Diagnostic accuracy of centralized assays for TB detection and detection of resistance to rifampicin and isoniazid: A systematic review and meta-analysis

Mikashmi Kohli[1,2]; Emily MacLean[1,2]; Claudia M. Denkinger[3,4]; Madhukar Pai[1,2]; Samuel G. Schumacher[3]

Supplementary Material

Corresponding author:
Samuel G. Schumacher, MSc, PhD
Deputy Head of TB
FIND, Geneva
Email: samuel.schumacher@finddx.org

Database: Embase <1996 to 2018 Week 26>
Search Strategy:

---

1 (rifampin* or rifampicin* or Isoniazid*).mp. (73472)

2 (MDR TB or MDRTB or RRTB or RR TB or DRTB or DR TB).mp. (4878)

3 exp tuberculosis/ or mycobacterium tuberculosis/ or tuberculosis control/ or rifampicin/ or isoniazid/ (189174)

4 (tubercul* or antitubercul* or tb).mp. (190575)

5 1 or 2 or 3 or 4 (231179)

6 (((Abbott or RealTime* or Real Time*) adj (mtb* or rif* or inh*)) or fluorotype* or bd max* or bdmax* or cobas* taqman*).mp. (1612)

7 *real time polymerase chain reaction/ (10598)

8 ((real time or realtime or rt or direct) and (pcr or polymerase chain reaction)).ti. (18450)

9 6 or 7 or 8 (22380)

10 5 and 9 (574)

11 limit 10 to yr="2009 -Current" (447)

12 limit 11 to dc=20171205-20180628 (17)

13 remove duplicates from 12 (17)

QUADAS-2 Protocol

Domain 1 Patient Selection:
Risk of Bias: Could the selection of patients have introduced bias?

- Signaling question 1: Was a consecutive or random sample of patients or specimens enrolled?
    - We scored ‘yes’ if the study enrolled a consecutive or random sample of eligible patients; ‘no’ if the study selected patients by convenience, and ‘unclear’ if the study did not report the manner of patient selection or this cannot be discerned.

- Signaling question 2: Was a case-control design avoided?
    - We scored ‘yes’ if the study enrolled only patients presumed of drug-resistant TB, including patients with confirmed TB. We scored ‘no’ if the study enrolled patients for whom resistance status was already known, and ‘unclear’ if the study did not report the design or this cannot be discerned.
- Signaling question 3: Did the study avoid inappropriate exclusions?
    - We scored ‘yes’ if no inappropriate exclusions were noted. We scored ‘no’ if studies note specific exclusions. Inappropriate exclusions could potentially occur if patients were excluded based on prior knowledge or testing about them or if the technician does not record performed test results but this was not anticipated for research studies in this review.

Applicability: Are there concerns that the included patients and setting do not match the review question?

We were interested in how the index tests (centralized molecular DST assays) performed in patients presumed of having TB who are evaluated. We judged ‘low concern’ when the specimens included in the study were from the patients with presumptive pulmonary TB and was conducted in high TB and/or high MDR-TB burden country as per the WHO list. We judged ‘high concern’ if the specimens were collected from patients in a low TB and/or MDR-TB burden country. We will judge ‘unclear concern’ if the study included specimens from both high and low TB/MDR-TB burden settings or we could not tell.

Domain 2: Index Test

Risk of Bias: Could the conduct or interpretation of the index test have introduced bias?

- Signaling question 1: Were the index test results interpreted without knowledge of the results of the reference standard?
    - We scored ‘yes’ for all studies because all the centralized molecular DST assay results are automatically generated and the user is provided with printable test results. Thus, there was no room for subjective interpretation of test results.

- Signaling question 2: If a threshold was used, was it prespecified?
    - As the threshold is prespecified in all centralized molecular DST assay in this review, we answered this question "yes" for all studies.

Applicability: Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Variations in test technology, execution, or interpretation may affect estimates of the diagnostic accuracy of a test.

We judged ‘low concern’ if the test was done as per recommendation of the manufacturer for PTB specimens. We judged ‘high concern’ it was stated and/or if additional steps were used for sample preparation and ‘unclear concern’ if we could not tell.

Domain 3: Reference Standard

Risk of Bias: Could the reference standard, its conduct, or its interpretation have introduced bias?

- Signaling question 1: Is the reference standard likely to correctly classify the target condition?
    - For detection of TB, culture is generally considered the best reference standard. We scored ‘yes’ if the studies used MGIT 960 as the reference standard (higher quality reference standard). We scored ‘no’ if the studies used only solid media-based culture (lower quality reference standard) as all these index tests are for centralized settings, we expect the laboratory settings to have liquid culture for detecting TB. LJ culture has lower diagnostic accuracy than liquid culture and would over or under-estimate the diagnostic accuracy of the index test. We scored ‘unclear’ if we could not tell.
    - For detection of rifampicin resistance, culture-based drug susceptibility testing (DST, also called conventional phenotypic method) is considered to be the best reference standard. As we extracted data for studies that used culture-based DST, we will score “ yes” for all studies.

- Signaling question 2: Were the reference standard results interpreted without knowledge of the results of the index test?
    - We scored 'yes' if the reference test provided was culture e.g. MGIT 960 DST where an automated result is generated (except for LJ with confirmation of MTB by a NAAT-based test), if blinding was explicitly stated, or if it was clear that the reference standard was performed at a separate laboratory and/or performed by different people. We will score ‘no’ if the study stated that the reference standard was interpreted with knowledge of the index test result. We scored 'unclear' if this was not stated or answered inadequately.

- Signaling question 3: (Rifampicin resistance) Were the reference standard results interpreted without knowledge of the results of the index test?
    - We added a signaling question for rifampicin resistance detection. We scored "yes" if the reference test provided an automated result (for example, MGIT 960), blinding was explicitly stated, or it was clear that the reference standard was performed at a separate laboratory or performed by different people, or both. We scored "no" if the study stated that the reference standard result was interpreted with knowledge of the index test result. We scored "unclear" if we could not tell.

Applicability: Are there concerns that the target condition as defined by the reference standard does not match the question?
We judged applicability to be of ‘low concern’ for all studies.

Domain 4: Flow and Timing
Risk of Bias: Could the patient flow have introduced bias?

- Signaling question 1: Was there an appropriate interval between the index test and reference standard?
    - We scored ‘‘yes’ if the tests were paired or separated by less than 48 hours after treatment initiation. We scored ‘no’ if the reference and index tests were not performed on paired specimens or were separated by more than a week. We scored ‘unclear’ if this was not stated in the paper or answered inadequately. In the majority of included studies, we expected specimens for index tests and culture to be obtained at the same time (i.e. to be performed on paired specimens for the majority of studies), when patients are presumed of having TB or MDR-TB.

- Signaling question 2: Did all patients receive the same reference standard?
    - For the diagnosis of TB, we scored this question "yes" if all participants in the study or a subset of participants in the study (for whom we will extract data) received the acceptable reference standard (solid culture, liquid culture, or both), which we specified as a criterion for inclusion in the review. However, we acknowledge that it is possible that some specimens could undergo solid culture and others liquid culture as the reference standard. This variation was recorded.
    - For rifampicin resistance detection, we scored "yes" if all participants received the same reference standard (either culture-based DST or MTBDR*plus*), "no" if not all participants received the same reference standard, and "unclear" if we could not tell.

- Signaling question 3: Were all patients included in the analysis?
- The answer to this question was determined by comparing the number of patients enrolled with the number of patients included in the two-by-two tables. We noted if authors record the number of indeterminate results. We scored ‘yes’ if the number of participants enrolled was clearly stated and corresponded to the number presented in the analysis or if exclusions were adequately described. We scored 'no’ if there were participants missing or excluded from the analysis and there was no explanation given; and 'unclear ' if not enough information was given to assess whether participants were excluded from the analysis

QUADAS-2 summaries – Risk of bias and applicability concerns

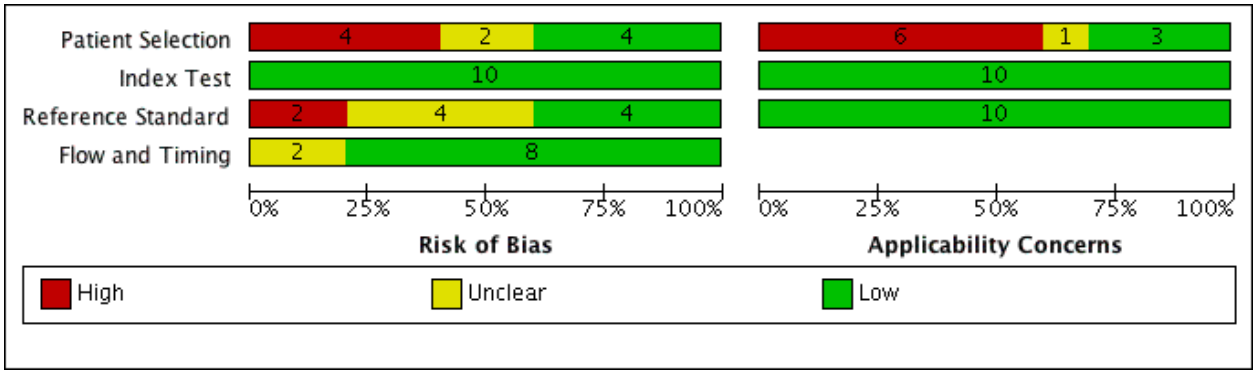

Figure S1. Risk of Bias and Applicability Concerns summary about each QUADAS-2 domain for Abbott RealTime MTB assay

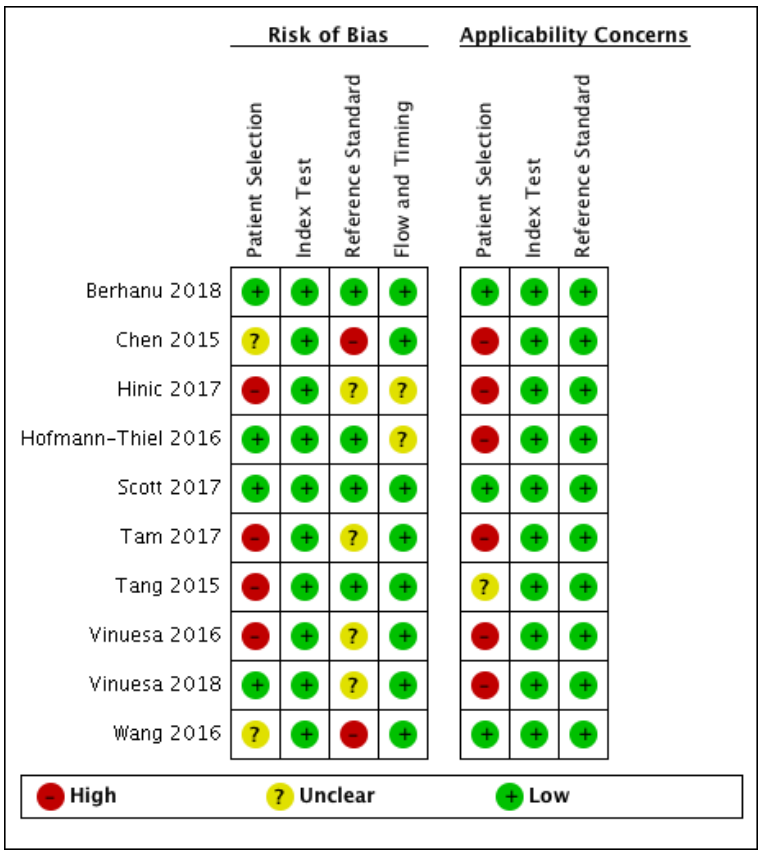

Figure S2. Risk of Bias and Applicability Concerns summary for QUADAS-2 domains in each study evaluating Abbott RealTime MTB assay

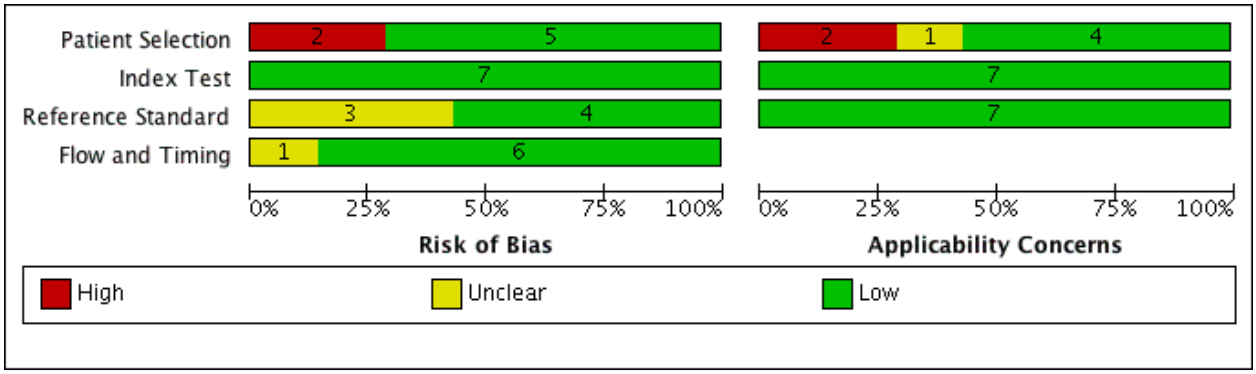

Figure S3. Risk of Bias and Applicability Concerns summary about each QUADAS-2 domain presented as percentages for Abbott RealTime MTB RIF/INH assay

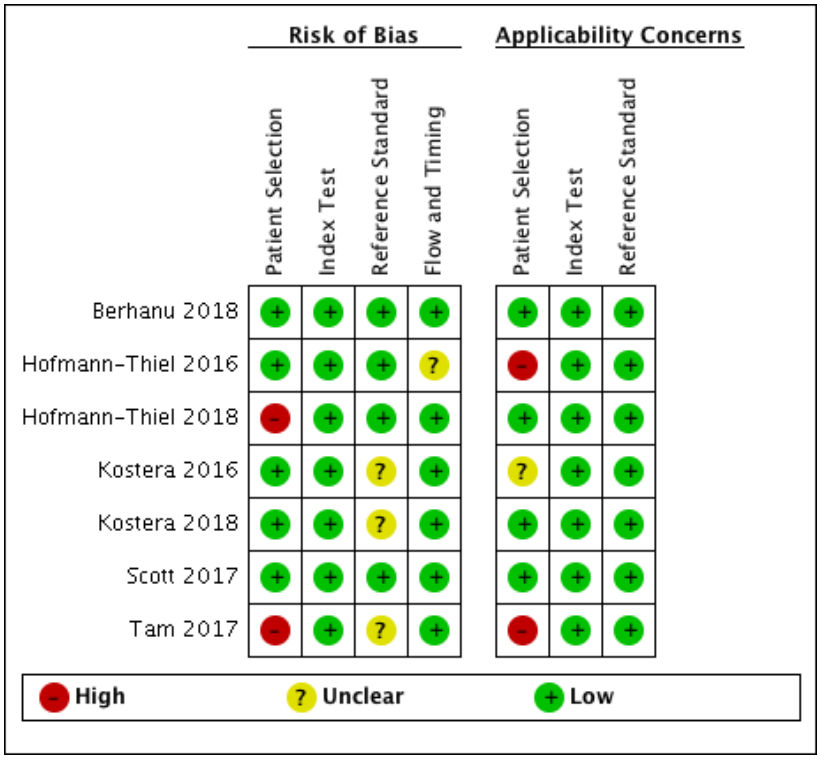

Figure S4. Risk of Bias and Applicability Concerns summary for QUADAS-2 domains in each study evaluating Abbott RealTime MTB RIF/INH assay

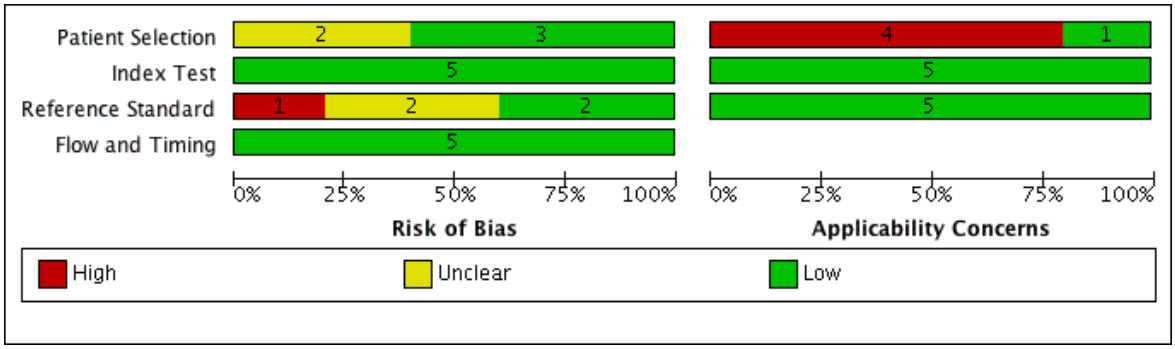

Figure S5. Risk of Bias and Applicability Concerns summary about each QUADAS-2 domain presented as percentages for FluoroType MTB assay

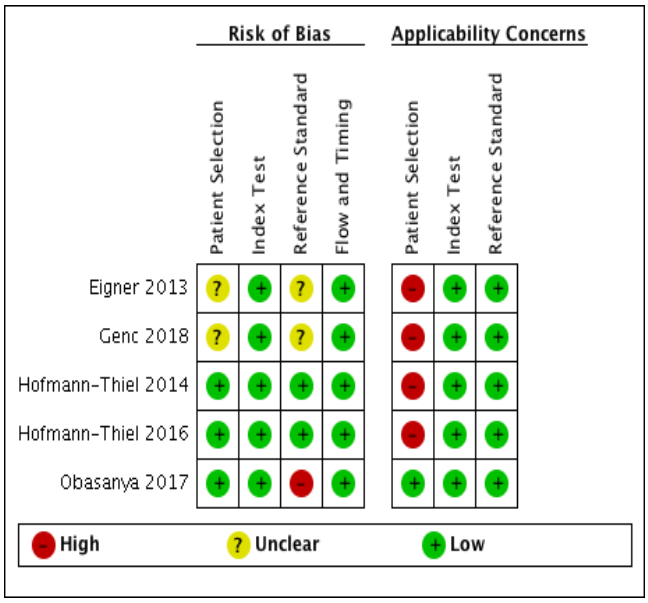

Figure S6: Risk of Bias and Applicability Concerns summary for QUADAS-2 domains in each study evaluating FluoroType MTB assay

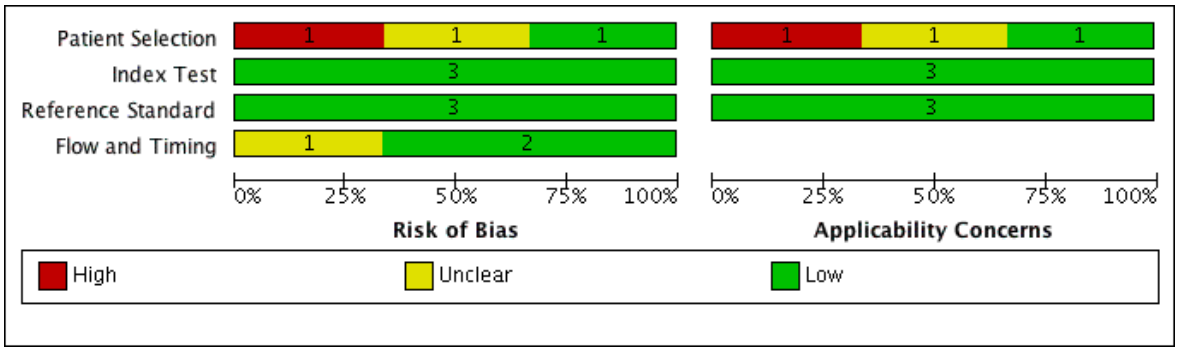

Figure S7. Risk of Bias and Applicability Concerns summary about each QUADAS-2 domain presented as percentages for FluoroType MTBDR assay

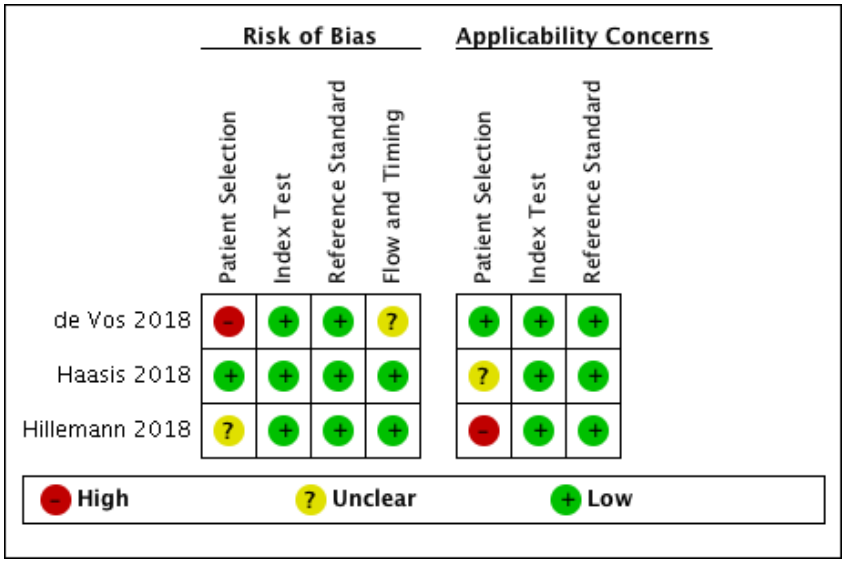

Figure S8. Risk of Bias and Applicability Concerns summary for QUADAS-2 domains in each study evaluating FluoroType MTBDR assay

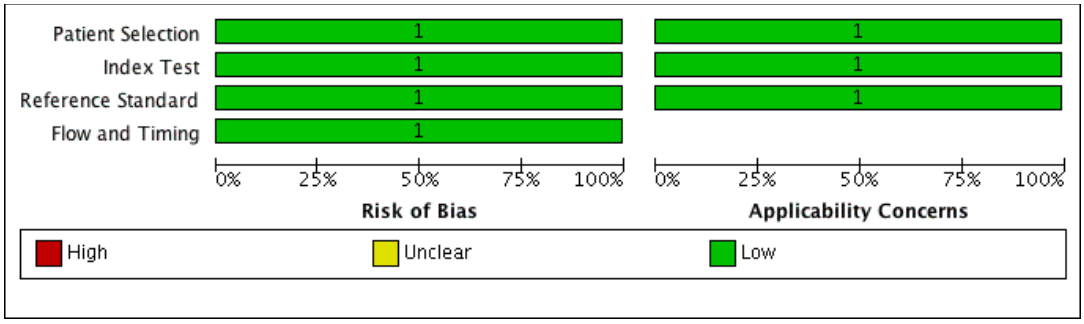

Figure S9: Risk of Bias and Applicability Concerns summary about each QUADAS-2 domain presented as percentages for BD Max MDR-TB assay

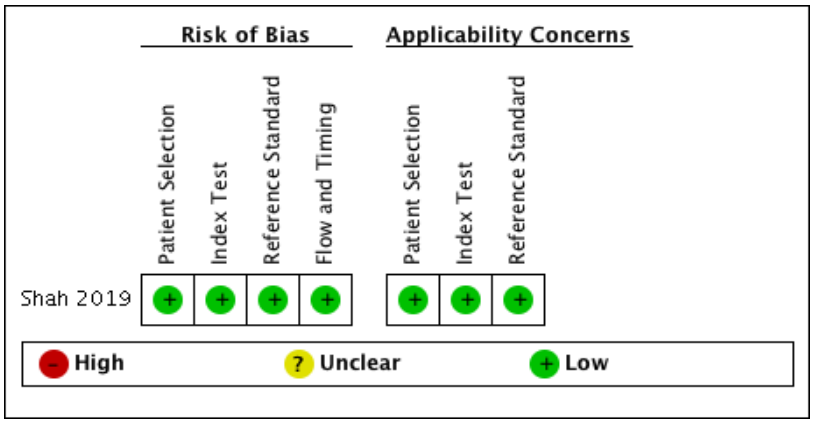

Figure S10: Risk of Bias and Applicability Concerns summary for QUADAS-2 domains in each study evaluating BD Max MDR-TB assay

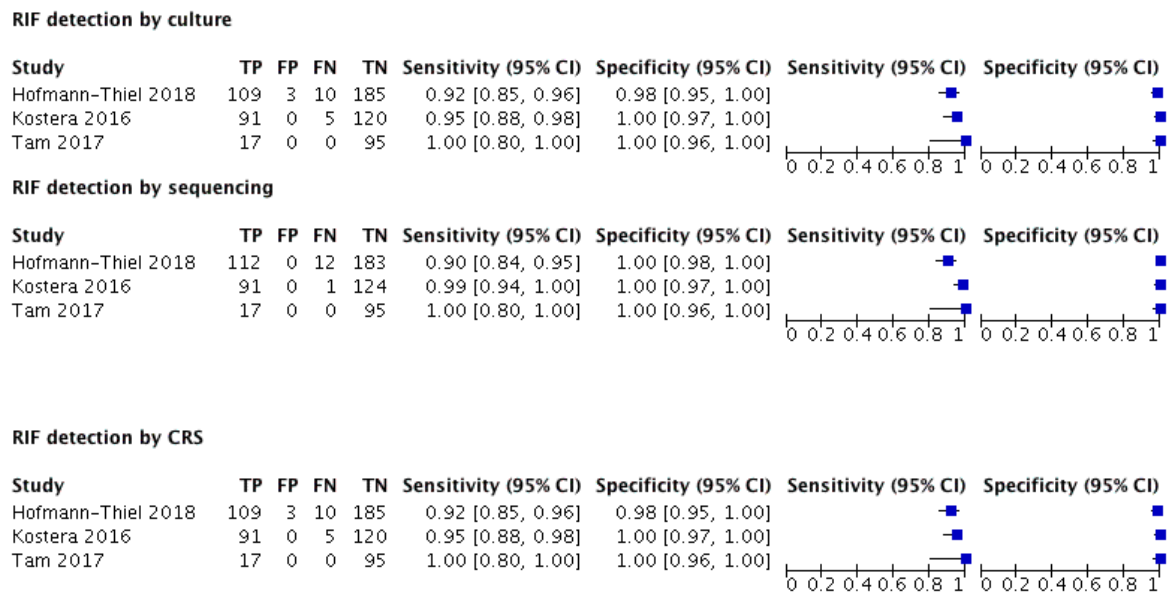

Figure S11: Forest plots for rifampicin resistance detection by Abbott RIF/INH assay using phenotypic DST, sequencing and composite reference standard

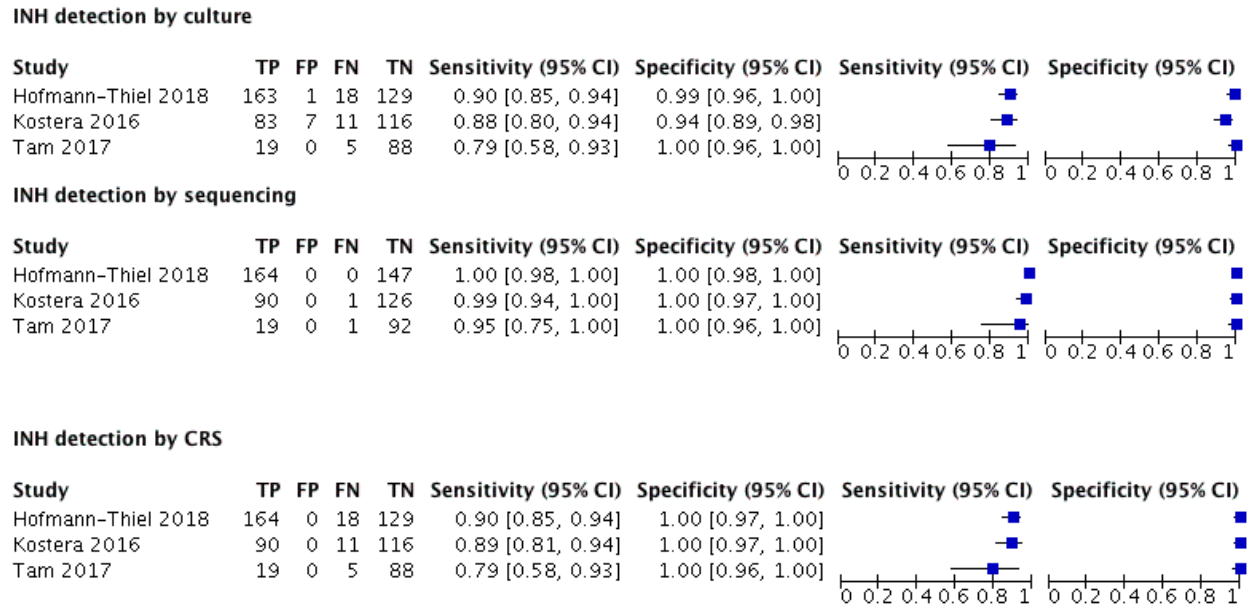

Figure S12: Forest plots for isoniazid resistance detection by Abbott RIF/INH assay using phenotypic DST, sequencing and composite reference standard