NAAT for tuberculosis detection in blood: systematic reviews and meta-analyses for Research in context box, Lancet Microbe submission

Linda Boloko, David Barr

13/12/2020

Contents

1 Preface

This document reports two systematic reviews and meta-analyses of literature on use of nucleic acid amplification technology (NAAT) to detect *M. tuberculosis* in patient blood samples. The first meta-analysis is focused on blood NAAT for TB diagnosis; the second is on bloood NAAT to diagnose *M. tuberculosis* blood stream infection (MTBBSI).

The motivation for these reviews is to inform the "Evidence before this study" section of the *Research in context* box for the KDHTB blood Xpert-ultra manuscript submission to *Lancet Microbe*.

This document is an Rmarkdown knitted as a pdf. This means alll the code for the analysis from raw data to final figures is embedded (so 100% reproducible). The code chunks are suppressed for readability but available at git repository [url:].

2 Systematic Review 1: TB diagnosis using blood NAAT

2.1 Introduction

2.1.1 Aims & objectives

We want to systematically review literature on use of nucleic acid amplification technology (NAAT) to identify *M. tuberculosis* in patients' blood samples as a diagnostic for tuberculosis. Aims are to summarise:

1. What NAAT methods including blood pre-processing have been used for identifying *M. tuberculosis* in blood.
2. Reported sensitivity and specificity of blood *M. tuberculosis* NAAT **for tuberculosis**.

Systematic review and meta-analysis of studies using NAAT on blood to diagnose TB will be performed.

2.1.2 Inclusions & exclusions

Studies in which investigators used NAAT to identify *M. tuberculosis* in peripheral blood samples (whole blood or component) from patients identified prospectively with either suspected tuberculosis (e.g. cohort design) or confirmed tuberculosis diagnosis (e.g. case-control design) will be included.

Studies where it is unclear if patients were identified prospectively for blood NAAT testing (e.g. studies where inclusion was based on opportunistic receipt of a blood sample), studies where it is unclear what reference standard for tuberculosis diagnosis was, and studies reporting artificially spiked sample experiments (patient samples spiked with *M. tuberculosis ex vivo*) will be excluded.

2.1.3 Data for extraction

We will extract data on patient populations (adult versus paediatic, HIV status, TB prevalence, pulmonary versus extra-pulmonary, inpatient versus outpatient) and NAAT method (commercial v in-house, blood pre-processing, blood volume).

2.1.4 Analysis plan

Descriptive summaries using figures and tables. Bivariate random-effects regression acounting for correlation between sensitivity and specificity will be used to summarised central tendancies and heterogeneity. These will be fit using a Baysian (MCMC) approach implemented with the package *brms* in R studio. Meta-regression

on selected covariates will be performed using bivariate regression to test association between study and method covariates and diagnostic performance.

2.1.5 Bias assessment

Identified studies will be assessed for risk of bias using questions adapted from the QUADRA-2 tool:

- **Patient selection** . Are methods of patient selection adequately described (prior testing, presentation, intended use of index test and setting)? Was a consecutive or random sample of patients enrolled in a cohort design? Was HIV status of patients ascertained?
- **Index test** . Were the index test results interpreted without knowledge of the results of the reference standard?
- **Reference standard** . Were the reference standard results interpreted without knowledge of the results of the index test? Did all patients have at least 2 TB diagnostic tests (e.g. culture, NAAT, antigent testing, excluding the index test) performed from at least 2 different body sites (e.g. sputum and urine)? Was the index test excluded from the reference standard?
- **Flow and timing** . If blood NAAT was performed on only a subgroup of patients are the inclusion/exclusion criteria for this subgroup given? Was blood NAAT performed from samples taken at same timepoint as reference standard samples?

If answers to >1 or >3 of these questions are "no" or "unclear" risk of bias will be rated as moderate or high.

2.2 Search strategies

2.2.1 PubMed

Terms used in main PubMed search engine:

tuberculosis AND (blood OR mycobacteraemia OR "blood stream infection" OR bacteraemia OR bacillaemia) AND (NAAT OR PCR OR Xpert) AND diagnosis

This is translated by PubMed algorithm into an expanded search query which we have edited to remove irrelevant search terms (e.g. "blood" is linked to a range of haematology terms which are irrelavant and these have been removed by editing in the advanced search editor). This give a final expanded search query of:

("tuberculosi"[All Fields] OR "tuberculosis"[MeSH Terms] OR "tuberculosis"[All Fields] OR "tuberculoses"[All Fields] OR "tuberculosis s"[All Fields])
AND
("blood"[MeSH Subheading] OR "blood"[All Fields] OR "blood"[MeSH Terms] OR "mycobacteraemia"[All Fields] OR "blood stream infection"[All Fields] OR ("bacteraemia"[All Fields] OR "bacteremia"[MeSH Terms] OR "bacteremia"[All Fields] OR "bacteraemias"[All Fields] OR "bacteremias"[All Fields]) OR "bacillaemia"[All Fields])
AND
("nucleic acid amplification techniques"[MeSH Terms] OR ("nucleic"[All Fields] AND "acid"[All Fields] AND "amplification"[All Fields] AND "techniques"[All Fields]) OR "nucleic acid amplification techniques"[All Fields] OR "naat"[All Fields] OR "PCR"[All Fields] OR "Xpert"[All Fields]) AND ("diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[MeSH Subheading])

This retured 951 results which have been exported as a .nbib file (from above URL, [send to]–>[Citations Manager]–>[all results]–>[create file]) saved as *pubmed.nbib* in the working directory.

2.2.2 Scopus

Scopus was searched with query:

(TITLE-ABS-KEY (tuberculosis AND diagnosis) AND TITLE-ABS-KEY (pcr OR naat OR xpert) AND TITLE-ABS-KEY (blood OR mycobacteraemia OR bacteraemia OR bacillaemia OR "blood steam infection"))

Returning 537 results on 12/12/2020; these are exported as a .bib file using [select all–>BibTeX export, including abstract] saved as scopus.bib in the working directory.

2.2.3 Combining search results, removing duplicates

The .bib files are read in and combined, with duplicates identified by doi and removed:

[embedded code supressed]

2.3 Screening results

Summary of identified, screened, eligible and included study shown in PRISMA flow diagram in figure 1.

2.4 Description of included studies

Characteristics of review 1 included studies are shown in figure 2 (study design, setting, patient population/cases, assessed risk of bias) and figure 3 (NAAT methods). A wide variety of study designs, patient populations and NAAT methods have been reported, with little replication of specific approaches. Initial reports of blood NAAT for TB diagnosis had a peak in the 1990s then a relative hiatus, followed by an incresae again in last 10 years.

2.5 Descriptive summaries of reported diagnostic performance

Reported diagnostic performance in included studies are shown in figure 4-6. There are several striking findings. First, there is very marked heterogeneity in sensitivity (ranging from 0 to 100%), with very little evidence of correlation between sensitivity and specificity. Heterogeneity in sensitivity seems unrelated to study design, setting, HIV prevalance in study, or NAAT methods (blood volume, method of blood preprocessng) other than the finding that the studies using Xpert Rif/MTB had amongst the lowest reported sesnitivities. Average reported sensitivity of blood NAAT seems to have slightly decreased over time since initial reports in 1990s.

However, reported sensitivity does appear to vary somewhat by assessed risk of bias, with studies assessed as lower risk of bias, and larger studies, reporting lower sensitivity.

There is also evidence of reporting bias with an asymetrical funnel plot suggesting smaller studies with high sensitivity are over represented (figure 7).

951 studies from	537 studies from
PubMed search	Scopus search

246 duplicates by
automated search

1242 for manual
title/abstract screen

1108 excluded by
title/abst review

134 for manual
full–text review

Excluded, 92:
Duplicate reference = 24
NAAT not performed on peripheral blood = 15
Non–eligible study design = 11
Patient selection not prospective or unclear = 34
Reference standard absent or unclear = 8

Included, 42 uses of NAAT on blood to diagnose TB
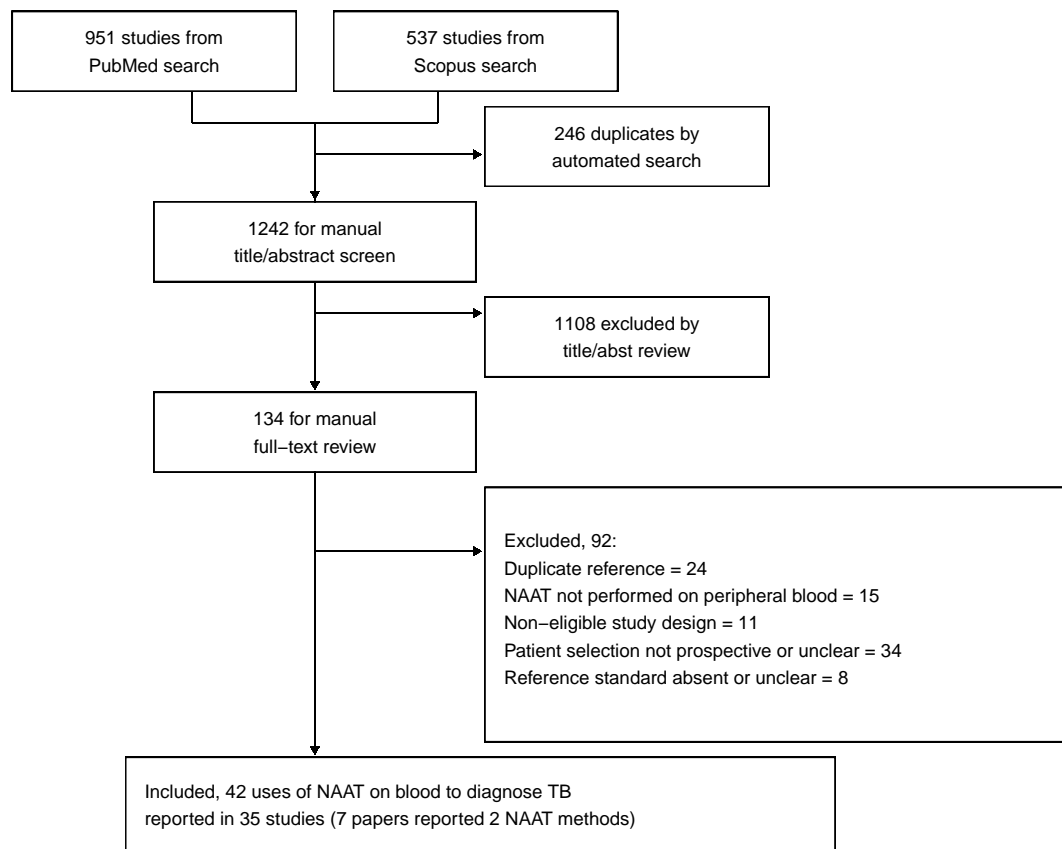reported in 35 studies (7 papers reported 2 NAAT methods)

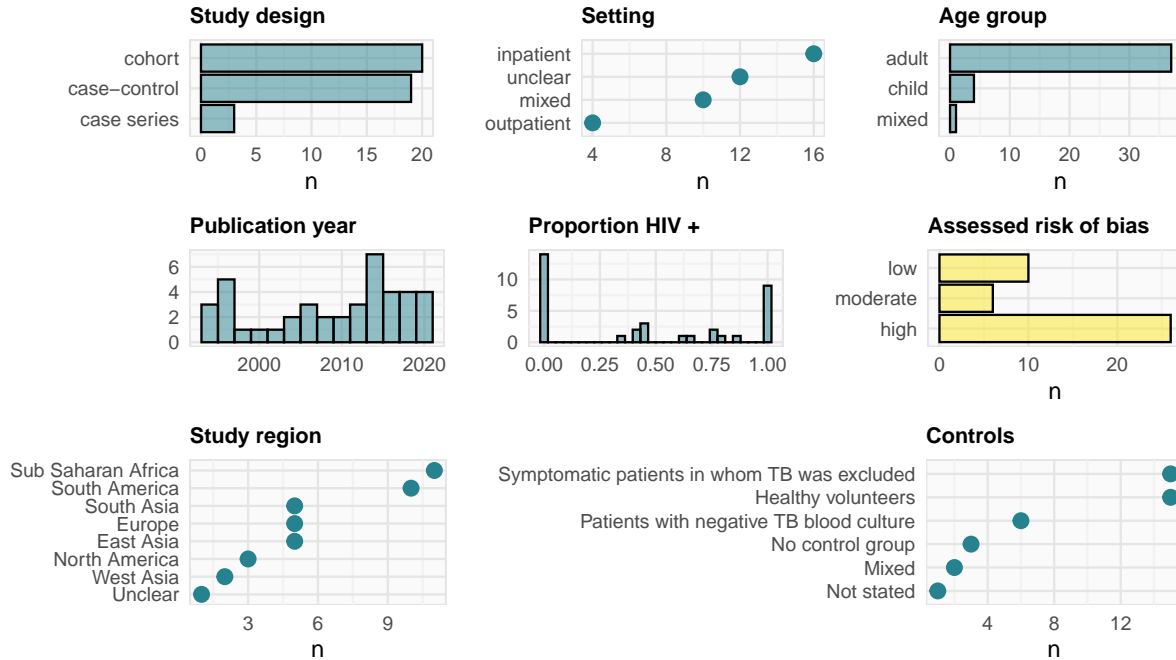Figure 1: PRISMA flow chart for systematic review 1

Figure 2: Description of review 1 included studies: study design, setting, patient population/cases, assessed risk of bias.

2.6 Bivariate regression modelling

Formal bivariate modelling of the reported sensitivities and specificities largely confirms impressions from descriptive plots in previous section. Heterogeneity in reported sensitivity is extreme, with the 90% prediction interval (in which the model estimates 90% of studies from the "population" of studies will lie) encompassing nearly all possible values of sensitivity (90%PI = 0.088 to 0.97).

2.6.1 Meta-regression: covariates versus diagnostic performance

Association of study-level covariates with reported diagnostic performance of blood NAAT was formally investigated using bivariate random-effects modelling.

These was some evidence that reported sensitivity was associated with risk of bias assessed through adapted QUADAS-2: studies assessed as low risk of bias reported lower sensitivity on average (figure 9). Posterior probability that low-bias risk studies had lower reported sensitivity than 'high or moderate' bias risk studies was 93%.

There was also a 95% posterior probability that larger sample size studies reported lower sensitivity than smaller sample size studies (figure 10).

There was no significant evidence that reported sensitivity for TB diagnosis was improving over time as NAAT technologies have evolved (rather there was a 'non-significant' weak downward trend, with 75% posterior probability that sensitivity was *decreasing* by year of publication (figure 10).

Proportion of study participants who were HIV positive was not convincingly related to reported diagnostic performance of NAAT with 67% posterior probability that sensitivity was higher in studies recruiting more patients who were HIV positive (figure 10).
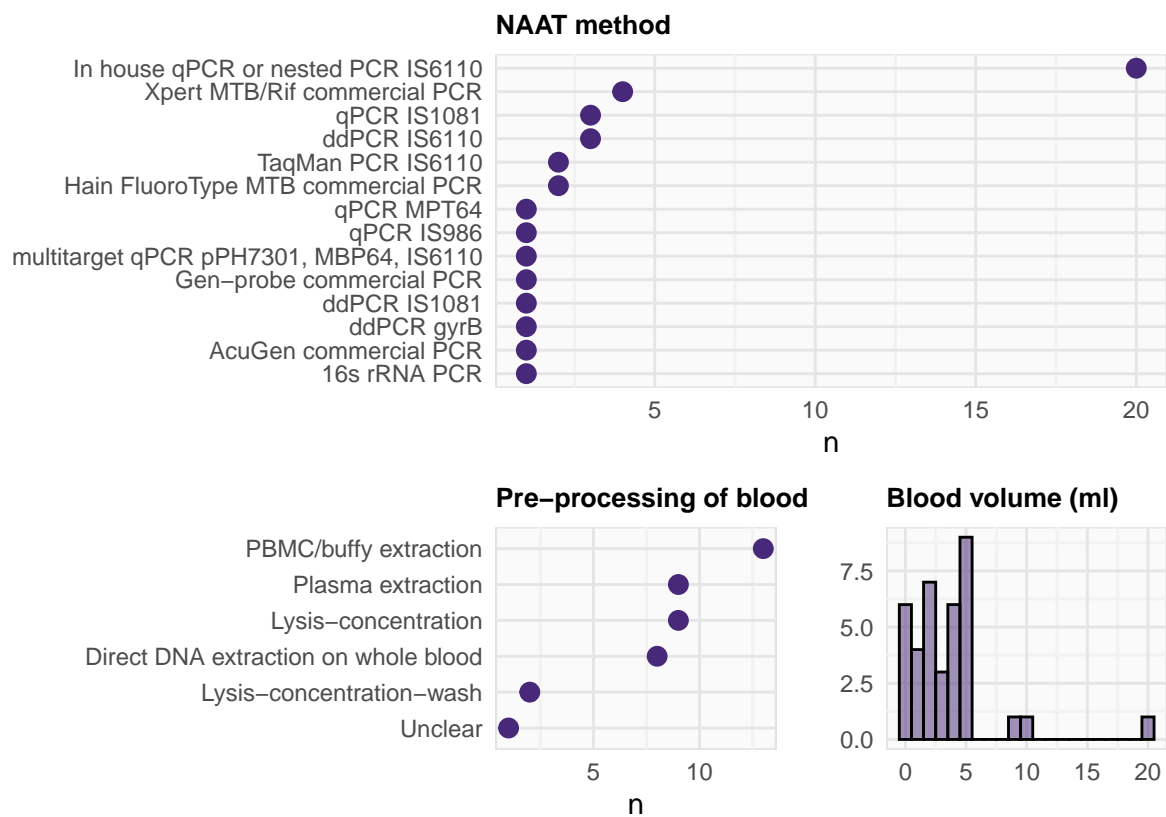
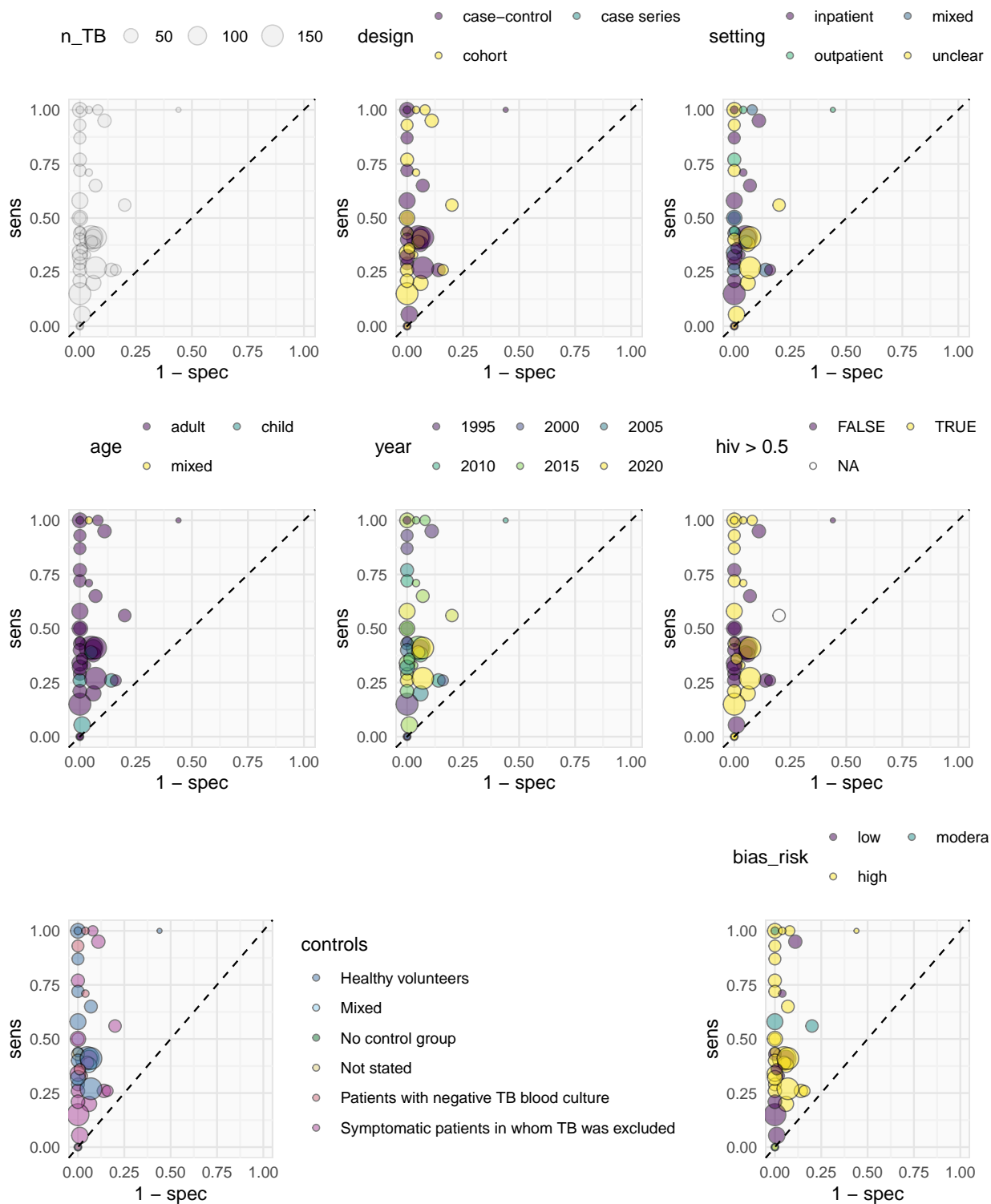Figure 3: Description of review 1 included studies: NAAT methods used.

Figure 4: Reported diagnostic performance (sens v 1-spec, blood NAAT for TB diagnosis, raw data, review 1) by study characteristic covariates
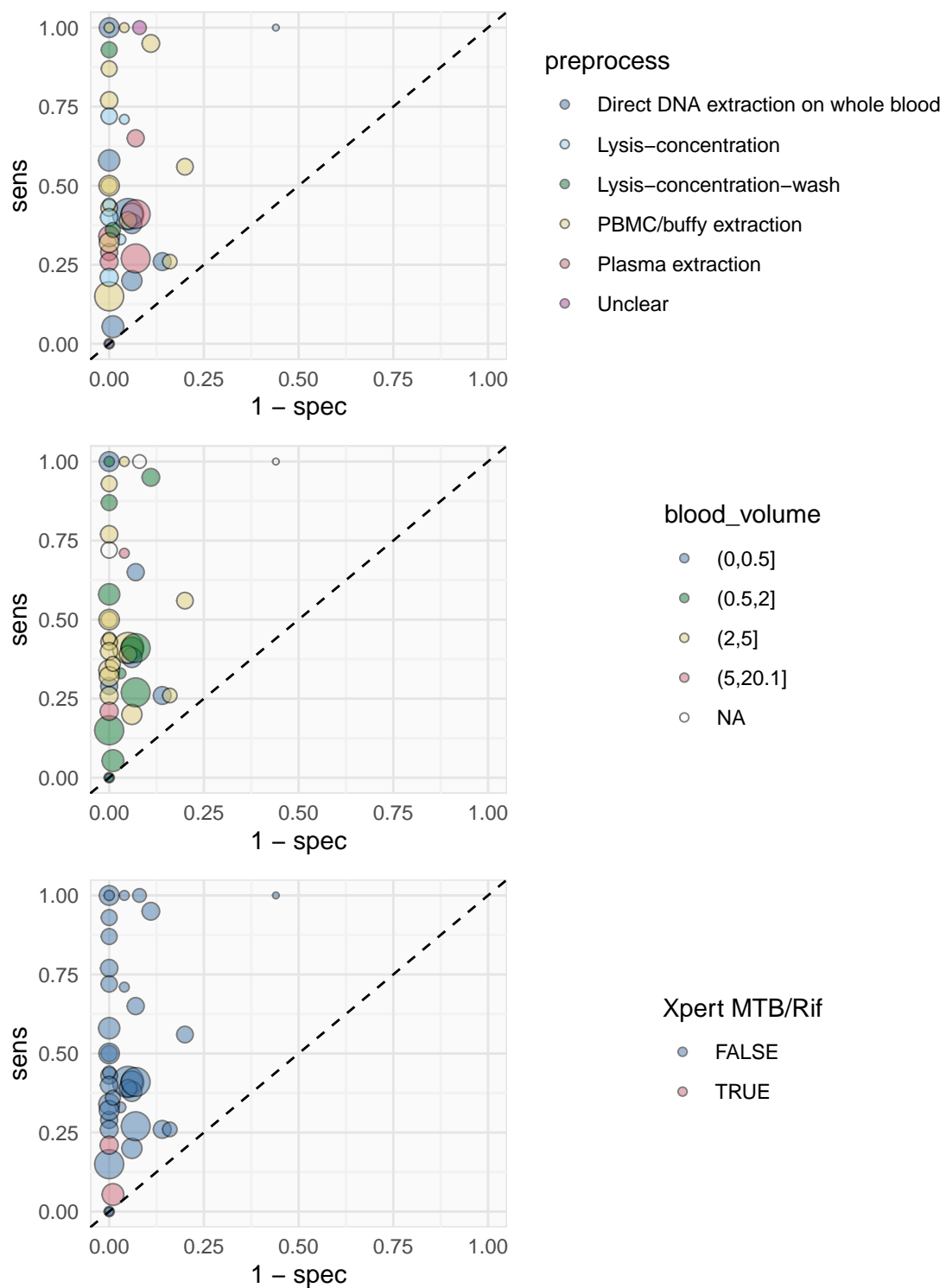
Figure 5: Reported diagnostic performance (sens v 1-spec, blood NAAT for TB diagnosis, raw data, review 1) by study NAAT methodology
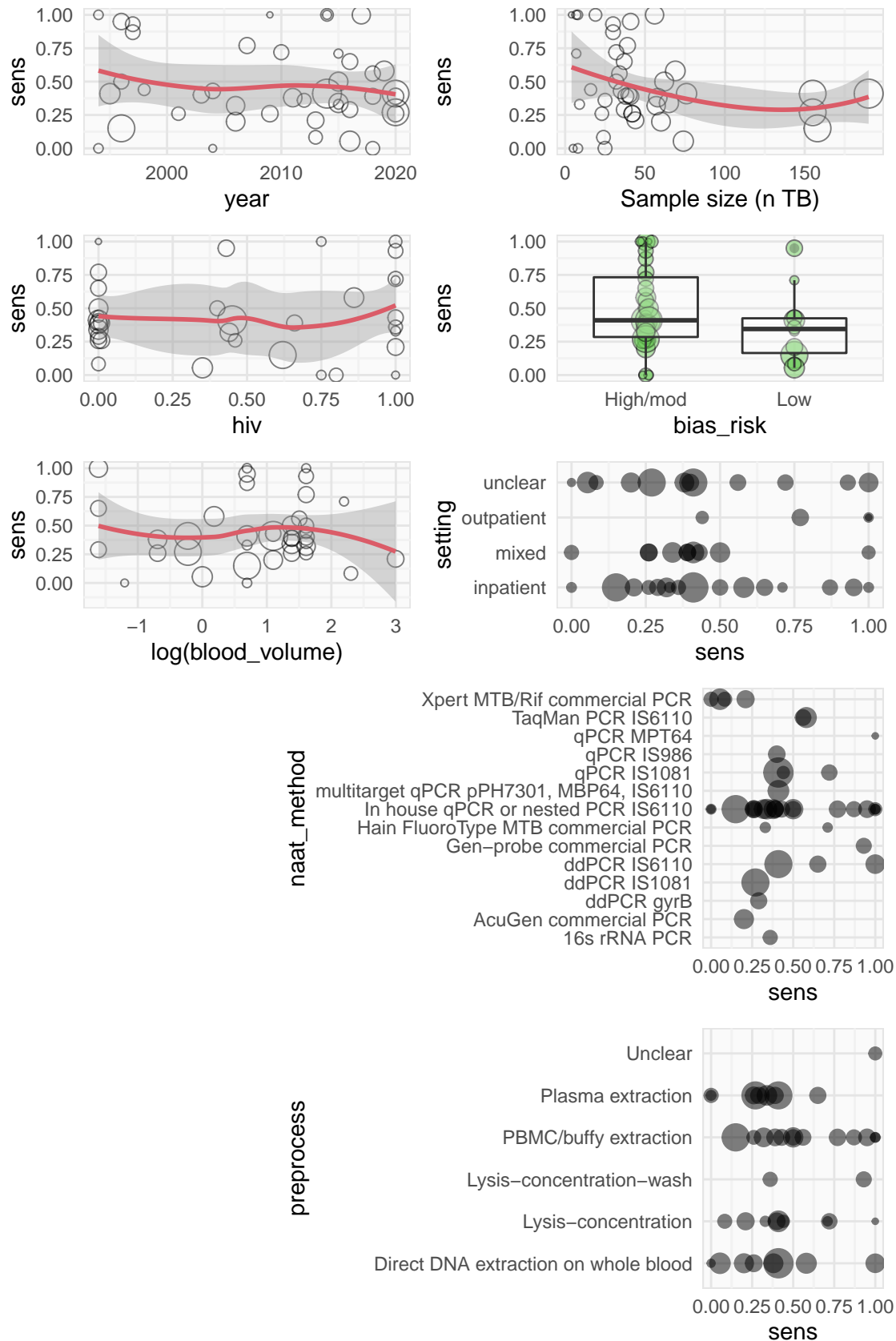
Figure 6: Reported sensitivity of blood NAAT for TB diagnosis (review 1, raw data) by selected covariates
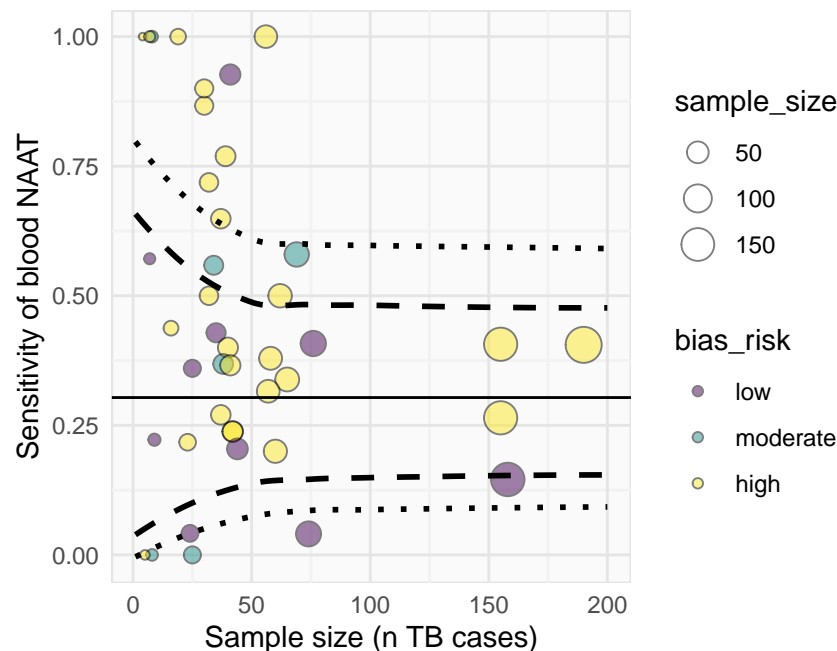
Figure 7: Evidence of bias in design and reporting versus reported sensitivity of blood NAAT for TB diagnosis (review 1).

2.7   Summary review 1

Since the 1990s dozens of reports describing use of NAAT on patient blood samples for TB diagnosis have been published, with extreme heterogeneity in reported sensitivity and specificity, not obviously related to plausible biological or technical covariates. Most studies are poorly reported and are assessed to have high risk of bias. Mostly, in house PCR methods have been used with a wide variety of specific methodologies. Promising results in smaller reports have not been replicated in larger, low-bias studies. Results from scalable / widely available commercial PCR platforms have been disappointing.

3   Systematic Review 2: *M. tuberculosis* blood stream infection diagnosis using blood NAAT

3.1   Introduction

Marked heterogeneity in sensitivity of blood NAAT for diagnosing tuberculosis could be related to disease spectrum of included TB cases in studies, not captured by gross study-level covariates such as inpatient v outpatient and proportion of patients HIV positive examined in review 1 above. We know that severity of HIV-associated tuberculosis is closely related to presence of *M. tuberculosis* blood stream infection (MTBBSI) and the presence of MTBBSI is clearly a plausible determinant of probability that TB is detected by blood NAAT testing. Detection of MTBBSI by blood culture is therefore a useful reference standard against which blood NAAT detection of tuberculosis can be assessed. We hypothesised that comparing blood NAAT to blood culture for detection of *M.tb* could resolve some of the heterogeneity in reported sensitivity by accounting for variance in disease spectrum of recruited cases in different studies, and therefore proposed a second systematic review limited to studies which peformed both blood NAAT and blood culture for detection of MTBBSI.
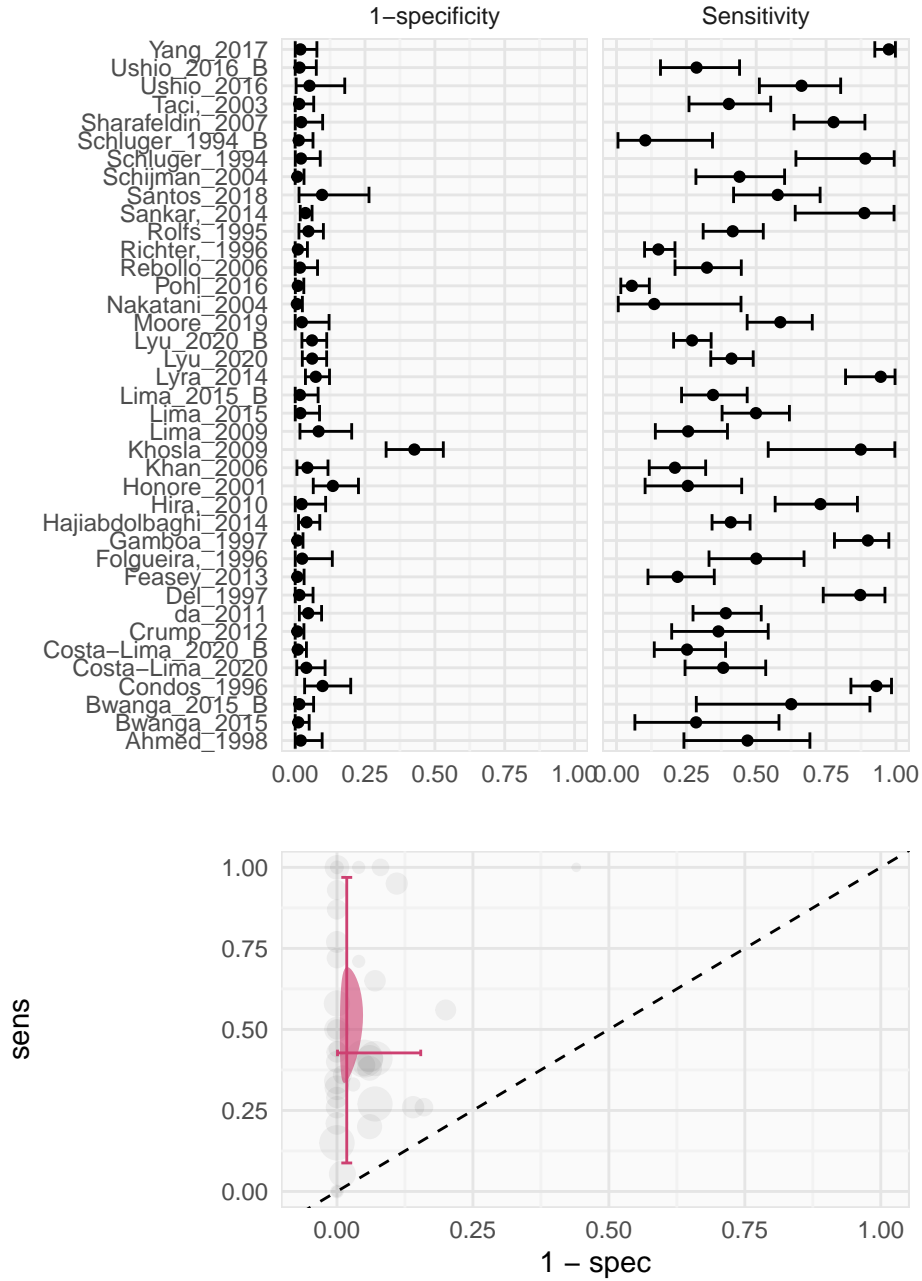
Figure 8: Bivariate random-effects modelling reported sensitivity and specificity blood NAAT for TB diagnosis (review 1). Forest plots show model estimates of mean values with 95% credibility intervals for each study. Lower plot shows 95% credibility ellipse for mean (expected) values, with 90% prediction intervals in which 90% of studies are expected to lie.
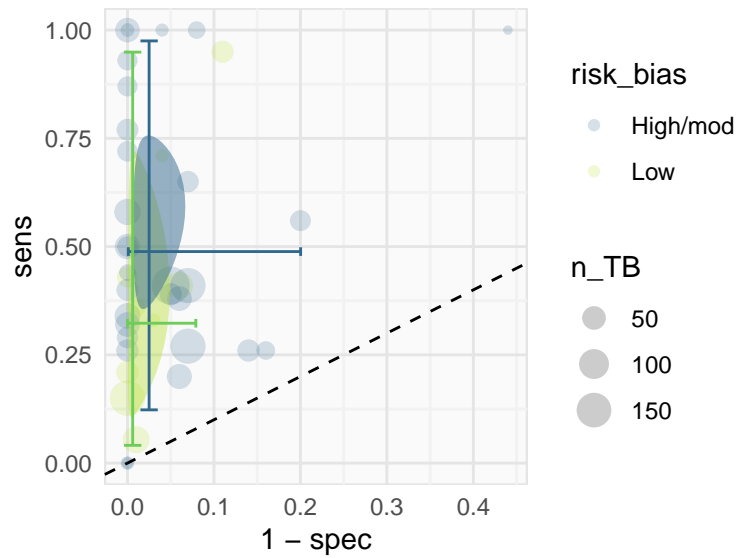
Figure 9: Bivariate random-effects model with assessed risk of study bias as covariate (high/moderate bias = blue; low bias = green) for performance of blood NAAT for TB diagnosis (review 1). Ellipses show 95%CrI for mean (expected) values; 90% prediction intervals shown with whisker intervals.
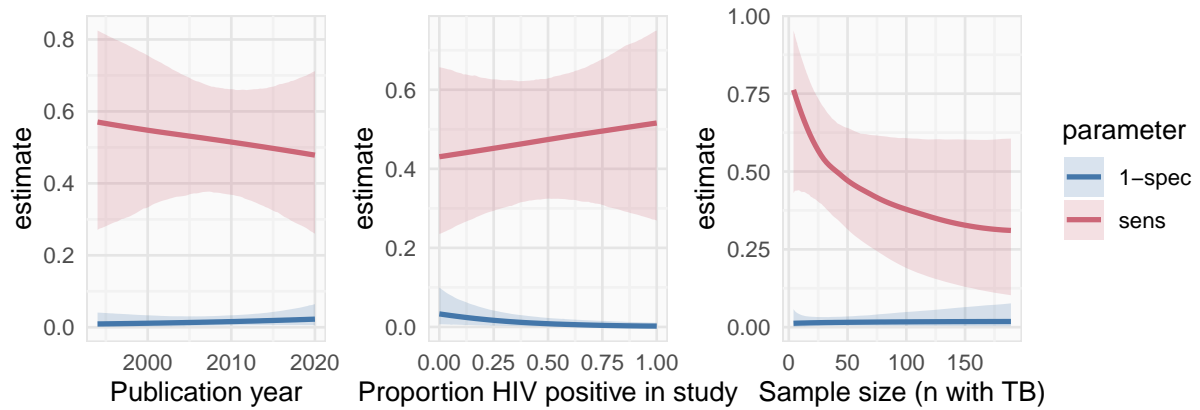

Figure 10: Conditional effects of 3 study-level covariates on reported performance in blood NAAT studies for TB diagnosis (review 1).

3.1.1 Aims and objectives

Objective is to summarise reported sensitivity of blood NAAT compared to mycobacterial blood culture.

Aims:

1. Summarise reported relative sensitivity of blood TB-NAAT and TB blood culture against an external reference standard.
2. Summarise reported sensitivity of blood TB-NAAT against TB blood culture as the reference standard. Note that specificity estimation is not an objective: this is justified as heterogeneity in sensitivity in systematic review 1 is the problem to be addressed.

3.1.2 Included studies

Included studies will be the subset of studies in review 1 that also performed a TB blood culture (liquid or solid media).

Studies which do not report results such that a 2x2 table cross-tabulating blood culture and NAAT results could be extracted were excluded.

[post hoc protocol edit:] *Studies where the TB reference standard was based on TB blood culture were excluded for aim 1 but were retained in aim 2 analysis].*

3.1.3 Data for extraction

1. **n_tb_diagnosis** : Number of patients who were classified as having TB diagnosis who had a valid TB blood culture and a blood NAAT performed.
2. **n_bloodculture** : Number of TB patients who were TB blood culture positive.
3. **n_bloodnaat** : Number of TB patients who were TB blood NAAT positive.
4. **n_bc_naat** : Number of TB blood culture positive patients who were also blood NAAT positive.

In addition to covariates assessed in review 1.

3.1.4 Analysis plan

Descriptive summaries using figures and tables.

For aim 1 (*Summarise reported relative sensitivity of blood TB-NAAT and TB blood culture against an external reference standard*) we anticipate correlation between reported sensitivity of blood culture and blood NAAT by study. Therefore a bivariate random-effects regression acounting for this correlation will be used to summarise central tendancies and heterogeneity. Ratio or difference measures for sensitivity of the two methods will then be derived.

For aim 2 (*Summarise reported sensitivity of blood TB-NAAT against TB blood culture as the reference standard*) a univariate binomial regression model with random-effects by study will be used to summarise central tendancies and heterogeneity.

Models will be fit using a Baysian (MCMC) approach implemented with the package *brms* in R studio. Meta-regression on selected covariates will be performed using bivariate regression to test association between study and method covariates and diagnostic performance if sufficient studies are identified to do so.

3.2 Screening results

From the 42 studies identified for review 1, 16 performed a mycobacterial blood culture; 2 of these did not report the results such that a 2x2 cross-tabulation of blood culture and blood NAAT could be extracted,

leaving 14 studies with both blood culture and blood NAAT for inclusion in review 2 (figure 11, PRISMA flow chart). Of these 14 studies, 5 used TB blood culture result as the reference standard for TB diagnosis, meaning n=9 studies were available for aim 1. In one study all TB blood cultures were negative, meaning n=13 studies were available for aim 2 analysis.
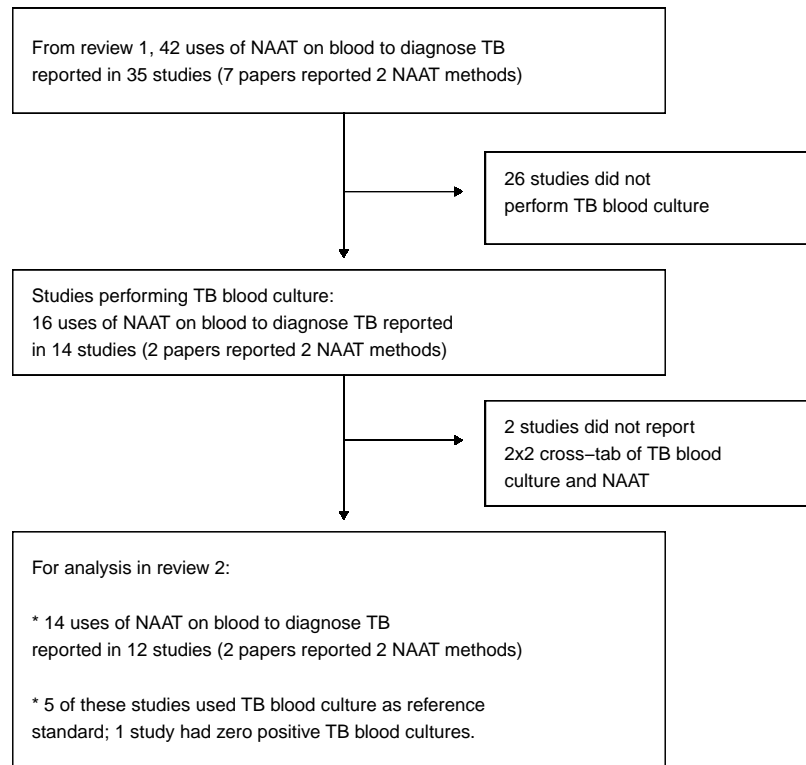

Figure 11: PRISMA flow chart for systematic review 2

3.3 Comparing sensitivity of blood TB-NAAT and TB blood culture against an external reference standard

In studies reporting sensitivity of blood NAAT and bood culture for TB diagnosis against an external reference standard there was, as expected, correlation between the sensitivities of the two methods across studies (r = 0.48 estimated from bivariate mixed-effects regression). Most studies reported higher sensitivity of NAAT compared to culture, but with substantial heterogeneity resulting in uncertainty and 95% credibility intervals encompassing both better and worse sensitiviy for NAAT (figure 12). Larger sample size and lower risk-of-bias studies reported lower relative sensitivity of NAAT on average, but the number of studies available did not support formal meta-regression.

3.4 Assessing sensitivity of blood TB-NAAT against TB blood culture as the reference standard

13 studies have reported TB blood culture and TB blood NAAT results in same patients. Estimated sensitivity of TB blood NAAT for TB blood culture cases (population median across all 13 studies, figure 13) was 0.71,
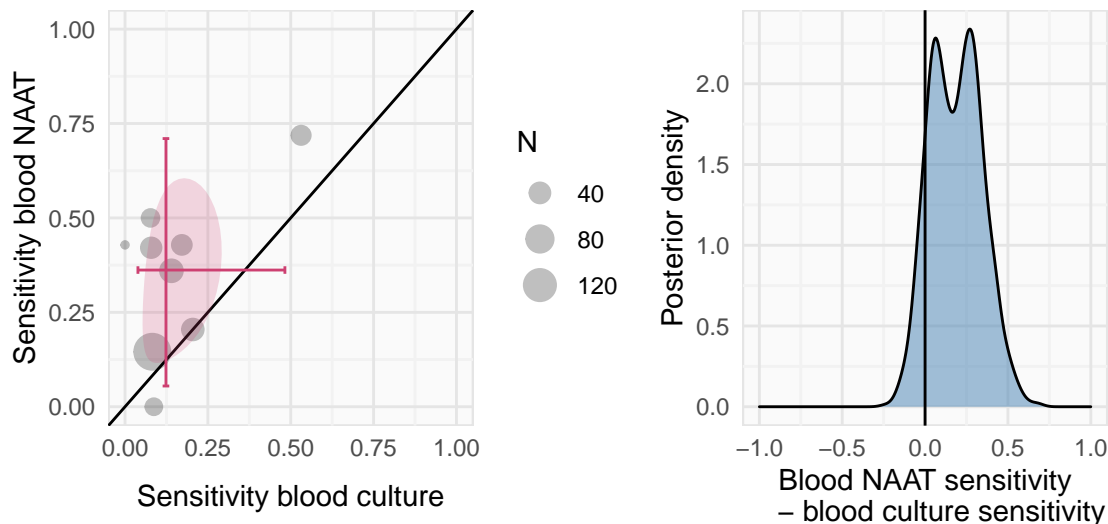
Figure 12: Bivariate mixed-effects regression relative sensitivity of TB blood culture and blood NAAT versus an external reference standard for TB diagnosis (review 2, aim 1). Left panel shows individual studies raw data for sensitivities (grey circles) and the 95%CrI for the median population value (red shaded area) and 90% prediction intervals for a new unobserved study (red lines). Right panel shows distribution of posterior estimates from the model for difference in sensitivity (estimate for sensitivity of NAAT minus sensitivity of culture); most probability was asigned by the model to higher sensitivity of NAAT, to the right of vertical black line.

Table 1: Mixed-effects meta-regression blood NAAT sensitivity for detection of blood culture positive TB: estimated difference in sensitivity of blood NAAT for two levels of study level variables.

effect	comparator	delta sensitivity	95% CrI
low bias	high/moderate bias	-0.03	-0.62 to 0.61
year 2015	year 1995	-0.21	-0.68 to 0.63
commercial_kit	in house kit	-0.42	-0.94 to 0.15
Study size, N=50	Study size, N=5	0.00	-0.02 to 0.02

but with substantial uncertainty for this population estimate (95% CrI 0.41 to 0.92, and 90% prediction interval for a new, unobserved study 0.08 to 0.99) due to heterogeneity and the limited amount of published data (in total, blood TB NAAT results have only been reported for 174 TB blood culture positive patients, with median of 7 patients per study).

This means there is limited power to support meta-regression. More recent studies, studies using commercial NAAT kits, and studies assessed to be at lower risk of bais all had lower reported sesnitivity of NAAT for TB blood culture positive disease but none of these associations reached a >95% posterior probability statistical significance level (table 1).

3.5 Summary review 2

Published data on blood NAAT for TB diagnosis where sensitivity can be related to a concomitant TB blood culture is sparse. This data was reviewed because sensitivity of blood TB NAAT relative to TB blood culture allows a degree of adjustment for disease spectrum which we hypothesised might underlie the extreme heterogeneity described in review 1. However, variance in reported sensitivity of TB blood NAAT was still
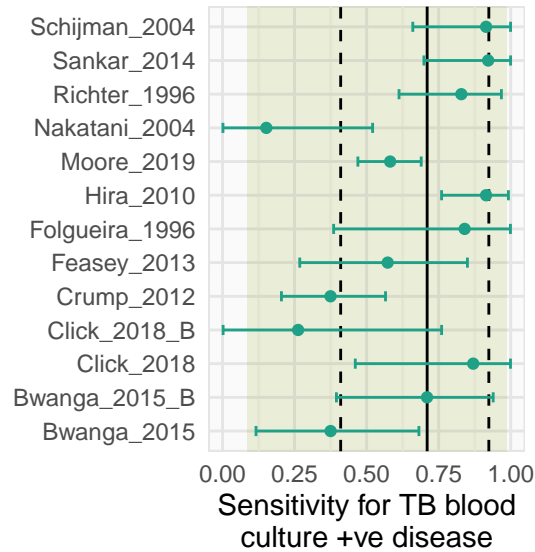
Figure 13: Mixed-effects regression blood NAAT sensitivity for detection of blood culture positive TB (review 2, aim 2). Fit and 95%CrI for individiual studies shown with green dots and whisters; estimated population median and 95%CrI shown with vertical solid and dashed lines respectively; 90% prediction intervals for a new unobserved study indicated by shaded green area.

pronounced relative to TB blood culture and within the strata of patients who were TB blood culture positive. Reported sensitivity of TB blood NAAT was again lower on average in low bias rated studies, studies using commercial NAAT kits, and in more recently reported studies compared to initial reports in 1990s; these associations were not statistically significant, which may relate to the limited amount of data available in review 2.