# Diagnosis of Pulmonary Tuberculosis in Children: Assessment of the 2012 National Institutes of Health Expert Consensus Criteria

**Heather J. Zar,[1] Lesley J. Workman,[1] Francesca Little,[2] and Mark P. Nicol[3]**

[1]Department of Paediatrics and Child Health, Red Cross War Memorial Children's Hospital, University of Cape Town and Medical Research Council Unit on Child and Adolescent Health, [2]Department of Statistical Sciences, and [3]Division of Medical Microbiology and Institute for Infectious Diseases and Molecular Medicine, University of Cape Town, National Health Laboratory Service, Groote Schuur Hospital, South Africa

*Background.* The 2012 National Institutes of Health (NIH) consensus criteria for standardized diagnostic categories of pulmonary tuberculosis in children have not been validated. We aimed to assess the NIH diagnostic criteria in children with culture-confirmed pulmonary tuberculosis and those in whom tuberculosis has been excluded.

*Methods.* We performed a retrospective analysis of consecutive children hospitalized with suspected pulmonary tuberculosis in Cape Town, South Africa, who were enrolled in a diagnostic study. Children were categorized as definite tuberculosis (culture positive), probable tuberculosis (chest radiograph consistent), possible tuberculosis (chest radiograph inconsistent), or not tuberculosis (improved without tuberculosis treatment). We applied the NIH diagnostic categories to the cohort and evaluated their performance specifically in children with definite tuberculosis and not tuberculosis.

*Results.* Four hundred sixty-four children (median age, 25.1 months [interquartile range, 13.5–61.5 months]) were included; 96 (20.7%) were HIV infected. Of these, 165 (35.6%) were definite tuberculosis, and 299 (64.4%) were not tuberculosis. If strict NIH symptom criteria were applied, 100 (21.6%) were unclassifiable including 21 (21.0%) with definite pulmonary tuberculosis, as they did not meet the NIH criteria due to short duration of symptoms; 71 (71%) had cough <14 days, 48 (48%) had recent weight loss, and 39 (39%) had fever <7 days. Of 364 classifiable children, there was moderate agreement ($\kappa = 0.48$) with 100% agreement for definite tuberculosis and moderate agreement for not tuberculosis (220 [60.4%] vs 89 [24.5%]).

*Conclusions.* Entry criteria for diagnostic studies should not be restrictive. Data from this analysis have informed revision of the NIH definitions.

*Keywords.* tuberculosis; child; diagnosis; NIH consensus.

Diagnosis of pulmonary tuberculosis in children can be challenging due to nonspecific symptoms, signs, and radiological changes and the difficulty of making a definitive microbiologic diagnosis [1]. Advances in specimen collection and diagnostic tests have improved the ability to make a microbiologic diagnosis in children [2–6]; however, microbiologic confirmation is still unavailable in many settings. Further, among children with suspected pulmonary tuberculosis, a confirmed microbiologic diagnosis is only achieved in a minority; thus, clinical diagnosis remains the predominant method for determining whether a child has tuberculosis and initiating treatment. However, clinical diagnosis may vary widely. Several clinical scoring systems have been developed, but there is wide variability in their performance [7, 8].

The development of new diagnostic tests for tuberculosis has been an important advance. New diagnostics have highlighted the need for standardized clinical

case definitions for pulmonary tuberculosis in children in the context of clinical research, so as to enable comparative evaluation of such tests [9]. To obtain consensus on clinical case definitions for diagnosis of pulmonary tuberculosis in children, the National Institutes of Health (NIH) convened a workshop of expert clinicians, researchers, and other opinion leaders in June 2011, to develop consensus diagnostic criteria for tuberculosis diagnostic studies for use in clinical research, based on clinical, radiological, and microbiological features [10, 11]. Using these, the expert panel recommended 5 clinical diagnostic categories—confirmed tuberculosis, probable tuberculosis, possible tuberculosis, tuberculosis unlikely, and not tuberculosis [10]. These diagnostic categories and criteria are based on expert opinion and have not yet been validated. Moreover, there has been confusion regarding interpretation of some of the definitions and classifications.

Since 2009, we have been performing clinical research on the performance of new diagnostic tests for pulmonary tuberculosis in children at a pediatric hospital in Cape Town, South Africa. Clinical, radiologic, and microbiologic data were used to classify children into 1 of 4 locally defined diagnostic categories: confirmed tuberculosis, probable tuberculosis, possible tuberculosis, or not tuberculosis. We used data collected in this research to evaluate NIH consensus criteria for tuberculosis diagnostic studies.

## METHODS

A retrospective analysis of consecutive children hospitalized with suspected pulmonary tuberculosis in Cape Town, South Africa, from 1 February 2009 to 30 June 2012, who were enrolled in a tuberculosis diagnostic study, was done. Children aged <15 years were eligible for the diagnostic study if they had a cough and at least 1 of the following: a household tuberculosis contact within the preceding 3 months, weight loss or failure to gain weight within the preceding 3 months, a positive tuberculin skin test (TST) to purified protein derivative (PPD 2TU, PPD RT23, Statens Serum Institut, Denmark), or a chest radiograph suggestive of pulmonary tuberculosis. Children were excluded if they had received tuberculosis drug(s) for >72 hours, did not live in Cape Town, and could not attend follow-up visits if informed consent was unavailable or if a respiratory specimen was not obtained. Written informed consent was obtained from a parent or legal guardian. The study was approved by the Ethics Committee of the Faculty of Health Sciences, University of Cape Town.

Signs, symptoms, and exposure history were carefully recorded. Routine investigations included chest radiography (anteroposterior and lateral views), a TST, and human immunodeficiency virus (HIV) testing when HIV status was unknown (HIV rapid test in all children, followed by a confirmatory polymerase chain reaction for children <18 months of age or HIV enzyme-linked immunosorbent assay for children aged ≥18 months). Two experienced reviewers, blind to microbiological and other results, reported chest radiographs using a standardized format and classified the radiograph as consistent or inconsistent with pulmonary tuberculosis. When there was a discrepancy in the radiographic assessment, a third reviewer's report was obtained. The same reviewers were used for the duration of the study.

Tuberculosis treatment was at the discretion of the hospital doctor based on clinical, radiological, and microbiologic information. Follow-up visits were done by study staff at 1, 3, and 6 months for children on tuberculosis therapy and at 1 and 3 months for those untreated. Response to treatment was assessed at follow-up by recording symptoms, signs, weight, and a repeat chest radiograph at completion of tuberculosis treatment.

### Microbiological Investigations

Children were investigated for pulmonary tuberculosis as has been previously described [2, 6]. Specimens for microbiologic testing included 2 sequential induced sputa specimens and 2 nasopharyngeal aspirates and, if clinically indicated and feasible, gastric lavage, tracheal aspirate, or bronchoalveolar lavage. Specimens were obtained for acid-fast staining, culture confirmation using automated liquid culture (BACTEC MGIT, Becton Dickinson), and rapid diagnostic testing with Xpert MTB/RIF assay (Xpert, Cepheid).

### Clinical Diagnostic Categories
#### Cape Town Categories
Children were categorized as "definite tuberculosis" (culture positive on any specimen for *Mycobacterium tuberculosis* [*Mtb*]), probable tuberculosis (chest radiograph suggestive of tuberculosis), possible tuberculosis (chest radiograph inconsistent with tuberculosis), or "not tuberculosis" (documented resolution of symptoms and signs at follow-up in children who did not receive tuberculosis treatment) (Table 1) [2, 6].

#### NIH Categories
Children were classified according to the NIH criteria into 1 of 5 diagnostic categories as follows [10]:

1. Confirmed tuberculosis (at least 1 sign or symptom and 1 positive culture);

2. Probable tuberculosis (at least 1 sign or symptom and a chest radiograph consistent with tuberculosis and 1 of: response to tuberculosis therapy/documented exposure to *Mtb*/immunologic evidence of infection with *Mtb*);

3. Possible tuberculosis (at least 1 sign or symptom and 1 of: chest radiograph consistent with tuberculosis/response to tuberculosis therapy/documented exposure to *Mtb*/immunologic evidence of infection with *Mtb*);

4. Tuberculosis unlikely—symptomatic but no other criteria;

**Table 1. Summary of Criteria Used to Classify Tuberculosis Disease in Children**

| | Cape Town Criteria | NIH Criteria |
|---|---|---|
| Signs/symptoms of TB | One of the following:<br>• Cough<br>• Reported loss of weight or failure to gain weight in past 3 mo | One of the following:<br>• Cough >14 d<br>• Documented weight loss in past 3 mo or HAZ or WAZ ≤−2<br>• Persistent lethargy<br>• Fever >7 d |
| Microbiology | At least 1 positive culture from a respiratory specimen for *Mtb* | At least 1 positive culture for *Mtb* |
| CXR consistent with TB | • CXRs reviewed by 2 independent experts<br>• Discordant results were read by a third expert<br>• Standardized reporting form<br>• CXR was considered as consistent with TB if 2/2 or 2/3 expert reviewers considered the CXR to be consistent with TB | • CXRs reviewed by 2 independent experts<br>• Discordant results were read by a third expert<br>• Standardized reporting form<br>• CXR is classified as consistent with TB if there is positive response for any 1 of the radiological features of TB, at the same location by at least 2 expert readers |
| Immunological evidence of *M. tuberculosis* infection | • Positive TST ≥10 mm in HIV-uninfected children; ≥5 mm in HIV-infected children<br>OR<br>• Positive IGRA | • Positive TST ≥10 mm in HIV-uninfected children; ≥5 mm in HIV-infected children<br>OR<br>• Positive IGRA |
| Tuberculosis exposure | History of TB contact in the preceding 3 mo | Reported exposure to a case of tuberculosis (household/close contact) within the preceding 24 mo |

Abbreviations: CXR, chest radiography; HAZ, height-for-age *z* score; HIV, human immunodeficiency virus; IGRA, interferon-γ release assay; *Mtb*, *Mycobacterium tuberculosis*; NIH, National Institutes of Health; TB, tuberculosis; TST, tuberculin skin test; WAZ, weight-for-age *z* score.

5. Not tuberculosis—symptomatic but no other criteria and an alternative diagnosis established.

To assign a diagnostic category using the NIH algorithm, assessment was consecutively made in 3 areas—symptom criteria, microbiologic confirmation, and radiologic evaluation. To enter the algorithm for classification, children had to meet the NIH-defined symptom criteria, followed by microbiologic confirmation and radiological evaluation [10]. Symptoms listed in the NIH criteria [10] include persistent cough (for >14 days), weight loss, or failure to thrive within the preceding 3 months (documented on the growth curve or a weight-for age *z* score [WAZ] or height-for-age *z* score [HAZ] of ≤−2), persistent unexplained fever (for >7 days), or persistent, unexplained lethargy (Table 1). Similar to the Cape Town study, radiologic classification required a minimum of 2 readers, with a third expert reader in the case of discordant reports. The chest radiograph was classified as consistent or not consistent with tuberculosis according to a standard format. Documented exposure to *Mtb* was defined as reported exposure to a household contact within the preceding 24 months or a positive TST or interferon-γ release assay as per the NIH guidelines.

**Analysis**

The data were analyzed using Stata 11 statistical software (StataCorp, College Station, Texas). Simple descriptive statistics were used to characterize the study population; normally distributed continuous data were summarized by mean and 95% confidence interval, and nonnormally distributed continuous data by median and interquartile range. WAZ and HAZ were calculated using EpiInfo version 6 (database and statistical software for public health professionals, Centers for Disease Control and Prevention [CDC]). Categorical data were summarized as number and proportion. For this analysis, the NIH categories of *Mtb* infection (which appears on NIH algorithm but is not listed in the diagnostic categories), unlikely tuberculosis, and not tuberculosis were combined into a single "not tuberculosis" category. Agreement between the 2 algorithms was measured using the κ statistic. The strength of agreement between the 2 algorithms was described as κ = 0–0.2, slight; 0.2–0.4 as fair; 0.4–0.6 as moderate; 0.6–0.8 as substantial; and 0.8–1.0 as almost perfect [12].

**RESULTS**

Among 924 children with suspected pulmonary tuberculosis enrolled in the Cape Town study, 25 without symptoms or a microbiologic specimen, 327 classified as possible tuberculosis, and 108 classified as probable tuberculosis were excluded from this analysis (Figure 1). Thus, 464 children with a median age of 25.1 months (interquartile range [IQR], 13.5–61.5 months) were included; 247 (53.2%) were male and 96 (20.7%) were HIV infected. Of these, 165 (35.6%) were classified as definite tuberculosis, and 299 (64.4%) as not tuberculosis in our study (Figure 1).

Using the strict symptom-based definitions of the NIH criteria, 100 of 464 (21.6%) children were unclassifiable, as they did not meet the symptom criteria for inclusion. Of these, 71 (71%)
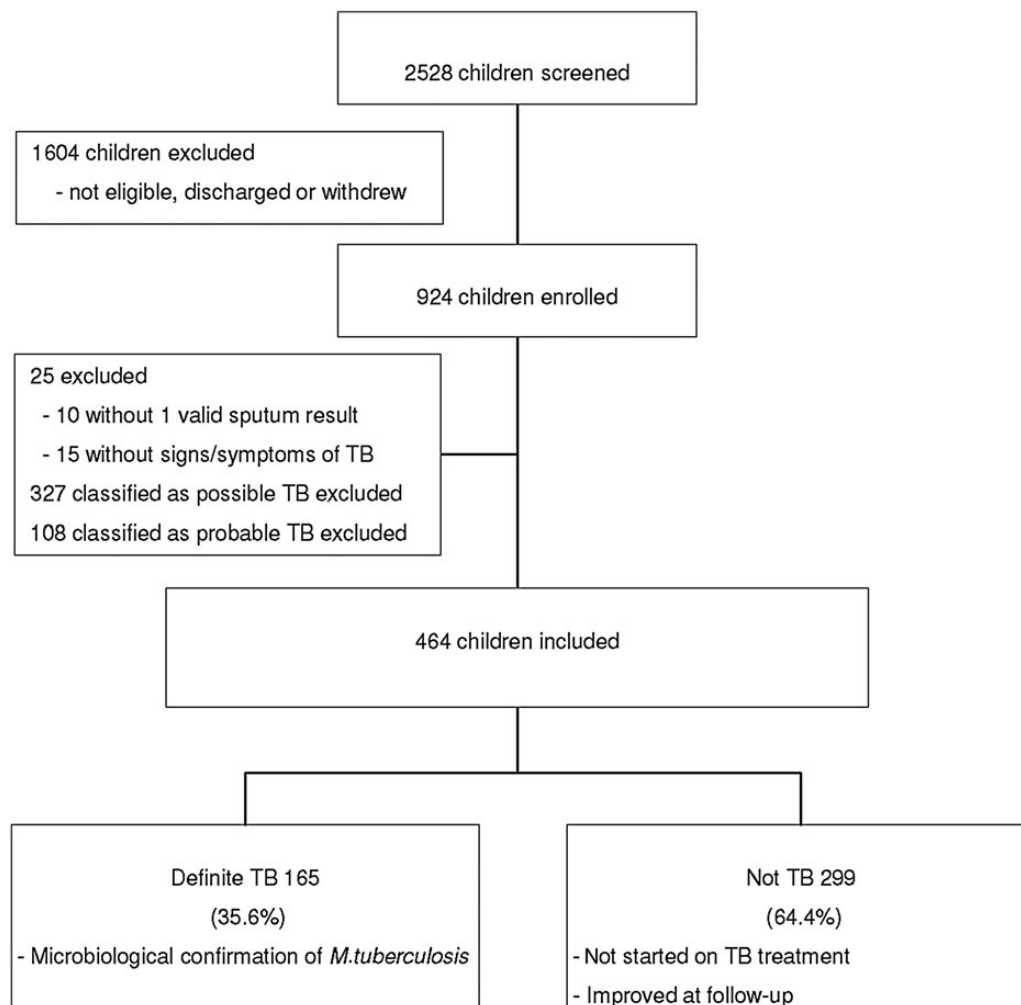
**Figure 1.** Study profile. Abbreviation: TB, tuberculosis.

presented with a cough for <14 days (median duration of cough, 4 days [IQR, 3–7 days]), 39 (39%) presented with a history of fever for <7 days (median duration of fever, 3 days [IQR, 2–4 days]), and 48 (48%) had reported weight loss but did not meet the NIH criteria as there was no documented evidence for preceding weight loss, and WAZ and HAZ were both >−2. Among these 100 children, 21 (21%) were classified as definite tuberculosis and 79 (79%) not tuberculosis in the Cape Town diagnostic study.

Of 364 children who met the NIH strict symptom criteria, all had at least 1 microbiologic specimen obtained. Among these, there was moderate agreement for the diagnostic categories between the Cape Town and NIH categories ($\kappa = 0.48$; Table 2). There was 100% agreement for definite tuberculosis, with 144 children classified as definite tuberculosis by both. There was moderate agreement between those classified in the not tuberculosis groups, with 220 (60.4%) children classified by the Cape Town study and 89 (24.5%) in the NIH categorization. Of the

220 children classified as not tuberculosis in the Cape Town study, 89 (40.5%) were classified as not tuberculosis by the NIH criteria, whereas 20 (9.1%) were classified as probable and 111 (50.4%) were considered possible tuberculosis (Table 2).

**Table 2. Diagnostic Categories of Children (n, %) Classified as Definite or Not Tuberculosis With the Cape Town Compared to the National Institutes of Health Classification Criteria**

| | | Cape Town Algorithm | | |
|---|---|---|---|---|
| | Categories | Definite TB | Not TB | Totals |
| NIH algorithm | Definite TB | 144 (100.0%) | 0 (0.0%) | 144 (39.6%) |
| | Probable TB | 0 (0.0%) | 20 (9.1%) | 20 (5.5%) |
| | Possible TB | 0 (0.0%) | 111 (50.4%) | 111 (30.5%) |
| | Not TB | 0 (0.0%) | 89 (40.5%) | 89 (24.4%) |
| | Totals | 144 (39.6%) | 220 (60.4%) | 364 (100.0%) |

$\kappa = 0.48$.

Abbreviations: NIH, National Institutes of Health; TB, tuberculosis.

## DISCUSSION

The development of validated, standardized diagnostic categories as an endpoint for tuberculosis diagnostic or vaccine studies or entry into clinical trials of novel therapy is critical [9]. However, among this cohort of children classified as definite or not tuberculosis using the Cape Town algorithm, the strict symptom-based definitions contained in the NIH classification could not be applied to approximately 21% of children, including a substantial proportion who had culture-confirmed disease. A major difference between the NIH and Cape Town classification methods is in the criteria for clinical symptoms and duration of symptoms, particularly the NIH requirement of cough for >14 days or fever for >7 days. As a result, several children did not meet the strict symptom definition criteria for classification by the NIH criteria, including 21% with culture-confirmed, definite tuberculosis. Acute symptoms have increasingly been reported for children with pulmonary tuberculosis; studies of children hospitalized with acute pneumonia in tuberculosis-endemic areas have reported 8%–15% of children to have culture-confirmed tuberculosis [13–15]. A recent meta-analysis reported that in high-tuberculosis-burden areas, tuberculosis was common in children presenting with acute pneumonia [16]. Consideration should therefore be given to modifying the NIH symptom criteria, specifically the duration of cough and the duration of fever. This important issue has since been clarified in the revised 2015 NIH definitions; the revised definitions include symptoms or signs of any duration as entry criteria.

There was a substantial difference between the number of children diagnosed as not having tuberculosis in the Cape Town study compared to when the NIH criteria were applied. This is partly due to the lack of a category for children who improved without tuberculosis treatment without immunological evidence of tuberculosis infection or documented tuberculosis exposure in the NIH algorithm. However, follow-up 2 months after initial evaluation is recommended in the NIH consensus document [10]. Therefore, consideration should be given to including an improvement in signs or symptoms without tuberculosis therapy as a criterion for the diagnostic category of not tuberculosis irrespective of tuberculosis exposure, as in the Cape Town study [2, 6]. Furthermore, many of the children classified as not tuberculosis by the Cape Town criteria were classified as possible or probable tuberculosis using the NIH criteria, based on the radiological assessment and documented exposure or immunologic evidence of infection. However, as diagnostic studies are commonly done in high-tuberculosis-prevalence areas, and given the NIH criteria of household exposure within the preceding 24 months, it is likely that many children will be considered exposed [17]. For example, a child with a chest radiograph that is inconsistent with tuberculosis but who had a household contact with tuberculosis 2 years prior would be cat-

egorized as possible tuberculosis by NIH criteria even if they improved without tuberculosis treatment. Consideration should therefore be given to modifying the criteria for "not tuberculosis" to include an improvement without tuberculosis therapy. The revised NIH definition has also now reduced the period for household exposure to 12 months.

Furthermore, classification as not tuberculosis in the NIH algorithm depends on an alternate diagnosis being established, which may be challenging [18, 19]. Bacteremia occurs only in a minority of pneumonia episodes [18]. Improvements in molecular diagnostic testing have enabled identification of viruses, bacteria, or atypical organisms on respiratory secretions; however, it is currently not possible to distinguish colonizing organisms from those causing disease [20]. Moreover, dual or multiple pathogens are common, especially in severe pneumonia, so identification of an alternate pathogen does not necessarily exclude pulmonary tuberculosis [15, 16, 18]. The NIH algorithm recommends that children be classified as "not tuberculosis" where an alternate diagnosis has been made and "tuberculosis unlikely" where no other diagnosis has been established [10]. However, given the challenges in making an alternate diagnosis, consideration should be given to consolidating the 2 categories to a single category irrespective of whether an alternate diagnosis is made, provided there is documented improvement without tuberculosis treatment (the revised NIH definition now includes a single category of "unlikely tuberculosis" [21]).

A limitation of the current study is that the study population comprised children with relatively severe disease who required hospitalization for suspected pulmonary tuberculosis. Most diagnostic studies in children have been done in hospitalized populations, so validating these criteria is important. However, the large burden of childhood tuberculosis occurs in primary care settings, where fewer diagnostic studies have been undertaken [22]; further validation of the NIH criteria in children in these settings is needed. A further limitation is the exclusion of children for whom a respiratory specimen was not obtained, as microbiological confirmation was required for assignment of a diagnostic category. However, in our study setting, an induced sputum specimen has been reliably obtained in >98% of children with suspected pulmonary tuberculosis [2, 6, 22]. Furthermore, a revision included in the new NIH definition that may strengthen microbiological diagnosis is a positive World Health Organization–endorsed molecular test such as Xpert MTB/RIF (from induced or expectorated sputum, nasopharyngeal aspirate, gastric aspirate, string test, or other intrathoracic sample or stool) for a diagnosis of confirmed tuberculosis.

## CONCLUSIONS

Entry criteria for diagnostic studies should be broad. These results have informed revised NIH consensus definitions in this

supplement that should strengthen collection of standardized data for pediatric diagnostic studies enabling valid, global comparisons [21]. These definitions are a useful tool for ensuring harmonization and comparability between studies; however, re-evaluation and refinement of the revised NIH criteria should be undertaken as further evidence emerges.

## Notes

## References

1. Zar HJ, Connell TG, Nicol M. Diagnosis of pulmonary tuberculosis in children: new advances. Expert Rev Anti Infect Ther 2010; 8:277–88.
2. Nicol MP, Workman L, Isaacs W, et al. Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in hospitalized children in a high HIV-prevalence area. Lancet Infect Dis 2011; 11:819–24.
3. 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.
4. 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.
5. 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; 13:36–40.
6. 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.
7. Hatherill M, Hanslo M, Hawkridge T, et al. Structured approaches for the screening and diagnosis of childhood tuberculosis in a high prevalence region of South Africa. Bull World Health Organ 2010; 88:312–20.
8. Hesseling AC, Schaaf HS, Gie RP, Starke JR, Beyers N. A critical review of diagnostic approaches used in the diagnosis of childhood tuberculosis. Int J Tuberc Lung Dis 2002; 6:1038–45.
9. Cuevas LE. The urgent need for new diagnostics for symptomatic tuberculosis in children. Indian J Pediatr 2011; 78:449–55.
10. 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.
11. Cuevas LE, Browning R, Bossuyt P, et al. Evaluation of tuberculosis diagnostics in children: 2. Methodological issues for conducting and reporting research evaluations of tuberculosis diagnostics for intrathoracic tuberculosis in children. Consensus from an expert panel. J Infect Dis 2012; 205(suppl 2):S209–15.
12. McGinn T, Wyer P, Newman T, et al. Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). CMAJ 2004; 171:1369–73.
13. Zar HJ, Hanslo D, Tannenbaum E, et al. Aetiology and outcome of pneumonia in human immunodeficiency virus-infected children hospitalized in South Africa. Acta Paediatr 2001; 90:119–25.
14. Madhi SA, Petersen K, Madhi A, Khoosal M, Klugman KP. Increased disease burden and antibiotic resistance of bacteria causing severe community-acquired lower respiratory tract infections in human immunodeficiency type 1-infected children. Clin Infect Dis 2000; 31:170–6.
15. McNally LM, Jeena PM, Gajee K, et al. Effect of age, polymicrobial disease and maternal HIV status on treatment response and cause of severe pneumonia in South African children: a prospective descriptive study. Lancet 2007; 369:1440–51.
16. Oliwa JN, Karumbi JM, Marais BJ, Madhi SA, Graham SM. Tuberculosis as a cause or comorbidity of childhood pneumonia in tuberculosis-endemic areas: a systematic review. Lancet Respir Med 2015; 3:235–43.
17. World Health Organization Stop TB Partnership Childhood TB Subgroup. Chapter 4: childhood contact screening and management. Int J Tuberc Lung Dis 2007; 11:12–5.
18. Gilani Z, Kwong YD, Levine OS, et al. A literature review and survey of childhood pneumonia etiology studies: 2000–2010. Clin Infect Dis 2012; 54(suppl 2):S102–8.
19. Zar HJ, Madhi SA, Aston SJ, Gordon SB. Pneumonia in low and middle income countries – progress and challenges. Thorax 2013; 68:1052–6.
20. Bhat N, O'Brien KL, Karron RA, Driscoll AJ, Murdoch DR; Pneumonia Methods Working Group. Use and evaluation of molecular diagnostics for pneumonia etiology studies. Clin Infect Dis 2012; 54(suppl 2):S153–8.
21. Graham SM, Cuevas LE, Jean-Philipe P, et al. Clinical case definitions for classification of intrathoracic tuberculosis in children: an update. Clin Infect Dis 2015; 61(suppl 3):S179–87.
22. 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 using Xpert MTB/RIF on respiratory specimens: a prospective study. Lancet Global Health 2013; 1:e97–104.