TB prevalence in people tested is a strong predictor of Xpert specificity in community and risk group screening

Nicky McCreesh*, Indira Govender[1,2], Alison D. Grant[1,2], Palwasha Y Khan[1]

[1]TB Centre, London School of Hygiene & Tropical Medicine

[2]Africa Health Research Institute, School of Laboratory Medicine & Medical Sciences, College of Health Sciences, University of KwaZulu-Natal*Corresponding author.
Nicky.mccreesh@lshtm.ac.uk LSHTM, Keppel Street, London, WC1E 7HT, UK

Keywords: Tuberculosis; diagnostics; specificity; Xpert

We thank Veeken *et al* for their excellent review of studies estimating the specificity of Xpert in community settings[1]. We agree with their conclusion that the specificity of Xpert may often be higher in community settings than is suggested by estimates from testing people with presumptive TB, and would like to highlight an additional finding that can be drawn from the data they assembled.

Using data from Veeken *et al*[1] and a Cochrane systematic review[2], we plotted estimated Xpert specificity against estimated TB prevalence, from studies conducting community-wide screening (community active case finding or prevalence surveys[1], Figure 1a and b) and from studies testing people in high-risk groups[2] (Figure 1c). For studies conducting community-wide screening, we plotted the data separately by estimated prevalence in all people screened (Figure 1a) and estimated prevalence in people who were screen positive according to the study criteria only (Figure 1b). For studies in high-risk groups, only estimates of prevalence in people tested were available. Finally, we split the data by TB NAAT diagnostic test and interpretation of trace results: algorithm 1 includes both Xpert MTB/RIF and Xpert Ultra with trace considered negative, and algorithm 2 consists of Xpert Ultra with trace considered positive.

The plots show that there are very strong relationships between prevalence and specificity for the community-wide screening studies, with specificity decreasing with increasing prevalence. Estimated specificity correlated less strongly with the estimated prevalence in the people *screened*, versus in people who were *screen positive* (r = -0.75 *vs* -0.96 respectively for algorithm 1, and r = -0.91 *vs* -0.98 respectively for algorithm 2). The observed relationship between prevalence and specificity was less strong for studies conducted in high-risk groups (r = -0.72, Figure 1c), and absent for studies conducted in people with presumptive TB (r = 0.50 for algorithm 1 and r = 0.38 for algorithm 2, using data from a second Cochrane systematic review[3], plot not shown).

Figure 1d shows the predicted relationship between estimated prevalence in people tested and the ratio of false positive to true positive diagnoses in community screening. For both algorithms, the ratio of false positive to true positive diagnoses is predicted to be relatively stable at higher prevalences, but increases sharply when prevalences are low.

We suggest two factors contributing to the relationship between prevalence and specificity. The first is that people with recent TB, resolved through treatment or self-cure, have an increased risk of false positive Xpert results[4]. The second is that the 'gold standard' test, culture, has an imperfect sensitivity[5], and some people will therefore be incorrectly classified as false positives. The prevalence of both people with recent TB and 'false false positives' are likely to correlate

strongly with the prevalence of current TB in a population. People with recent TB and 'false false positives' are also likely to have an increased probability of screening positive on symptom screen and/or chest x-ray, providing an explanation for why specificity correlates more strongly with prevalence in people *tested* than in people *screened*. In more highly selected groups of participants, such as people with presumptive TB, it is probable that many additional factors will have affected study participation, masking the relationship between prevalence and specificity.

The World Health Organization recommends that community screening can be conducted in settings with TB prevalences of 0.5% or higher[6]. At this prevalence, we estimate that the ratio of false positive to true positive diagnoses will be 0.85 (1.73) with algorithm 1 (algorithm 2), with only moderate reductions to 0.74 (1.34) and 0.63 (1.01) if symptom screening (sensitivity 71%, specificity 64%[6]) or chest x-ray screening (any abnormality, sensitivity 94%, specificity 89%[6]) are used. These figures are likely to be overestimates, as they do not allow for the imperfect sensitivity of culture.

There are a number of limitations to this analysis. We assumed that the estimated prevalence reported by the studies was the true prevalence. Xpert MTB/RIF and Xpert Ultra (trace as negative) were grouped based on observed similarities in the data, as opposed to an *a priori* decision. The results are based on a post-hoc analysis and limited data, and should therefore be treated as hypothesis-generating. Nevertheless, the strength and biological plausibility of the observed relationship lends substantial support to the hypothesis that TB prevalence is a strong predictor of Xpert specificity in community and risk group screening. Additional data from future prevalence surveys and research studies can be used to validate the observed relationships, and well as improve estimates of the relationship between prevalence and specificity.

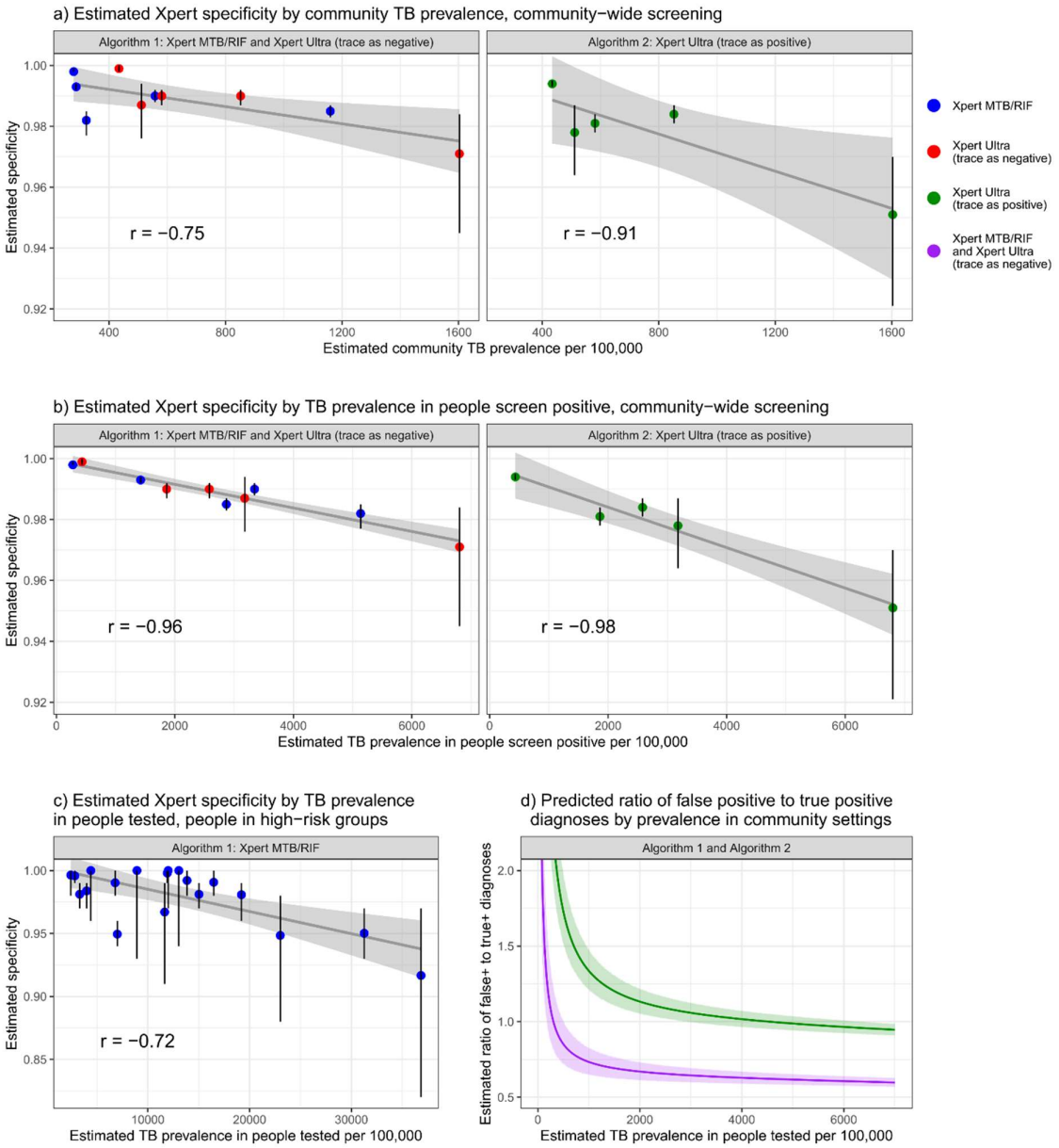

Figure 1. a – c Relationship between TB prevalence and estimated Xpert specificity. a) and b) in community settings and c) in people in high-risk groups; a) estimated TB prevalence in the community; b) estimated TB prevalence in people screen positive; c) estimated TB prevalence in people tested. d) Relationship between TB prevalence in people tested and estimated ratio of false positive to true positive diagnoses, calculated as (1 – prevalence) × (1 – specificity) / (prevalence × sensitivity) , and assuming a sensitivity of 0.618 for Xpert MTB/RIF and Xpert Ultra (trace as negative) and 0.690 for Xpert Ultra (trace as positive). Uncertainty intervals are 95% confidence intervals incorporating the uncertainty in the ratio of false positive to true positive diagnoses

Author contributions

Conceptualisation: NM; Methodology: NM, PYK; Writing – original draft: NM; Writing – review & editing: PYK, ADG, IG; Visualisation: NM

Funding sources

NM is funded by the UK Medical Research Centre (MR/Z504270/1)

Conflicts of interest

The authors report no conflict of interests

Data availability

Data and code are available at https://github.com/NickyMcC/SpecificityLetter

References

1. Veeken LD, Schwalb A, Horton KC, et al. Mind the clinic-community gap: how concerned should we be about false positive test results in mass tuberculosis screening? *The Journal of Infectious Diseases*. 2025:jiaf268.
2. Zifodya JS, Kreniske JS, Schiller I, et al. Xpert Ultra versus Xpert MTB/RIF for pulmonary tuberculosis and rifampicin resistance in adults with presumptive pulmonary tuberculosis. *Cochrane Database of Systematic Reviews*. 2021(2).
3. Shapiro AE, Ross JM, Yao M, et al. Xpert MTB/RIF and Xpert Ultra assays for screening for pulmonary tuberculosis and rifampicin resistance in adults, irrespective of signs or symptoms. *Cochrane Database of Systematic Reviews*. 2021(3).
4. Theron G, Venter R, Smith L, et al. False-Positive Xpert MTB/RIF Results in Retested Patients with Previous Tuberculosis: Frequency, Profile, and Prospective Clinical Outcomes. *Journal of clinical microbiology*. 2018;56(3):10.1128/jcm.01696-01617.
5. Yilmaz A, Bayram Selvi U, Damadoğlu E, et al. Diagnostic value of repeated sputum examinations in pulmonary tuberculosis: how many sputum specimens are adequate? *Tuberk Toraks*. 2008;56(2):158-162.
6. World Health Organization. *WHO consolidated guidelines on tuberculosis. Module 2: screening – systematic screening for tuberculosis disease*. Geneva2021.