashed lines the predicted number due to new-exposure. Bottom row: Sensitivity estimates by time since testing. Observed (white) and adjusted (grey) sensitivty estimates by time since screening (left hand panels: RISK11 threshold of 60%; right hand panels: RISK11 threshold of 26%).


RISK11 threshold = 60
RISK11 threshold = 26
Cumulative TB cases
Month
Observed
New exposure
RISK11-
RISK11+
RISK11 threshold = 60
RISK11 threshold = 26
Sensitivty
Month
Observed
Adjusted

Figure