HHS Public Access

Author manuscript

Lancet Respir Med. Author manuscript; available in PMC 2016 February 17.

Published in final edited form as:

Lancet Respir Med. 2015 June ; 3(6): 451–461. doi:10.1016/S2213-2600(15)00095-8.

Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis

Anne K Detjen, MD, **Andrew R DiNardo, MD**, **Jacinta Leyden**, **Karen R Steingart, MD**, **Dick Menzies, MD**, **Ian Schiller, MSc**, **Nandini Dendukuri, PhD**, and **Anna M Mandalakas, MD**

The International Union Against Tuberculosis And Lung Disease (The Union), New York, NY, USA (A K Detjen MD); Infectious Diseases, Department of Internal Medicine, Baylor College of Medicine, Houston, TX, USA (A R DiNardo MD); Department of Bioengineering, Rice University, Houston, TX, USA (J Leyden); Cochrane Infectious Diseases Group, Liverpool School of Tropical Medicine, Liverpool, UK (K R Steingart MD); Respiratory and Epidemiology Clinical Research Unit, Montreal Chest Institute, McGill University, Montreal, QC, Canada (D Menzies MD); Division of Clinical Epidemiology, McGill University Health Centre, Montreal, QC, Canada (I Schiller MSc, N Dendukuri PhD); The Global Tuberculosis Program, Texas Children's Hospital, Houston, TX, USA (A R DiNardo, J Leyden, A M Mandalakas MD); and Retrovirology and Global Health, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA (A M Mandalakas)

Summary

Background—Microbiological confirmation of childhood tuberculosis is rare because of the difficulty of collection of specimens, low sensitivity of smear microscopy, and poor access to culture. We aimed to establish summary estimates for sensitivity and specificity of of the Xpert MTB/RIF assay compared with microscopy in the diagnosis of pulmonary tuberculosis in children.

Methods—We searched for studies published up to Jan 6, 2015, that used Xpert in any setting in children with and without HIV infection. We systematically reviewed studies that compared the diagnostic accuracy of Xpert MTB/RIF (Xpert) with microscopy for detection of pulmonary tuberculosis and rifampicin resistance in children younger than 16 years against two reference standards—culture results and culture-negative children who were started on anti-tuberculosis therapy. We did meta-analyses using a bivariate random-effects model.

Findings—We identified 15 studies including 4768 respiratory specimens in 3640 children investigated for pulmonary tuberculosis. Culture tests were positive for tuberculosis in 12% (420 of 3640) of all children assessed and Xpert was positive in 11% (406 of 3640). Compared with culture, the pooled sensitivities and specificities of Xpert for tuberculosis detection were 62% (95% credible interval 51–73) and 98% (97–99), respectively, with use of expectorated or induced

Correspondence to: Dr Anne Detjen, The International Union Against Tuberculosis and Lung Disease, New York, NY 10006, USA, adetjen@theunion.org.

Contributors

AKD, ARD, KRS, DM, ND, and AMM worked on protocol development. AKD, ARD, and JL collected data. AKD, ARD, KRS, DM, IS, ND, and AMM reviewed, analysed, and interpreted data. All authors contributed to manuscript preparation.

Declaration of interests

We declare no competing interests.

sputum samples and 66% (51–81) and 98% (96–99), respectively, with use of samples from gastric lavage. Xpert sensitivity was 36–44% higher than was sensitivity for microscopy. Xpert sensitivity in culture-negative children started on antituberculosis therapy was 2% (1–3) for expectorated or induced sputum. Xpert's pooled sensitivity and specificity to detect rifampicin resistance was 86% (95% credible interval 53–98) and 98% (94–100), respectively.

Interpretation—Compared with microscopy, Xpert offers better sensitivity for the diagnosis of pulmonary tuberculosis in children and its scale-up will improve access to tuberculosis diagnostics for children. Although Xpert helps to provide rapid confirmation of disease, its sensitivity remains suboptimum compared with culture tests. A negative Xpert result does not rule out tuberculosis. Good clinical acumen is still needed to decide when to start antituberculosis therapy and continued research for better diagnostics is crucial.

Funding—WHO, Global TB Program of Texas Children's Hospital.

Introduction

The 530 000–999 792 cases of tuberculosis every year in children account for at least 6% of the global burden of the disease.[1–3] These numbers underestimate the burden of childhood tuberculosis, which is higher due to difficulty in diagnosis of childhood tuberculosis, emphasising the need for improved diagnostics. Smear microscopy remains the most used and widely available tuberculosis diagnostic method in low-income and middle-income countries, particularly in peripheral settings that do not have access to higher-level laboratories. Microscopy is of little value in children, who typically have paucibacillary tuberculosis and have difficulty producing sputum. In children, culture methods have a greater, yet highly variable, sensitivity. For these reasons, microbiological confirmation of childhood tuberculosis is rare and clinical diagnosis relies on a combination of signs, symptoms, radiological findings, and identification of a tuberculosis contact.[4]

The ongoing rollout of Xpert MTB/RIF (Xpert; Cepheid, Sunnyvale CA, USA) in low-income and middle-income countries offers an opportunity for investigators to provide access to diagnosis for children beyond smear microscopy. We did a systematic review and meta-analysis on the use of Xpert in children, which informed the recent WHO update of guidelines on the use of Xpert in adults and children. This Article includes results updated up to December, 2014.[5] We aimed to establish summary estimates for the accuracy of Xpert in diagnosis of pulmonary tuberculosis and rifampicin resistance in children, with the secondary objective of investigation of heterogeneity of comparison studies in relation to age, smear-test status**,** HIV-status, and an inpatient versus outpatient setting.

Methods

Study inclusion

We searched Medline (through PubMed and Ovid) and Scopus for published work without language and date restrictions. Our last search was done on Jan 6, 2015. We searched through reference lists of included studies and review articles for additional studies. We contacted authors from published studies and a broad network of researchers of childhood tuberculosis to identify continuing and unpublished studies.

We included studies assessing Xpert for the diagnosis of pulmonary tuberculosis in HIV-infected and HIV-uninfected children aged 0–15 years with presumed pulmonary tuberculosis. Studies used Xpert on routine respiratory specimens such as expectorated or induced sputum, gastric lavage, and nasopharyngeal aspirates, and included more than five participants. We included published articles, articles in press, and unpublished studies when authors agreed to share methods and results. We included cross-sectional studies, cohort studies, and randomised controlled trials from settings with a high, moderate, and low tuberculosis burden if they compared Xpert to an acceptable reference standard. We excluded case-control studies, case reports, and studies only presented as an abstract.

Xpert MTB/RIF was the index test. We considered one result per index test per child; ideally, this corresponded to the first specimen provided. Smear microscopy was the comparator test for studies that reported a direct comparison of smear and Xpert against a reference standard. Two reference standards were selected for pulmonary tuberculosis: culture and clinical diagnosis. We considered studies that used one or more solid media or commercial liquid culture per child or both, including studies that assessed several specimen sources (eg, sputum and gastric aspirate). We assigned culture-positive children to the group named “confirmed tuberculosis”. Recognising the limitations of culture, we accepted a second reference standard (clinical tuberculosis) that was applied only in culture-negative children. Children were categorised as positive for clinical tuberculosis if a provider started antituberculosis therapy without knowing the results of Xpert testing. Children assigned to the group named “not clinical tuberculosis” either had an alternative diagnosis or did not start antituberculosis therapy and improved or did not worsen after at least 1 month. Most studies did not provide data to enable the use of existing consensus definitions for clinical tuberculosis in children.[6] The reference standard for rifampicin resistance was established by conventional phenotypic drug susceptibility testing or line-probe assays.[7]

Two authors (ARD, AKD) independently reviewed titles and abstracts, followed by full-text review of selected studies. Studies categorised as not meeting inclusion criteria by both authors were excluded; consensus was achieved from a third reviewer (AMM) if the authors disagreed (appendix). The two authors independently extracted data with use of a form adapted from a recent Cochrane review.[8] We contacted study authors for missing data, clarifications, and to reclassify children according to the clinical tuberculosis reference standard. All data were entered into Microsoft Excel version 14.4.1 and verified independently by the same two authors. We assessed study quality with Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2).[9] We did not assess publication bias because these methods are not applicable for studies of diagnostic accuracy.[10]

Statistical analysis

We did descriptive analyses with Excel and Review Manager 5 (RevMan) version 5.2 (The Cochrane Collaboration, Copenhagen, Denmark) to summarise study characteristics and quality based on QUADAS-2. Meta-analyses of estimated pooled sensitivity and specificity of Xpert were done seperately for tuberculosis detection and rifampicin resistance using a bivariate random-effects model.[8,11] This approach allowed calculation of pooled estimates while minimising potential sources of variation caused by imprecision of sensitivity and

specificity estimates within individual studies, correlation between sensitivity and specificity across studies, and variation in sensitivity and specificity between studies.[11] We used a Bayesian approach to estimate all meta-analysis models. In the Bayesian approach a prior distribution (which summarises information available on the parameters of the meta-analysis model from external sources) is combined with information from the included studies.[8] We used non-informative prior distributions to allow the observed data to dominate the results, and reported pooled estimates together with a 95% credible interval. We also reported a prediction interval, which would capture the uncertainty around the sensitivity or specificity estimates that could be expected in a future study. If there was no heterogeneity between studies, then the credible interval for the pooled estimate would be the same as the prediction interval. By contrast, if there was considerable heterogeneity between studies, then the prediction interval would be much wider than would the credible interval. To assess the heterogeneity of the accuracy of Xpert with respect to culture, we refitted the meta-analysis model within groups defined by smear-status or HIV-status or both. We fitted meta-analyses and meta-regression models in WinBUGS Version 1.4.3 (MRC Biostatistics Unit, London, UK).

Role of the funding source

Funders had no role in the study design, collection, analysis or interpretation of the data, or writing of the report. All authors had access to the raw data. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

Results

We identified 115 published studies, and, after title and abstract review, assessed 27 full-text articles (figure 1). We selected 15 studies for inclusion (figure 1).[12–25] Five studies were done in low-income countries, two in lower middle-income, six in upper middle-income, and two in high-income countries (table 1).[26] 12 studies were done at tertiary, university, or research facilities. Eight studies included inpatients only whereas four included both inpatients and outpatients; two laboratory-based studies collected minimal clinical information. One study[24] included outpatients from a primary care setting, but specimens were processed at a university hospital reference laboratory. With use of QUADAS, we considered most studies to have low risk for selection bias because children were recruited in a consecutive manner.

Studies included 3640 participants contributing to 4768 specimens assessed (median 265 specimens per study, range 20–948 specimens; appendix). Investigators of some studies collected the same specimen type from all children, whereas others collected different types of specimens from different subgroups of children—eg, expectorated sputum in older children and induced sputum or gastric fluid (gastric lavage) in younger children. Investigators of four studies[18,20,23] collected different types of specimen in each child (Walters E, Desmond Tutu TB Centre, South Africa, personal communication; appendix). Researchers had heterogeneous approaches to inclusion criteria and definition of presumed tuberculosis, the number of specimens collected and cultures needed per child to confirm

tuberculosis, and the definition of clinical tuberculosis in culture-negative children (table 2; appendix).

Of all studies and specimen types included in our findings, 12% (range 1–53%; 420 of 3640) of children had culture-confirmed tuberculosis (table 2). 80% of investigators (12 of 15) took multiple cultures (up to six) for each child to establish positivity. On average, 11% (406 of 3640; range 1–45%) of all children were positive with Xpert. 44% (688 of 1576) of culture-negative children in seven studies were started on empirical antituberculosis therapy; 2% (10 of 688) of them were Xpert positive.

The sensitivity of Xpert varied broadly across studies and specimen types, whereas specificities ranged from 93% to 100% (figure 2). In our meta-analysis, data for expectorated sputum and induced sputum were combined because these specimen sources are clinically similar; gastric lavage data were analysed separately. In a meta-regression model comparing the pooled estimates of different specimen types, expectorated sputum and induced sputum and gastric lavage did not significantly differ. Two studies assessed Xpert on nasopharyngeal aspirate samples and collected induced sputum from the same cohort of children.[18,23] Because we could not adjust for within-person correlation, we excluded the nasopharyngeal aspirates data from the meta-analysis (figure 2 shows individual study data). Table 3 shows summary estimates for Xpert against the reference standard of culture and the reference standard of clinical tuberculosis for tuberculosis detection. Sensitivity analysis that excluded the unpublished study did not change our findings (data not shown). Prediction intervals were wider than the pooled credible intervals for all sensitivity estimates (table 3), indicating much heterogeneity between studies. By contrast, estimates of specificity against both reference standards were all 98% or greater with narrow credible intervals.

Sensitivity estimates vary for smear microscopy compared with the reference standard of culture for the same studies and specimen types as assessed for Xpert (0–60%), whereas specificity was consistently high (>93%; figure 3). By comparison with smear microscopy, Xpert was 36% more sensitive on expectorated or induced sputum samples and 44% more sensitive on gastric lavage samples (table 3).

In a subset of studies with adequate data, we examined potential causes of heterogeneity for findings with Xpert compared with the reference standard of culture. Meta-analysis stratified by smear status shows a difference in sensitivity between smear-positive and smear-negative children (figure 4; appendix). The specificity of Xpert was greater than 93% in all studies except one in which specimens were stored at −20°C for at least 3 days; this study reported a specificity of 70% (95% credible interval 63–76).[24] We removed this particular study from the meta-analysis, which resulted in a pooled sensitivity of 62% (44–80), and specificity of 99% (97–99) for Xpert. The pooled and predicted sensitivity for children aged 0–4 years was 53% (37–67) for Xpert for expectorated or induced sputum and 57% (38–75) for Xpert for gastric lavage, compared with a sensitivity of 76% (61–87) for children aged 5–15 years in sputum samples. A scarcity of data precluded assessment of Xpert in children aged 5–15 years from gastric lavage. Specificity was 98% or greater for all groups assessed with relatively narrow pooled and predicted credible intervals. Recognising the association between smear status and Xpert sensitivity, we completed an age-stratified analysis (table

4). Xpert sensitivity was highest in smear-positive children in both age strata, and lowest among smear-negative children aged 0–4 years (table 4).

Pooled and predicted sensitivity was higher for HIV-positive children than for HIV-negative children when Xpert was used on expectorated and induced sputum samples (55% and 70%, respectively) with wide and overlapping pooled and predicted credible intervals (data not shown). Data were insufficient to do a meta-analysis of Xpert performance by HIV status with samples from gastric lavage. HIV-stratified analysis comparing Xpert for smear status showed high sensitivity among smear-positive HIV-positive (97%) and HIV-negative (94%) children with slight and overlapping credible intervals (table 4). Xpert sensitivity was lowest among smear-tnegative, HIV-uninfected children (44%). Credible intervals were wide and overlapping; predicted intervals significantly broadened, showing the heterogeneity of the studies (table 4).

In a meta-regression model simultaneously controlling for smear status and HIV status for Xpert diagnosis of expectorated or induced sputum samples, the odds of test positivity was four times greater in smear-positive children than in smear-negative children and was not associated with HIV status (appendix). Similar to the stratified analysis, pooled sensitivities calculated by the model had the highest sensitivity in HIV-positive, smear-positive children.

The pooled and predicted sensitivity of studies that included mainly inpatients was 70% (95% CI 57–82) compared with a sensitivity of 48% (31–65) in studies that included mainly outpatients (appendix). Similar to age and HIV, these differences were mainly due to a greater proportion of smear-negative individuals in outpatient settings.

We used four studies to assess the incremental yield of Xpert (appendix).[17,18,26,27] Using the total number of tuberculosis cases (culture-confirmed, plus cases meeting an author-defined diagnosis of clinical tuberculosis) as a denominator, the incremental yield of Xpert on a second specimen ranged from 8·3% to 17·5%, and from 0% to 12·5% on a third specimen. The incremental yield of culture (Löwenstein-Jensen or Mycobacteria Growth Indicator Tube values) ranged from 14·3% to 21·9% on the second specimen. One study[16] showed the incremental yield for all different diagnostic assays done, ranging from 3·1% for smear, 9·5% for Löwenstein-Jensen, 12·5% for Xpert to 21·9% for Mycobacteria Growth Indicator Tube on up to three specimens. One study[18] reported an increased yield when specimens were collected on different days.

Six studies provided data for Xpert used for rifampicin-resistance testing (table 1). Of 240 children, 11 (4·6%) were resistant on culture or line probe assay and ten (4·2%) on Xpert. Four children identified as drug-sensitive by culture had indeterminate Xpert rifampicin-resistance results. A meta-analysis of three studies (one being Walters' unpublished study),[12,18] collectively including 176 participants, showed a pooled sensitivity of 86% (95% credible interval 53–98) and a pooled specificity of 98% (94–100) for rifampicin-resistance.

Discussion

Our results show that the Xpert assay can diagnose tuberculosis equally well in different respiratory specimens and is better than smear microscopy, but that overall sensitivity remains suboptimum compared with culture.[5] The high variation of sensitivities between studies and specimen types reported for Xpert correlates with the high variation in yield of culture between studies and heterogeneity in study populations and settings. The wide pooled and predicted 95% credibility intervals emphasise the need for continued research in well defined populations to better understand the potential role of Xpert in routine care. Specificities for Xpert against culture in all analyses were consistently high with narrow 95% credibility intervals.

Our data show the association between Xpert accuracy and smear status. Hence, sensitivity with Xpert could be lower in children with clinically diagnosed disease that might be paucibacillary. Smear status also seems to confound estimates for the other subgroups assessed: age, HIV status, and study setting. Xpert sensitivity estimates are highest in smear-positive older children who might present in a similar manner to adults with tuberculosis and are lowest in smear-negative, young children.

Xpert identified additional tuberculosis cases compared with microscopy. However, given the difficulty in obtaining adequate paediatric specimens, in many clinical settings, diagnostic testing is rarely attempted in children. Increased access to Xpert and improved diagnosis might motivate health-care workers to obtain specimens from children; there is a need for training to optimise specimen collection methods, which might in turn increase diagnostic yield. This simple behavioural shift could substantially improve tuberculosis diagnosis in children. Future assessments should include studies in routine clinical settings to address alternative specimen collection methods and specimen type.

Mycobacterial culture, the gold standard for the diagnosis of tuberculosis and the main reference standard applied in studies used in this Article, is imperfect in children. The yield of culture in childhood tuberculosis ranges from 20% to 70% depending on factors such as age, disease severity, and type and quality of the specimen, and culture method used.[27,28] All these factors probably also interact with smear status. Disease severity might be a proxy for bacillary load as shown in one study assessing pulmonary tuberculosis in children, in which the yield of culture ranged from 35% in uncomplicated lymph-node disease and 82% in complicated parenchymal disease, to 93% in disseminated and 100% in adult-type disease.[24]

The diagnostic yield of culture can be improved by taking cultures from several specimens. 11 of 13 studies that we analysed did several cultures to confirm tuberculosis, with some taking as many as six cultures per child. Hence, for our reference standard of culture, we accepted the best performance of culture to ascertain the most accurate diagnosis. As a consequence, our approach underestimated Xpert sensitivity against culture because data from only one Xpert per child were included. This underestimation can be seen in the unpublished study by Walters and colleagues that used both analytical approaches: one Xpert compared with one culture (sensitivity of Xpert 64%) and one Xpert compared with

yield from up to four cultures (sensitivity 47%; Walters E, Desmond Tutu TB Centre, personal communication). Several studies have shown the incremental yield of Xpert.[14,16,18,23] However, the benefit has to be weighed against the cost of additional tests as well as additional transportation costs for patients.

Due to poor availability and capacity and the diagnostic uncertainty of culture in children, the diagnosis of childhood tuberculosis often relies upon clinical criteria with treatment often initiated empirically. We therefore compared Xpert with culture-negative children started on antituberculosis therapy. This pragmatic reference standard supported pooled analysis of studies that use different definitions of clinical tuberculosis, but only applied them to culture-negative children. The decision to treat in the studies we analysed was based on clinical grounds, without knowledge of Xpert results. The data show high rates of empirical treatment in children. The absence of an optimal gold standard precludes determination of proportion of true cases of tuberculosis. Nevertheless, all of these children received treatment and contribute to the costs of tuberculosis care including the potential cost of overtreatment and side-effects.

Ideally, studies would apply a standard set of criteria to define clinical tuberculosis in research that are applied to all children, not only to those who are culture-negative, enabling better comparison between studies and reference standards.[6] Our estimates of Xpert against this clinical reference standard are very low, indicating that very few culture-negative, clinically defined cases of tuberculosis can be detected by Xpert. This is expected because Xpert, like culture, requires a minimum amount of mycobacteria to be present in the specimen to detect the infection (limit of detection for the current version of Xpert is 131 colony forming units per mL, culture is 10–100 colony forming units per mL).[29]

The studies included in this Article were mainly done at higher levels of care and in inpatients, probably due to availability of culture at the tertiary care level enabling researchers to compare Xpert with culture. Particularly in children, inpatients and outpatients differ. Referral bias has probably affected our findings because study participants, compared with a community-based sample, will probably have more severe and complicated forms of disease with higher bacillary loads. There would therefore be a higher likelihood of smear-culture positivity.[28] This hypothesis is supported by our stratified analysis of studies that included mainly inpatients versus those that included mainly outpatients. Even though smear status affected the results, the findings further show potential differences between inpatient and outpatient populations. Our observation has relevance to clinical care because Xpert has been strategically rolled out as a near point-of-care test and is being used increasingly in outpatient settings.

The inpatient paediatric population might also have a higher likelihood of advanced HIV infection, especially in high-burden settings. An initially unexpected finding was the improved sensitivity among HIV-infected children compared with HIV-uninfected children, a result shown by a meta-regression model to be associated with smear status rather than HIV status. We hypothesise that HIV-infected, largely hospitalised children included in our analysis had greater disease severity, associated with a higher likelihood of smear positivity.[30]

The included studies assessed very different populations of children, with some having more rigorous definitions of presumed tuberculosis and consideration of exposure. Studies using broader definitions included laboratory-based studies with no clinical inclusion criteria and hospital-based studies including populations such as severely malnourished children, in which tuberculosis is one of the differential diagnoses.[19,20] The yield of Xpert in these groups was very low and shows the association between disease prevalence and test accuracy.

Other important outcomes are needed to show the true effect of Xpert on health systems and patients in routine settings. One study reported time to treatment initiation (median 8·5 days for culture-positive children, 17 days for culture-negative children), postulating that 25% of Xpert-positive children could have received antituberculosis treatment at least 31 days earlier, whereas 50% could have received treatment 6 or more days earlier compared with use if culture was used.[16] More studies would ideally reflect routine-care pathways similar to that of one large public health assessment of Xpert in India showing a doubling of confirmed tuberculosis cases among children.[31] Studies are needed to assess how the implementation of Xpert changes clinical diagnosis, empirical treatment, and outcomes for children. Important factors that affect time from the collection of a specimen to the initiation of antituberculosis treatment are related to specimen transport, transmission of results, etc. They are independent of the type of test done and emphasise the need to optimise diagnostic pathways so that any test can have optimum effect. Studies of adults have reported shortcomings and delays in health systems and diagnostic pathways that affect the impact of Xpert on timely treatment initiation and patient outcomes.[32–34] Costs for health systems, as well as for patients, have to be taken into account during consideration of the use of a new diagnostic, but no data were provided in the studies included. Additional data are needed to improve use of Xpert (panel).[21,27,28,35,36]

Panel

Research priorities for Xpert and tuberculosis diagnostics in children

1 Implementation research

How would results with Xpert differ if done at, or close to, point of care (for example, in clinics) as compared with in hospital laboratories?

2 Disease severity

How do results with Xpert differ in children with different stages of disease severity, from non-severe to very severe or disseminated?

3 Specimen collection and preparation

Are there ways to improve yield through improvements in specimen collection or preparation?

What is the role of other respiratory and non-respiratory specimens (eg, stool, urine, cerebrospinal fluid)?

4 Integration

What is the role of Xpert in non-traditional tuberculosis settings (eg, HIV clinics, malnutrition units)?

5 Routine programme data

What are the challenges of integrating Xpert into the health system?

6 Important outcomes for children

How does Xpert affect important outcomes for patients (eg, time to diagnosis, time to treatment, disease outcomes, health-system cost, and cost for families)?

7 Beyond Xpert

What lessons have been learnt from Xpert to inform optimum characteristics of tuberculosis diagnostic tests for children?

8 Empirical treatment

Will the rollout of Xpert affect rates of empirical antituberculosis treatment initiation?

9 New diagnostics and improved gold standards

What is the true incidence and prevalence of tuberculosis in children?

Any research study should apply well-defined and transparent definitions for the certainty of diagnosis (eg, confirmed tuberculosis, clinical tuberculosis, and not tuberculosis).

Our findings in this Article are based on comprehensive searching, strict inclusion criteria, and standardised data extraction. Its main limitations are the low number of studies and participants included, as well as heterogeneous methodological approaches used by individual studies. Although our data show that Xpert is equally effective in gastric lavage and expectorated and induced sputum, age might have confounded our results because gastric lavage tends to be done in younger children than does expectorated sputum. We did not include two studies that used Xpert on specimens obtained from bronchoalveolar lavage showing sensitivities of 53% and 78%, since this technique is more invasive and respiratory specimens are not frequently available.[37,38] Due to the small sample size we were unable to evaluate the prevalence of false-positive Xpert results for rifampicin resistance at different pretest probabilities. Furthermore, most investigators did Xpert at higher levels of care and among inpatients, limiting the generalisability of our findings for other settings. Culture is regarded as the best available reference standard for active tuberculosis in children but it has major limitations as discussed above. The reference standard clinical tuberculosis was applied to culture-negative children only. Although this approach mimics clinical practice, it is methodologically flawed. Ideally, studies would apply each reference standard to all included children. We assessed Xpert for the diagnosis of pulmonary tuberculosis, yet up to 25% of tuberculosis cases in children are extrapulmonary. Pooled sensitivities of Xpert using lymph-node tissue and cerebrospinal fluid are between 80% and 83%, but are only 46% using pleural fluid.[39]

The ongoing rollout of Xpert in low-income and middle-income settings should increase access to much needed diagnostics for tuberculosis and rifampicin resistance in children.

The effect of Xpert can be optimised if its implementation is complemented by efforts to strengthen health systems including improved specimen collection, linkage of specimens to diagnostics, and timely reporting of results. The ability to detect drug-resistant tuberculosis mandates the availability of multidrug-resistant treatment for children. However, the suboptimum sensitivity of Xpert for diagnosis of tuberculosis in children serves as a reminder that many children might need empirical antituberculosis therapy, despite negative Xpert and culture results. Disease severity might further change the accuracy of Xpert, further affecting the interpretation of Xpert results among different patient populations (eg, HIV-infected children, inpatient *vs* outpatient). More sensitive and non-sputum-based diagnostics for paediatric tuberculosis are still needed. In the interim, resource allocation should support training that optimises clinical diagnosis.

Acknowledgments

We received funding from WHO to write this Article, findings of which were presented at the Expert group meeting on Xpert MTB/RIF on May 20–21, 2013, in Veyrier-Du-Lac, France. Additional funding came from the Global TB Program of Texas Children's Hospital, Houston, TX, USA. We thank all authors of analysed studies for the effective and timely communication, discussions and for helping with data extraction. We thank authors of continuing studies for sharing unpublished data.

References

1. WHO. Global Tuberculosis Report. Geneva: World Health Organization; 2014.
2. Dodd PJ, Gardiner E, Coghlan R, Seddon JA. Burden of childhood tuberculosis in 22 high-burden countries: a mathematical modelling study. Lancet Glob Health. 2014; 2:e453–59. [PubMed: 25103518]
3. Jenkins HE, Tolman AW, Yuen CM, et al. Incidence of multidrug-resistant tuberculosis disease in children: systematic review and global estimates. Lancet. 2014; 383:1572–79. [PubMed: 24671080]
4. Marais BJ, Gie RP, Hesseling AC, et al. A refined symptom-based approach to diagnose pulmonary tuberculosis in children. Pediatrics. 2006; 118:e1350–59. [PubMed: 17079536]
5. WHO. [accessed March 15, 2015] Policy update: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. 2013. http://www.who.int/tb/laboratory/xpert_launchupdate/en/
6. Graham SM, Ahmed T, Amanullah F, et al. Evaluation of tuberculosis diagnostics in children: 1. Proposed clinical case definitions for classification of intrathoracic tuberculosis disease. Consensus from an expert panel. J Infect Dis. 2012; 205(suppl 2):S199–208. [PubMed: 22448023]
7. WHO. Policy guidance on drug—susceptibility testing (DST) of second-line antituberculosis drugs. Geneva: World Health Organization; 2008. WHO/HTM/TB/2008.392
8. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert(R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev. 2014; 1:CD009593. [PubMed: 24448973]
9. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155:529–36. [PubMed: 22007046]
10. Macaskill, P.; Gatsonis, C.; Deeks, JJ.; Harbord, RM.; Takwoingi, Y. Analysing and presenting results. In: Deeks, JJ.; Bossuyt, PM.; Gatsonis, C., editors. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 090. Copenhagen: The Cochrane Collaboration; 2010.
11. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005; 58:982–90. [PubMed: 16168343]

12. Bates M, O'Grady J, Maeurer M, et al. Assessment of the Xpert MTB/RIF assay for diagnosis of tuberculosis with gastric lavage aspirates in children in sub-Saharan Africa: a prospective descriptive study. Lancet Infect Dis. 2013; 13:36–42. [PubMed: 23134697]
13. Causse M, Ruiz P, Gutierrez-Aroca JB, Casal M. Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. J Clin Microbiol. 2011; 49:3065–67. [PubMed: 21653775]
14. Nguyen TQ, Dang TM, Nguyen DA, et al. Evaluation of Xpert MTB/RIF and MODS assay for the diagnosis of pediatric tuberculosis. BMC Infect Dis. 2013; 13:31. [PubMed: 23343418]
15. Nicol MP, Workman L, Isaacs W, et al. Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in children admitted to hospital in Cape Town, South Africa: a descriptive study. Lancet Infect Dis. 2011; 11:819–24. [PubMed: 21764384]
16. Rachow A, Clowes P, Saathoff E, et al. Increased and expedited case detection by Xpert MTB/RIF assay in childhood tuberculosis: a prospective cohort study. Clin Infect Dis. 2012; 54:1388–96. [PubMed: 22474220]
17. Tortoli E, Russo C, Piersimoni C, et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. Eur Respir J. 2012; 40:442–47. [PubMed: 22241741]
18. Zar HJ, Workman L, Isaacs W, et al. Rapid molecular diagnosis of pulmonary tuberculosis in children using nasopharyngeal specimens. Clin Infect Dis. 2012; 55:1088–95. [PubMed: 22752518]
19. LaCourse S, Chester F, Preidis G, et al. Use of Xpert for the diagnosis of pulmonary tuberculosis in severely malnourished hospitalized Malawian children. Pediatr Infect Dis J. 2014; 33:1200–02. [PubMed: 25361410]
20. Chisti M, Graham SM, Duke T, et al. A prospective study of the prevalence of tuberculosis and bacteraemia in Bangladeshi children with severe malnutrition and pneumonia including an evaluation of Xpert MTB/RIF assay. PLoS One. 2014; 9:e93776. [PubMed: 24695758]
21. Walters E, Gie RP, Hesseling AC, Friedrich SO, Diacon AH, Gie RP. Rapid diagnosis of pediatric intrathoracic tuberculosis from stool samples using the Xpert MTB/RIF Assay: a pilot study. Pediatr Infect Dis J. 2012; 31:1316. [PubMed: 23188101]
22. Sekadde MP, Wobudeya E, Joloba ML, et al. Evaluation of the Xpert MTB/RIF test for the diagnosis of childhood pulmonary tuberculosis in Uganda: a cross-sectional diagnostic study. BMC Infect Dis. 2013; 13:133. [PubMed: 23497044]
23. Zar HJ, Workman L, Isaacs W, Dheda K, Zemanay W, Nicol MP. Rapid diagnosis of pulmonary tuberculosis in African children in a primary care setting by use of Xpert MTB/RIF on respiratory specimens: a prospective study. Lancet Glob Health. 2013; 1:e97–104. [PubMed: 25104164]
24. Pang Y, Wang Y, Zhao S, Liu J, Zhao Y, Li H. Evaluation of the Xpert MTB/RIF assay in gastric lavage aspirates for diagnosis of smear-negative childhood pulmonary tuberculosis. Pediatr Infect Dis J. 2014; 33:1047–51. [PubMed: 25361186]
25. Reither K, Manyama C, Clowes P, et al. Xpert MTB/RIF assay for diagnosis of pulmonary tuberculosis in children: a prospective, multi-centre evaluation. J Infect. 2015; 70:392–99. [PubMed: 25312863]
26. World Bank. [accessed Jan 8, 2015] World Bank list of economies. 2014. http://siteresources.worldbank.org/DATASTATISTICS/Resources/CLASS.XLS
27. Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: a prospective study. Lancet. 2005; 365:130–34. [PubMed: 15639294]
28. Marais BJ, Hesseling AC, Gie RP, Schaaf HS, Enarson DA, Beyers N. The bacteriologic yield in children with intrathoracic tuberculosis. Clin Infect Dis. 2006; 42:e69–71. [PubMed: 16575719]
29. Helb D, Jones M, Story E, et al. Rapid detection of *Mycobacterium tuberculosis* and rifampin resistance by use of on-demand, near-patient technology. J Clin Microbiol. 2010; 48:229–37. [PubMed: 19864480]
30. Schaaf HS, Marais BJ, Whitelaw A, et al. Culture-confirmed childhood tuberculosis in Cape Town, South Africa: a review of 596 cases. BMC Infect Dis. 2007; 7:140. [PubMed: 18047651]
31. Raizada N, Sachdeva KS, Nair SA, et al. Enhancing TB case detection: experience in offering upfront Xpert MTB/RIF testing to pediatric presumptive TB and DR TB cases for early rapid

diagnosis of drug sensitive and drug resistant TB. PLoS One. 2014; 9:e105346. [PubMed: 25140877]

32. Naidoo P, du Toit E, Dunbar R, et al. A comparison of multidrug-resistant tuberculosis treatment commencement times in MDRTBPlus line probe assay and Xpert(R) MTB/RIF-based algorithms in a routine operational setting in Cape Town. PLoS One. 2014; 9:e103328. [PubMed: 25079599]
33. Theron G, Zijenah L, Chanda D, et al. Feasibility, accuracy, and clinical effect of point-of-care Xpert MTB/RIF testing for tuberculosis in primary-care settings in Africa: a multicentre, randomised, controlled trial. Lancet. 2014; 383:424–35. [PubMed: 24176144]
34. Hanrahan CF, Selibas K, Deery CB, et al. Time to treatment and patient outcomes among TB suspects screened by a single point-of-care xpert MTB/RIF at a primary care clinic in Johannesburg, South Africa. PLoS One. 2013; 8:e65421. [PubMed: 23762367]
35. Wiseman CA, Gie RP, Starke JR, et al. A proposed comprehensive classification of tuberculosis disease severity in children. Pediatr Infect Dis J. 2012; 31:347–52. [PubMed: 22315002]
36. Nicol MP, Spiers K, Workman L, et al. Xpert MTB/RIF testing of stool samples for the diagnosis of pulmonary tuberculosis in children. Clin Infect Dis. 2013; 57:e18–21. [PubMed: 23580738]
37. Walters E, Goussard P, Bosch C, Hesseling AC, Gie RP. GeneXpert MTB/RIF on bronchoalveolar lavage samples in children with suspected complicated intrathoracic tuberculosis: a pilot study. Pediatr Pulmonol. 2014; 49:1133–37. [PubMed: 24339262]
38. Yin QQ, Jiao WW, Han R, et al. Rapid diagnosis of childhood pulmonary tuberculosis by Xpert MTB/RIF assay using bronchoalveolar lavage fluid. Biomed Res Int. 2014; 2014:310194. [PubMed: 25165698]
39. Denkinger CM, Schumacher SG, Boehme CC, Dendukuri N, Pai M, Steingart KR. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2014; 44:435–46. [PubMed: 24696113]

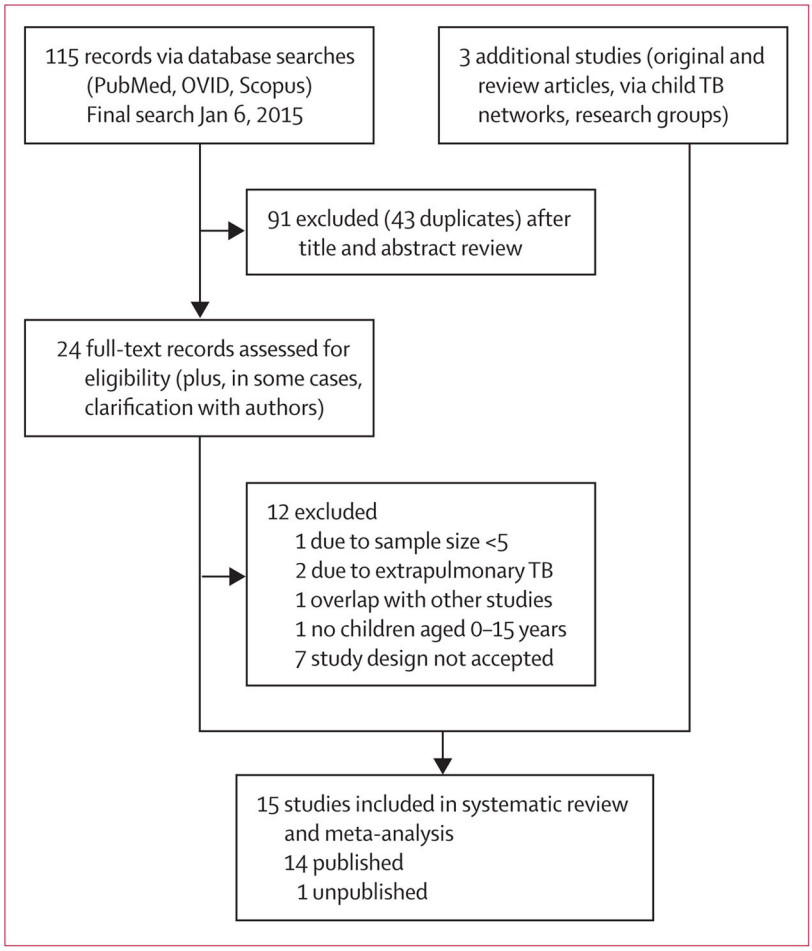

Figure 1. Study selection
TB=tuberculosis.

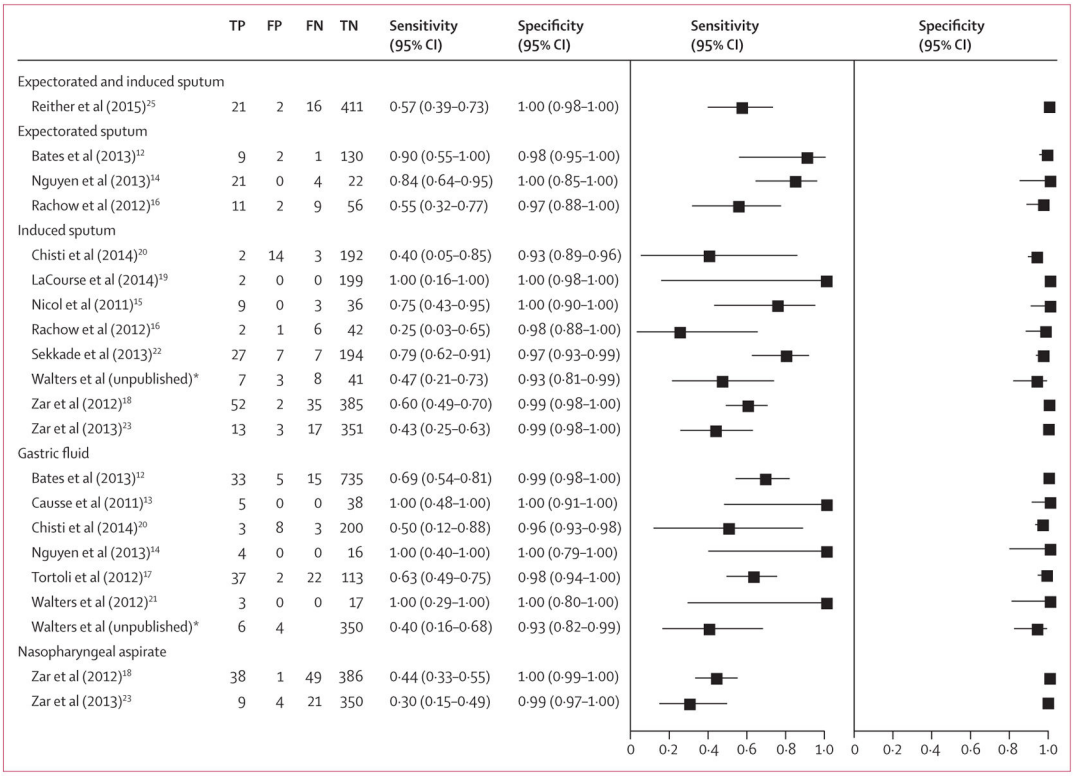

Figure 2. Sensitivity and specificity of Xpert against culture reference standard by study and specimen type

TP=true positive. FP=false positive. FN=false negative. TN=true negative. *Walters E, Desmond Tutu TB Centre, personal communication.

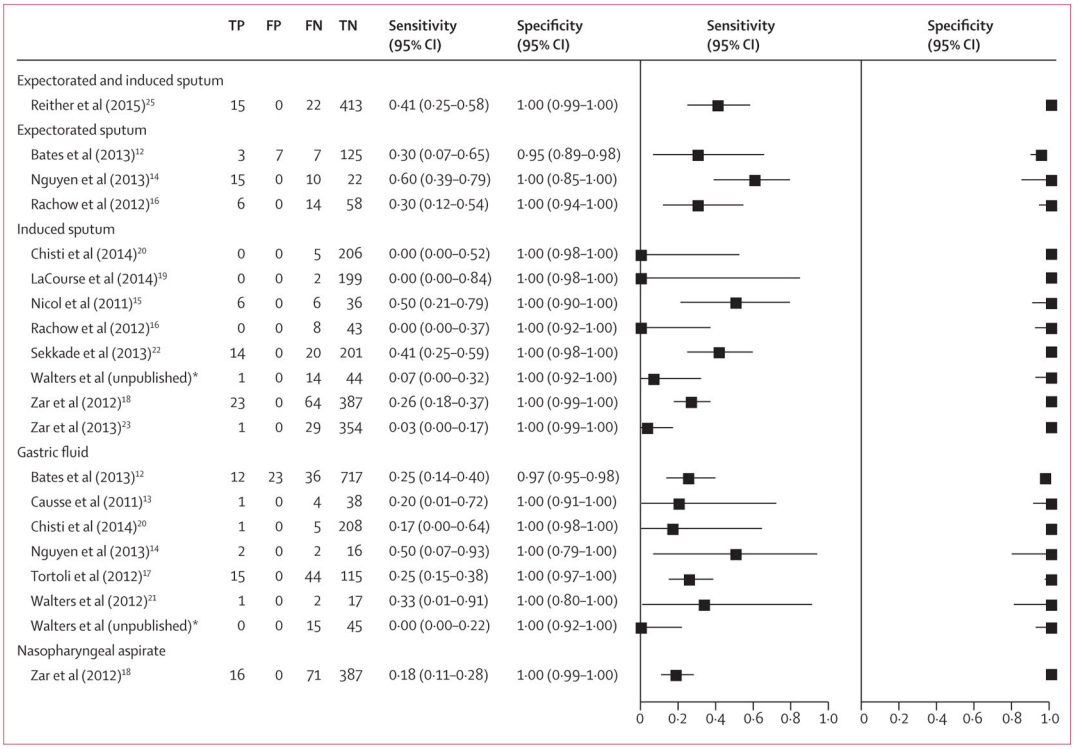

Figure 3. Sensitivity and specificity of smear microscopy against culture reference standard by study and specimen type

TP=true positive. FP=false positive. FN=false negative. TN=true negative. *Walters E, Desmond Tutu TB Centre, personal communication.

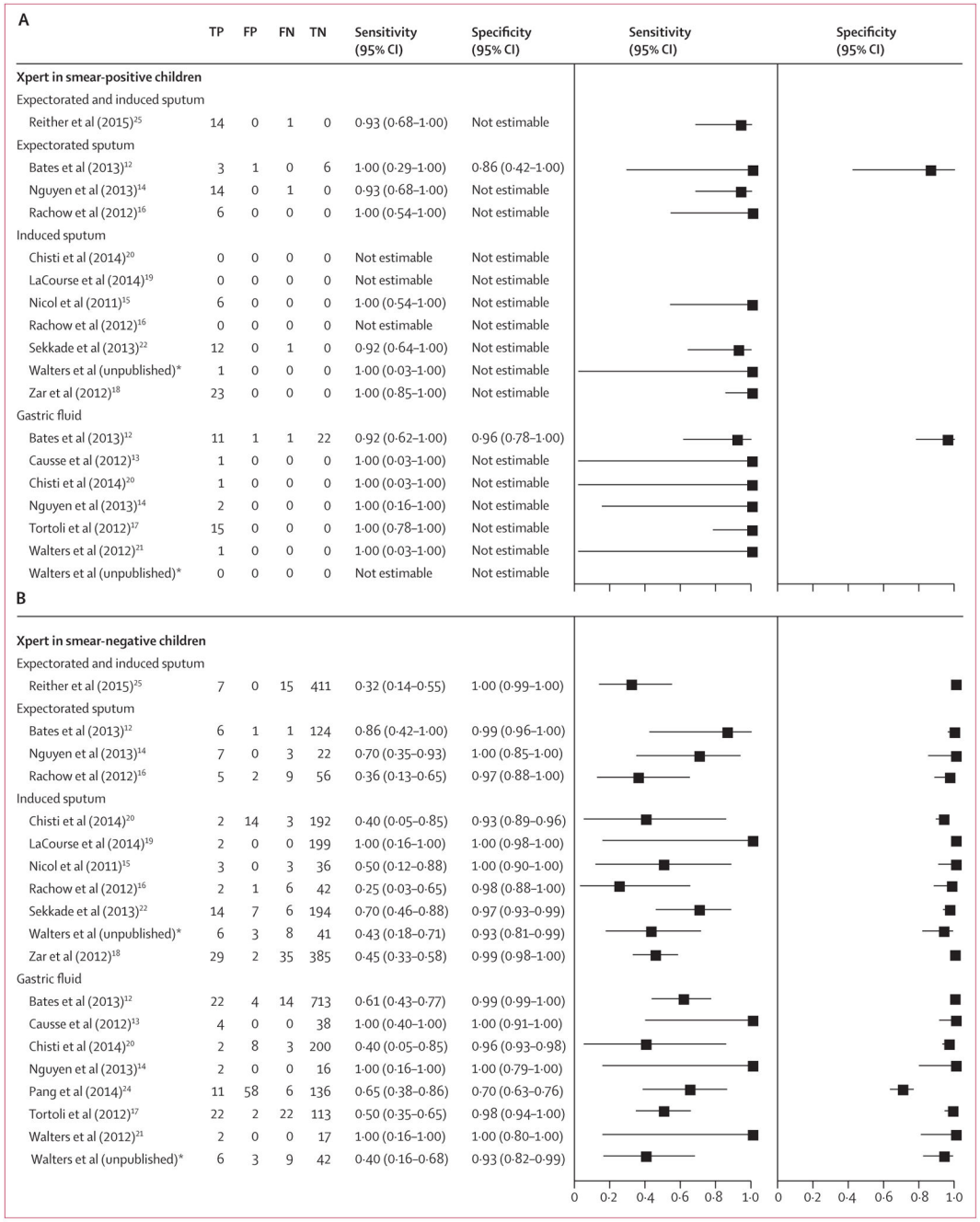

Figure 4. Sensitivity and specificity of Xpert against a culture reference standard by smear status

TP=true positive. FP=false positive. FN=false negative. TN=true negative. *Walters E, Desmond Tutu TB Centre, personal communication.

Table 1

Features of included studies

	Country	World bank income classification	TB incidence rate per population	Clinical setting	Laboratory setting	Sample size included for analysis (percentage culture-positive)	Age (months)	HIV-infection prevalence (%)
Bates et al (2013)[12]	Zambia	LMIC	410 per 100 000	Inpatients	University hospital laboratory (tertiary referral center)	142 for ES (7%); 788 for GLA (6%)	Median 24 (IQR 12–74); ES group 119 (IQR 84–44); GLA group 20 (IQR 11–40)	32%
Causse et al (2011)[13]	Spain	HIC	13 per 100 000	Unclear*	Central (reference)	43 for GLA (12%); 41 for RIF resistance	Mean 43; range 3–180	NR
Chisti et al (2014)[20]	Bangladesh	LIC	224 per 100 000	Inpatients	Research laboratory	211 for IS (2%); 214 for GLA (3%); same children, two different specimens	Mean 13; range 2–52	NR
LaCourse et al (2014)[19]	Malawi	LIC	156 per 100 000	Inpatients	Research laboratory	201 for IS (1%)	Mean 21; median 18; range 6–60	18%
Nguyen et al (2013)[14]	Vietnam	LMIC	144 per 100 000	Inpatients	Hospital laboratory	47 for ES (53%); 20 for GLA (20%)	Median 106; range 0–180	10% (7 of 73, but only 8 tested)
Nicol et al (2011)[15]	South Africa	UMIC	860 per 100 000	Inpatients	University hospital laboratory	48 for IS (25%); 18 for RIF resistance	Median 72; range 5–156	38%
Pang et al (2014)[24†]	China	UMIC	70 per 100 000	Inpatients	Central (reference)	211 for GLA (8·1%)	NR	NR
Rachow et al (2012)[16]	Tanzania	LIC	164 per 100 000	Inpatients (about 30%) and outpatients	Research laboratory	78 for ES (26%); 51 for IS (16%); 25 for RIF resistance	Median 70; IQR 29–113	54%
Reither et al (2015)[25]	Tanzania & Uganda	LIC	164 per 100 000; 166 per 100 000	Inpatients (20%) and outpatients	Research laboratory	177 for ES; 273 for IS (8%)	Mean 67; IQR 24–118	44%

	Country	World bank income classification	TB incidence rate per population	Clinical setting	Laboratory setting	Sample size included for analysis (percentage culture-positive)	Age (months)	HIV-infection prevalence (%)
Sekadde et al (2013)[22]	Uganda	LIC	166 per 100 000	Inpatients (84%) and outpatients	University microbiology laboratory	235 for IS (15%)	Median 36; range 2–144	42%
Tortoli et al (2012)[17]	Italy	HIC	5·7 per 100 000	Unclear*	8 different higher level routine laboratories	174 for GLA (20%)	Mean 91; range 0–156	NR
Walters et al (2012)[21]	South Africa	UMIC	860 per 100 000	Inpatients	University hospital laboratory	20 for GLA (15%); 14 for RIF resistance	Median 17; range 3–137	0%
Walters et al (unpublished) ‡	South Africa	UMIC	860 per 100 000	Inpatients (85%) and outpatients	University hospital laboratory	59 for IS (25%); 60 for GLA (25%); same children, two different specimens	Median 13; range 1–59	12%
Zar et al (2012)[18]	South Africa	UMIC	860 per 100 000	Inpatients	University hospital laboratory	474 for IS (18%); 474 for NPA (18%); same children, two different specimens; 125 for RIF resistance	Median 19; IQR 11–38	25%
Zar et al (2013)[23]	South Africa	UMIC	860 per 100 000	Outpatients	University hospital laboratory	384 for IS (8%); 384 for NPA; same children, two different specimens; 13 for RIF resistance	Median 38; IQR 21–57	8%

Data for TB incidence rate from WHO, 2013. TB=tuberculosis. LMIC=lower-middle-income country. ES=expectorated sputum. GLA=gastric lavage. HIC=high-income country. RIF=rifampicin. LIC=low-income country. NR=not reported. IS=induced sputum. UMIC=upper-middle-income country.

* These studies were led by laboratory researchers and reports had no or very little information for the origin of specimens—ie, from inpatients versus outpatients.

† This study only included smear-negative children and was therefore only included in the analysis of Xpert against culture in smear-negative children.

‡ Walters E, Desmond Tutu TB Centre, personal communication.

Table 2

Characteristics and findings of studies assessing culture and Xpert positivity in different subgroups

	Inclusion criteria*	Specimen type	Participants (N)	Culture positive† (n; %)	Xpert positive† (n; %)	Number of samples per child sent for culture‡	Culture-negative children initiated on ATT (n, %)§	Culture-negative children initiated on ATT that were Xpert positive (n; %)
Bates et al (2013)[12]	Broad	ES	142	10 (7%)	11 (8%)	1	..	..
Bates et al (2013)[12]	Broad	GLA	788	48 (6%)	38 (5%)	1	..	..
Causse et al (2011)[13]	Broad	GLA	43	5 (12%)	5 (12%)	Up to 2	..	..
Chisti et al (2014)[20]	Broad	IS	211¶	6 (3%)	16 (8%)	Up to 2		..
Chisti et al (2014)[20]	Broad	GLA	214¶	..	11 (5%)	..	..	..
LaCourse et al (2014)[19]	Broad	IS	201	2 (1%)	2 (1%)	Up to 2	130 (65%)	0
Nguyen et al (2013)[14]	Broad	ES	47	25 (53%)	21 (45%)	Up to 3	13 (28%)	0
Nguyen et al (2013)[14]	Broad	GLA	20	4 (20%)	4 (20%)		11 (55%)	0
Nicol et al (2011)[15]	Rigorous	IS	48	12 (25%)	9 (19%)	Up to 2	17 (35%)	0
Pang et al (2014)[24]	Rigorous	GLA	211	17 (8%)	69 (33%)	1	..	..
Rachow et al (2012)[16]	Rigorous	ES	78	20 (26%)	13 (17%)	Up to 6	45 (58%)	2 (<1%)
Rachow et al (2012)[16]	Rigorous	IS	51	8 (16%)	3 (6%)	..	34 (67%)	1 (<1%)
Reither et al (2015)[25]	Rigorous	ES/IS‖	450	37 (8%)	33 (7%)	At least 2	92 (34%)	2 (<1%)
Sekadde et al (2013)[22]	Rigorous	IS	235	34 (15%)	34 (15%)	Up to 2	..	..
Tortoli et al (2012)[17]	Broad	GLA	174	57 (33%)	39 (22%)	2	..	..
Walters et al (2012)[21]	Rigorous	GLA	20	3 (15%)	3 (15%)	1	..	..
Walters**	Rigorous	IS	59¶	15 (25%)	10 (17%)	Up to 4	..	..
Walters**	Rigorous	GLA	60¶	..	9 (15%)	..	..	
Zar et al (2012)[18]	Rigorous	IS	474	87 (18%)	54 (11%)	Up to 4	194 (41%)	2 (<1%)
Zar et al (2013)[23]	Rigorous	IS	384	30 (8%)	22 (6%)	Up to 2	152 (40%)	3 (<1%)
Total	..	..	3640	420 (12%)	406 (11%)	..	..	..

ATT=antituberculosis treatment. ES=expectorated sputum. GLA=gastric lavage. IS=induced sputum.

[*] Rigorous defined as studies that had clear criteria for a presumed tuberculosis case for eligibility. Broad defined as studies that had either unclear or very broadly defined criteria to define presumed tuberculosis.

[†] Culture positive is defined as any positive culture of all cultures done. Xpert results are based on one Xpert assay per child.

[‡] If more than one culture done: any positive culture meant the child was defined as confirmed TB. Depending on the study, multiple cultures were performed on the same specimen (different culture methods) or multiple specimens or specimen types per child.

[§] Empirical ATT was initiated without knowledge of Xpert results in culture-negative children.

[¶] In these studies, two different specimen types were taken from the same cohort of children. The numbers at the bottom of the column reflect the total and mean number of children per study, hence for Chisti et al (2014)[20] 214 children were used and for Walters (unpublished), 60 children were added to the calculations.

[||] Data for children with either ES or IS were analysed together.

[**] Walters E, Desmond Tutu TB Centre, personal communication.

Table 3

Meta-analysis findings for estimated Xpert and microscopy sensitivity and specificity against the reference standards

	Number of studies (number of children)	Pooled and predicted median sensitivity (pooled 95% credible interval; predicted 95% credible interval)	Pooled and predicted median specificity (pooled 95% credible interval; predicted 95% credible interval)
Xpert against culture			
Expectorated and induced sputum	12 (2380)[12,14–16,18–20,22,23,25*†]	62% (51–73; 30–87)	98% (97–99; 90–100)
Gastric lavage	7 (1319)[12–14,17,20,21†]	66% (51–81; 33–91)	98% (96–99; 91–100)
Xpert against culture-negative and started on ATT			
Expectorated and induced sputum	8 (1512)[14–16,18,19,23,25*]	2% (1–3; 0–6)	100% (99–100; 99–100)
Smear microscopy against culture			
Expectorated and induced sputum	12 (2380)[12,14–16,18–20,22,23,25*†]	26% (14–39; 4–69)	100% (99–100; 94–100)
Gastric lavage	7 (1319)[12–14,17,20,21†]	22% (12–35; 6–51)	99% (97–100; 93–100)

Includes published and one unpublished study. ATT=antituberculosis treatment.

*Expectorated and induced sputum cohorts from reference 16 included as separate studies.

†Also includes data from Walters E, Desmond Tutu TB Centre, personal communication.

Table 4

Meta-analysis for sensitivity of Xpert against reference standard of culture in children by age as well as HIV status, subdivided by smear status

	Specimen type (number of studies, number of children)	Median pooled and predicted sensitivity (pooled 95% credible interval; predicted 95% credibility level)
Smear-positive status		
Age 0–4 years	ES and IS (6; 27)[12,15,18,22,23,25]	94% (79–99; 69–99)
Age 0–4 years	GLA (4; 12)[12,14,20,21]	82% (51–96; 40–97)
Age 5–15 years	ES and IS (5; 35)[12,15,18,22,23,25]	96% (85–99; 78–100)
HIV-positive	ES and IS (6; 25)[12,15,16,18,22,25]	97% (87–100; 82–100)
HIV-negative	ES and IS (7; 41)[12,15,16,18,22,23,25]	94% (83–99; 75–99)
Smear-negative status		
Age 0–4 years	ES and IS (6; 109)[12,15,18,22,23,25]	44% (30–61; 18–74)
Age 5–15 years	ES and IS (5; 34)[12,15,18,22,23,25]	66% (44–84; 29–92)
HIV-positive	ES and IS (7; 36)[12,15,16,18,22,23,25]	60%* (40–77; 26–87)
HIV-negative	ES and IS (7; 125)[12,15,16,18,22,23,25]	44% (30–59; 18–73)

There were not enough data to assess GLA results in the age group 5–15 years. ES=expectorated sputum. IC=induced sputum. GLA=gastric lavage.

*Predicted sensitivity is 59% due to rounding.