Author’s response to reviews

Title: Agreement between QuantiFERON(R)-TB Gold In-Tube and the tuberculin skin test and predictors of positive test results in Warao Amerindian pediatric tuberculosis contacts

Authors:

Lilly M Verhagen (lillyverhagen@hotmail.com)
Mailis Maes (mailismaes@gmail.com)
Julian A Villalba (jvn1987@gmail.com)
Adriana d’Alessandro (adrianamdh@gmail.com)
Lazaro Perez Rodriguez (lachy8223@gmail.com)
Mercedes F España (mefiesce@gmail.com)
Peter WM Hermans (Peter.Hermans@radboudumc.nl)
Jacobus H de Waard (jacobusdeward@gmail.com)

Version: 2 Date: 13 April 2014

Author’s response to reviews:

Reviewer 1:
Reviewer’s report
Major Compulsory Revisions
General comments--This is an ambitious evaluation comparing the use of QuantiFERON-TB Gold In-Tube with TST for the diagnosis of latent TB infection (LTBI) and TB disease in Warao Amerindian contacts of TB patients. While the paper presents useful data and hypotheses, it will require significant re-working for finding and conclusions to be clearer.

Specific comments

Abstract: should be clear if the comparison is the use of these tests to diagnose LTBI, TB disease, or both. If results are not statistically significant (p=0.12, they should be described as similar). Suggest adding proportions and p-values to all results.

Response to specific comment Reviewer 1:
In order to clearly state our primary objectives in the abstract, we changed:

'We investigated the performance of the QuantiFERON®-TB Gold In-Tube (QFT-GIT) and TST in a prospective cohort of Warao Amerindian children in Venezuela.'

into
'We determined the prevalence of Mycobacterium tuberculosis infection by TST and QuantiFERON®-TB Gold In-Tube (QFT-GIT) and assessed agreement between the two test methods and factors associated with positivity in either test in Warao Amerindian children in Venezuela. Furthermore, progression to active TB disease was evaluated for up to 12 months.'

Furthermore, we deleted the paragraphs describing the results of the analyses assessing characteristics of CXR lesions and conversions/reversions of QFT-GIT results from the abstract, since this was not part of the primary objectives of our study. In addition, following this reviewer's suggestion, we added proportions and p-values to all results and we changed 'At baseline, the proportion of TST positive children was higher than the proportion of children with a positive QFT-GIT (47% vs. 42%, p=0.12).’ into 'At baseline, the proportion of TST positive children was similar to the proportion of children with a positive QFT-GIT (47% vs. 42%, p=0.12).'

1. Authors need to make very clear whether the outcome being compared is test performance to detect LTBI, TB disease, or both

Response to comment 1 Reviewer 1:

We agree with this reviewer that it should be clear what the outcome being compared is. The primary objectives of our study were 1) to determine the prevalence of M. tuberculosis infection by TST and QFT-GIT and assess agreement between the two test methods and 2) to determine the characteristics associated with TST positivity and QFT-GIT positivity in childhood TB contacts. Secondary objectives were 1) the performance of TST and QFT-GIT to detect TB disease and predict the development of TB disease and 2) to determine the prevalence of and characteristics associated with CXR lesions indicative of TB sequelae.

In order to make this clear for the reader, we changed the paragraph:

'We assessed the performance and concordance of the QFT-GIT and the TST in a prospective study of Warao pediatric TB contacts. In addition, we assessed the relationship between QFT-GIT and TST results and radiological features indicating the presence of TB sequelae.'

into:

'The primary objectives of our study were 1) to determine the prevalence of M. tuberculosis infection by TST and QFT-GIT and assess agreement between the two test methods and 2) to identify characteristics associated with TST positivity and QFT-GIT positivity in childhood TB contacts. Secondary objectives were the assessment of the QFT-GIT and TST performance to detect TB disease and predict the development of TB disease and their relationship with radiological features indicating the presence of TB sequelae.'

2. Later it is clear that Venezuelan guidelines define a positive TST as 10mm, but
this should be clear in the methods (as some programs use a 5mm cutoff for contacts). Also, definition of a positive and indeterminate QFT should be clearly included in the methods (e.g. under Definitions).

Response to comment 2 Reviewer 1:
We had already stated that the Venezuelan guidelines define a positive TST as 10 mm and we had already provided the definitions of QFT-GIT conversion or reversion QFT-GIT in the methods section under definitions:

'Venezuelan National TB Program guidelines regard a TST #10 mm 48-72 hours after injection as positive. A TST conversion was defined as having a negative TST at baseline and a TST #10 mm at any follow-up time point. QFT-GIT conversion or reversion was defined as respectively a negative to positive or a positive to negative change according to manufacturer’s criteria (ie, baseline IFN-# <0.35 IU/ml and follow-up IFN-# #0.35 IU/ml or baseline IFN-# #0.35 IU/ml and follow-up IFN-# <0.35 IU/ml respectively) [24].'

Following the Reviewer's suggestion, in the revised manuscript, we also stated the definition of a positive and indeterminate QFT-GIT in the following sentence which was added to the paragraph cited above:

'A positive QFT-GIT was defined as a value of TB antigen minus nil (negative control) value #0.35 IU/ml and #25% of nil value. A nil value of >8.0 IU/ml or a mitogen minus nil of <0.5 IU/ml was classified as indeterminate.'

3. Define “negative TST”—is it 0mm of induration, or could it include results <10mm?

Response to comment 3 Reviewer 1:
We added the following sentence to the Definitions section in the Methods:

'An induration of <10 mm was considered TST negative.'

4. Statistical analysis: were any power calculations done in advance to determine what sample size would be required for a statistically significant difference in study population?

Response to comment 4 Reviewer 1:
Rather than performing a power calculation in advance, we included all childhood household contacts of all adult patients from the municipalities Antonio Díaz and Pedernales in the Orinoco Delta, provided that they did not meet the exclusion criteria mentioned in the Methods section. Because sample size and power calculations are based on predictions of what the study results will be, they are particularly useful as tools to assist in planning. However, these predictions require some knowledge concerning the prevalence of the disease or condition of interest and its determinants. If there is a lack of knowledge on these factors in the population of interest, such as is the case for CXR lesions indicating TB
sequelae, conversion or reversion of QFT-GIT and TST or QFT-GIT positivity in active TB in many South American indigenous populations, explorative studies are necessary. Therefore, our study is a first attempt to obtain insight into these factors in order to use these insights in power calculations to assist in the planning of future clinical epidemiological studies.

5. If QFT and/or TST results are used in the case definition for TB, this presents a problem for assessing their ability to predict TB disease (as the authors point out on p.9)

Response to comment 5 Reviewer 1:
This is a limitation of our study that was indeed acknowledged in the manuscript.

6. The discussion of CXR findings is also difficult to follow. Suggest more clearly defining what is meant by “TB sequelae” and other terms (either in methods or elsewhere, and not only in discussion)

Response to comment 6 Reviewer 1:
We had already defined what is meant by TB sequelae in the methods section under 'Definitions':

'CXRs were classified as (1) normal (no abnormalities suggestive of current or past TB observed), (2) active TB or (3) radiographic lesions possibly related to past pulmonary TB. The latter category was defined as the presence of at least one of the three radiological features (calcification, parenchymal destruction with fibrosis, bronchiectasis) defined as 'consequences of previous pulmonary tuberculosis' by Marais et al [28]. CXR findings regarded as compatible with bronchiectasis were saccular changes or cylindrical outlines of airways that widened as airways extended into the lung periphery [27, 28].'

In order to make it more clear that the 'radiographic lesions possibly related to past pulmonary TB' described here are later described as 'TB sequelae', in the revised version, we added the word 'TB sequelae' to the description cited above in the Methods section.

7. Logistic regression—no information was really provided on how this model was constructed (and no data are shown in table form), so it is very difficult to interpret these findings

Response to comment 7 Reviewer 1:
Table 2 shows the results of the multivariable GEE analyses for characteristics associated with QFT-GIT and TST positivity and table 5 shows the results of the multivariable GEE regression analysis for the presence of TB sequelae on chest X-rays. All variables assessed were included in the multivariable model and no entry-criterium based on p-values in univariate analysis were used, ie, all clinical and other relevant variables as outlined in tables 2 and 5 were included in the multivariate models. We included all variables in the multivariable models
regardless of their significance in univariate analysis because they all formed part of the study objectives.


'Generalized estimation equations (GEEs) were used to fit a multivariable logistic regression model aimed at identifying possible associations between QFT-GIT or TST positivity (dependent variable) and BCG vaccination, age, sex, malnourishment and duration of exposure (independent variables) and between the presence of TB sequelae on CXR (dependent variable) and IFN-# level, TST induration, BCG vaccination, age, sex, malnourishment and duration of exposure (independent variables). GEEs account for correlation and lack of independence of responses for contacts with an index TB case in common (clusters within households).'

8. Discussion: several studies have now shown that QFT has poor predictive value as a longitudinal tool. As this study (along with several others cited) has a relatively small sample size (especially when not all results are available on all children over time). This is an important limitation and thus comparisons of PPV and NPV, sensitivity and specificity across studies should be interpreted with great caution

Response to comment 8 Reviewer 1:
We added this limitation to the Discussion section:

'Furthermore, almost 30% of the study subjects in the German study moved away from the study region during the observation period of 2 years and no follow-up data were available for these participants. The proportion of children that was lost to follow-up in our study was even higher (62%) and estimations and comparisons of positive and negative predictive values should thus be interpreted with caution.'

9. CXR findings of calcification and bronchiectasis. As authors point out, these may be due to other factors (e.g. histoplasmosis for calcifications). As there is no comparison group (e.g. children who were not TB contacts), these findings (and their association with TB exposure or infection) should be interpreted with caution

Response to comment 9 Reviewer 1:
We added this limitation to the Discussion section:

'Studies including a control group of children without a TB contact would be able
to distinguish TB-related CXR lesions from CXR findings related to other pathologies. However, in a high burden setting where under-diagnosis and under-registration of cases occur it is difficult to identify such a comparison group.’

Minor Essential Revisions
Suggest changing “recommendable” to “recommended”

Response to minor revision Reviewer 1:
We replaced 'recommendable' by 'recommended'.

Introduction:
1. In general, the introduction is very long and should focus more on the specific comparisons of interest. Other relevant information might better fit in the discussion

Response to minor revision comment 1 Reviewer 1:
We agree with the reviewer that the introduction was very long with too little focus. Therefore, we deleted the third paragraph as well as the background information on the difficulty of diagnosing active TB in children. Furthermore, we deleted repetitive parts, further shortening the introduction.

2. Rapid disease progression is especially a concern in young children (less than 3 years of age), so consider adding “young”

Response to minor revision comment 2 Reviewer 1:
Following this Reviewer's suggestion, we added "young" to the sentence stating that rapid progression is a unique aspect of TB in children.

3. Discussion of CXR findings is a bit confusing and should relate to the methods and findings of the study. Not sure if authors mean “hilar adenopathy” when they refer to “regional adenopathy”?

Response to minor revision comment 3 Reviewer 1:
We agree with this reviewer that the discussion of CXR findings in the Introduction was a bit confusion. Since the value of CXR in the diagnosis of active TB disease was not part of our study objectives, we deleted this paragraph, which also considered the sentence containing 'regional lymphadenopathy'. By 'regional' we meant 'hilar and paratracheal', which are the most common diagnostic features expected (Swingler GH. Arch Dis Child 2005;90:1153-6), and we agree with the reviewer that we should have specified this, had we kept this statement. The paragraph that was kept introduces one of the secondary objectives of our study: the assessment of the relationship of QFT-GIT and TST results with TB sequelae on CXR.

4. 6th paragraph, suggest changing “extraordinary” to “extraordinarily”
Response to minor revision comment 4 Reviewer 1:
Following this Reviewer's suggestion, we changed "extraordinary" into "extraordinarily".

5. To understand potential bias in the sample, it would be useful to know if these were all child TB contacts during the time period, or what % and how many refused to participate

Response to minor revision comment 5 Reviewer 1:
The included contacts were all child TB contacts of TB registered TB patients in the Pedernales and Antonio Díaz municipality. Six patients that were registered were untraceable and in one household, parents were not willing to enroll their children in the study. However, due to the under-diagnosis and under-registration of TB in this area, it is difficult to state whether these were all child TB contacts during the time period.

We added this information to the Results section in:

'Six registered adult TB patients were untraceable and in one household, parents were not willing to enroll their children in the study.'

Furthermore, we added the difficulty of estimating the exact number of child TB contacts to the limitations section in the Discussion:

'Finally, although we included all child TB contacts of registered TB patients that were encountered and willing to participate, we cannot rule out the possibility that child TB contacts were missed due to under-registration of TB cases in the Orinoco Delta Department of the Venezuelan National TB Control Program.'

6. See consensus definition of childhood TB cases from NIH and whether it can be applied to these cases

Response to minor revision comment 6 Reviewer 1:
The definitions we used are slightly different from the JID NIH definitions. In May 2012, Graham et al proposed clinical case definitions for standardizing the classification of intrathoracic tuberculosis in children, for research evaluating new diagnostic tests in children in whom tuberculosis is suspected but rarely confirmed. Our study started in May 2010, before the publication of the NIH definitions, and we had then already defined our definitions. However, differences were minimal and included the fact that we did not take documented exposure to M. tuberculosis as a possible criterium (since all children included in our study were exposed to an infectious adult) and the fact that we did not take into account a positive clinical response to anti-TB treatment (since the proportion of children that are lost to follow-up after initiation of TB treatment is high as was observed in our study). Our CXR classification was based on the radiologic classification proposed by Marais et al for diagnosis of intrathoracic tuberculosis in children up to 15 years of age (Marais BJ et al. Pediatr Radiol
Possible TB was defined as the presence of abnormal findings on chest radiograph that did not meet the radiological criteria for probably TB as classified by Marais et al, a definition also used in the study of Cohn et al (Pediatr Infect Dis J 2009) who also based TB diagnosis on the criteria of Marais BJ et al. The descriptive terminology for abnormal radiological features proposed by Marais et al was also cited in the NIH definitions.

However, as the JID NIH definitions focus on infants and children < 10 years of age, it is questionable whether these definitions would be the most suitable for our study population aged 0 to 15 years of age. In addition, the purpose of the NIH definitions was to use these case definitions in studies validating new diagnostic methods, these definitions were not designed for use in household contact evaluation or clinical management.

7. BCG vaccination status. Not all children with documented vaccination develop a scar, particularly in settings of immunosuppression (e.g. HIV, malnutrition). Consider adding this as a limitation, as you did not examine BCG records

Response to minor revision comment 7 Reviewer 1:
Following this Reviewer's suggestion, we added the following paragraph to the limitations section in the Discussion:

'Study subjects with a BCG scar were reported as BCG vaccinated and vaccination cards were not examined. This may have led to under-reporting of the number of children with BCG vaccination since some BCG vaccinated children may not develop a scar. However, household-retained vaccination cards have also been identified as an insufficient source of information for estimating vaccination coverage in other areas [61].'

8. Suggest describing as TB treatment (not anti-TB)

Response to minor revision comment 8 Reviewer 1:
We replaced 'anti-TB treatment' by 'TB treatment' throughout the manuscript.

9. Suggest avoiding the terminology t=6 or t=12, and instead writing “at six-month follow-up”, etc.

Response to minor revision comment 9 Reviewer 1:
We followed up this Reviewer's suggestion and made these changes throughout the manuscript.

10. Suggest rephrasing as “reported or documented fever”. Methods imply that a physical exam was part of all contact investigations—this should be made clear

Response to minor revision comment 10 Reviewer 1:
Following this Reviewer's suggestion, we rephrased 'parental history or clinical signs of' as 'reported or documented'.
A physical examination was indeed part of all contact investigations. We had stated this in the sentence: 'To this end, a TST and clinical examination were performed in all encountered children.' For clarity, in the revised manuscript, we replaced 'clinical' by 'physical'.

11. As not all tests were available on all children, it becomes difficult to follow the denominators for the results. Suggest making this clearer in the figure and in the text.

Response to minor revision comment 11 Reviewer 1:
Figure 1 was revised and percentages of available tests were added. Additionally, we added calculations of numbers to the text, eg, '163-57' to show how numbers were derived from the total number of children included.

12. Suggest adding in the % of indeterminate results (e.g. for QFT)

Response to minor revision comment 12 Reviewer 1:
The percentages of indeterminate results for QFT-GIT were added to Table 1.

13. Was the child with “hip bone TB” bacteriologically confirmed? If so, suggest classifying as culture positive

Response to minor revision comment 13 Reviewer 1:
This child had intrathoracic lesions compatible with probable TB (not bacteriologically confirmed) and had clinical signs of as well as radiological evidence of hip bone TB (not bacteriologically confirmed). We described this child separately ('One child was diagnosed with intrathoracic and hip bone TB') rather than adding it to any of the other categories (confirmed, probable or possible intrathoracic TB).

14. As the authors do not really describe findings related to malnutrition (other than prevalence in this population of contacts) or other immunologic parameters, the inclusion in the discussion seems out of place

Response to minor revision comment 14 Reviewer 1:
We agree with this reviewer and deleted the studies discussing malnutrition and immunologic parameters from the Discussion. We only kept the studies on BCG vaccination and lung lesions, since these may provide an explanation for the significant association between BCG vaccination and CXR lesions of possible TB sequelae we observed.

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Needs some language corrections before being published
Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests

Reviewer 2:

General comments on manuscript:
This is an interesting study completed in a S American population about which there is relatively little information in the literature, despite the reportedly very high rates of childhood TB. The most important finding (though not surprising) is the fact that children exposed to a SSP/Cx proven TB case really do need IPT and this data should be fed back to the Venezuelan national Tb programme.

Response to general comment Reviewer 2:
We have reported these data to the Venezuelan National TB Programme. In fact, one of the co-authors of our study is the current head of the Venezuelan National TB Programme.

I have a few general comments:
1. the title is not very specific – what are the authors comparing with the TST and QFG? Suitability for detecting infection, detection of disease, progression to disease, comparison with CXR findings? All of these factors are incorporated into the analysis and yet the title gives no indication of this.

Response to general comment 1 Reviewer 2:
In order to clearly state our primary objectives in the titled, we changed the title into:

'Agreement between QuantiFERON®-TB Gold In-Tube and the tuberculin skin test and predictors of positive test results in Warao Amerindian pediatric tuberculosis contacts'

2. Likewise the abstract needs to state more clearly what the aims and objectives were. Otherwise it looks as tho’ this was a cohort used for another study and in whom someone thought it was a good idea to screen all childhood TB contacts with a TST and QFT and CXR wihtout really thinking about how were going to analyse the data. Currently this has the feel very much of a post-hoc analysis.

Response to general comment 2 Reviewer 2:
We agree with this reviewer that it should be clear what the aims and objectives were. The primary objectives of our study were 1) to determine the prevalence of M. tuberculosis infection by TST and QFT-GIT and assess agreement between the two test methods and 2) to determine the characteristics associated with TST
positivity and QFT-GIT positivity in childhood TB contacts. Secondary objectives were 1) the performance of TST and QFT-GIT to detect TB disease and predict the development of TB disease and 2) to determine the prevalence of and characteristics associated with CXR lesions indicative of TB sequelae.

In order to clearly state our primary objectives in the Abstract, we changed:

'We investigated the performance of the QuantiFERON®-TB Gold In-Tube (QFT-GIT) and TST in a prospective cohort of Warao Amerindian children in Venezuela.'

into:

'We determined the prevalence of Mycobacterium tuberculosis infection by TST and QuantiFERON®-TB Gold In-Tube (QFT-GIT) and assessed agreement between the two test methods and factors associated with positivity in either test in Warao Amerindian children in Venezuela. Furthermore, progression to active TB disease was evaluated for up to 12 months.'

3. Many features are evaluated in this manuscript making for a very long paper with many inconclusive findings. Could this not have made two papers with a separate one for the comparison of CXR lesions with QFT-IT results? It would certainly make the paper more succinct. Alternatively, the QFT and TST do not really add anything new to the already very extensive literature on this area of paediatric TB diagnostics and could potentially be left out all together, whereas the CXR vs QFT analysis is relatively novel.

Response to general comment 3 Reviewer 2:

Although QFT-GIT and TST performance have been evaluated in many other studies from other parts of the world, the influence of previous BCG vaccination on TST outcome varies between geographical regions and had not been investigated in South American indigenous populations. Second, the QFT-GIT could be of specific value in the effort to decrease under-diagnosis of TB disease in children in this area as outlined in the Introduction section. However, no studies to estimate the value of QFT-GIT in TB disease diagnosis in children in South American indigenous populations had taken place. In the revised manuscript, we significantly shortened the manuscript text by leaving out parts of the Introduction and Discussion that were not related to our research objectives in response to the comments of this and the previous Reviewer.

Discretionary Revisions
1. The introduction is well written but overlong, repetitive in parts and would benefit from being more concise

Response to Introduction comment 1 Reviewer 2:

We agree with the reviewer that the introduction was very long with too little
focus. Therefore, we deleted the third paragraph as well as the background information on the difficulty of diagnosing active TB in children. Furthermore, we deleted repetitive parts, further shortening the introduction.

Methods:
1. repetitive and overlong. Would benefit from shortening though needs some further clarification on specifics as detailed below.

Response to comment 1 Reviewer 2:
We shortened the Methods section in order to make it easier to read.

3. Was the cohort actually prospectively screened for HIV as part of this study? It's not entirely clear from the methods. It would be good to know what the HIV prevalence is estimated to be in this population.

Response to comment 3 Reviewer 2:
The cohort was indeed prospectively screened for HIV. However, due to the high number of children that was lost to follow-up, it is difficult to estimate the HIV prevalence from this study. More reliable estimates of HIV prevalence in this population are published in Villalba JA et al. AIDS 2013;27:1783-91 and Rangel HR et al. PLoS ONE 2012;7:e40626. We added the information on HIV testing to the methods in:

'HIV testing was performed upon inclusion and at 12-month follow-up in all children.'

4. It would be clearer to give detailed description of F/U timepoints and illustrate all of this with results of recruitment and classification in a consort style flow diagram. Figure 1 is not clear and should be revised.

Response to comment 4 Reviewer 2:
Children that were not encountered at six-month follow-up but were encountered at 12-month follow-up were still included in the 12-month follow-up analyses. The flow diagram is therefore not a continuum, but, rather, 2 groups (six-month follow-up and 12-month follow-up) emerge from each inclusion group. We agree that Figure 1 was not clear. We think that it would be more clear if the children initially diagnosed with TB disease are separated from those initially classified as asymptomatic, since the analyses described in the manuscript generally focus on one of these two groups. Therefore, we re-designed Figure 1. Also, in the revised figure, we included percentages of children in which the examinations were performed.

5. Where was the QFG-IT assay completed – on site or transported to a lab elsewhere? If transported elsewhere what was the time delay before getting into an incubator?
Response to comment 5 Reviewer 2:
The QFT-GIT assay was completed on site with a portable incubator. After completion of the incubation time, sera were separated and frozen at -20°C for a maximum of 8 weeks until ELISA’s were performed in a laboratory in Caracas. So, samples were transported after completion of the incubation time. We added this information to the Methods section.

6. I wonder why the authors didn’t use SSP Tb cases (culture confirmed) for screening contacts since this might have increased their yield of secondary paediatric cases.

Response to comment 6 Reviewer 2:
We agree with this Reviewer that limiting our analyses to smear-positive TB cases could have increased the relative proportion of secondary pediatric cases. However, approximately 10-20% of TB transmission is due to source cases with smear-negative pulmonary TB (Tostmann A et al. Clin Infect Dis 2008;47:1135-42, Behr MA et al. Lancet 1999;353:444-9, Hernandez-Garduno E et al. Thorax 2004;59:286-90). We decided to include all culture confirmed source cases registered in the study region in order not to miss any potential TB cases.

7. It’s surprising that the mean age of secondary cases in the initial (T0) screen was relatively high given that one would expect the highest rates to be in young children. No discussion of this finding is included in the discussion. Similarly the large gender difference is surprising in this pre-pubertal age group but may simply reflect the fact that the numbers are too small to draw any meaningful conclusions.

Response to comment 7 Reviewer 2:
As the Reviewer points out, numbers of active TB cases at initial screen (n=11) are indeed small and, therefore, it is highly likely that these differences in age and gender between TB cases and asymptomatic children are a result of chance alone and statistical analyses on such small numbers are therefore not reliable. Considering this, we removed this statistical analysis from the Results section.

8. A table summarizing key findings eg median/mean age, gender diffs, nutritional status, % TST positive by age, as well as breakdown of Tb cases by age would have been easier to view rather than reading in text.

Response to comment 8 Reviewer 2:
We added a table summarizing key findings classified by age category (Table 1 in the revised manuscript) and we deleted this information from the text.
9. Eleven children were diagnosed with TB but details of the site of TB is only given for 9! What was the site of the two Cx proven cases?

Response to comment 9 Reviewer 2:
The site of the two Cx proven cases was pulmonary TB. We rephrased the sentence

'At inclusion, 11 children were diagnosed with active TB, of which two were confirmed, four were probable and four were possible intrathoracic TB.'

into:

'At inclusion, 11 children were diagnosed with active TB, of which two were confirmed pulmonary TB, four were probable and four were possible intrathoracic TB.'

10. Why exclude indeterminate IGRA results when calculating sensitivity and specificity? Does this not result in an over estimate of the sensitivity.?

Response to comment 10 Reviewer 2:
The exclusion of indeterminate IGRA results when calculating sensitivity and specificity was done because there is no TST outcome equivalent to an indeterminate QFT-GIT result. An indeterminate result could be due to a laboratory error (and thus be equivalent to a 'not performed' IGRA) or to anergy and a lack of response to the antigen (and thus probably equivalent to a negative TST) or to high background interferon gamma levels. Other studies have also excluded IGRA results when these were indeterminate from statistical analysis for sensitivity (Nicol MP et al. Comparison of T-SPOT.TB assay and tuberculin skin test for the evaluation of young children at high risk for tuberculosis in a community setting. Pediatrics 2009;123:38-43, Kampmann et al. Interferon-# release assays do not identify more children with active tuberculosis than the tuberculin skin test. Eur Respir J 2009;1374-82.). To get an idea of the influence of excluding indeterminate results on sensitivity, we re-calculated the sensitivity of QFT-GIT in our study after grouping the indeterminate with the negative results and the resulting sensitivity was 78% (95% CI 40-97%). This is indeed lower than the identified sensitivity of 88% but still higher than the sensitivity of TST testing (55%).

11. How were the TSTs read? Needs documenting.

Response to comment 11 Reviewer 2:
We added the following sentence to the Methods section:

'Reading was performed by trained professionals measuring the palpable transverse induration on the volar surface of the forearm between 48 and 72 hours after administration.'

Results Section:
1. Progression to active TB during F/U- what was the diagnostic certainty and the site of disease – PTB or EPTB? – it would be helpful if this was detailed.

Response to comment 1 Reviewer 2:
Following this Reviewer’s suggestion, we added ‘pulmonary’ to the sentence ‘The mean age of the five children that developed active TB during follow-up was lower than the mean age of all included children (respectively 4 vs. 8 years).’ Table 3 in the previous manuscript (which is table 4 in the revised manuscript) provides the reader with detailed information on the demographics and diagnostic test results, including CXR findings, for the five contacts who progressed to active TB.

2. The change in TST and QFT results from neg to positive with time may relate to on-going TB transmission within the community. Without knowing if there were further adult SSP TB cases within the households included in the study it is difficult to draw any conclusions about these findings.

Response to comment 2 Reviewer 2:
We agree with this Reviewer that ongoing transmission, either in the household or in the community, may have caused the change in TST and QFT-GIT results from negative to positive with time and we addressed this in the Discussion of the revised manuscript by adding the following sentence:

'Notably, we cannot exclude the possibility that TST and QFT-GIT conversions were related to ongoing transmission within the community.'

3. TST and QFT-GIT results during F/U:
In the final sentence re QFT recruitment responses, there seem to be quite a lot of negative responses. While this is obviously due to high IFNg levels in the negative control tube it does make it difficult to interpret the data (range -7.6 to 0.17).

Response to comment 3 Reviewer 2:
The negative response was indeed due to high IFN-gamma levels in the negative control tube. The child with a negative response of -7.6 IU/ml was, however, the only extreme outlier. Apart from this child, there was 1 child with a negative response of -0.372 whereas all other children had responses #0.139.

4. Overall there appears to be a lot of data but relatively small numbers at follow-up. I am therefore concerned that the very inconclusive findings re conversions and reversions relates to the fact that this study is underpowered to show any significant and meaningful trends.

Response to comment 8 Reviewer 1:
We had already added the limitation of the limited number of children with active follow-up data to the Discussion section. Although this is an important limitation, we still believe that our data are valuable. No studies assessing TST and QFT-GIT performance in South American indigenous populations have been carried out, while indigenous people comprise around 10% of the population. In addition, upon inclusion, TST, CXR and QFT-GIT results of virtually all children were available and a large part of the analyses were based on these inclusion results. Finally, passive follow-up was also performed by 4 of the authors residing in the study area at the time of research and by revising the TB register of the Orinoco Delta Department of the National TB Control Program. In the revised manuscript, in addition to the acknowledgement of this limitation in the last paragraph of the Discussion, we added a section specifically stressing the limitation related to the interpretation of QFT-GIT data in:

'Furthermore, almost 30% of the study subjects in the German study moved away from the study region during the observation period of 2 years and no follow-up data were available for these participants. The proportion of children that was lost to follow-up in our study was even higher (62%) and estimations and comparisons of positive and negative predictive values should thus be interpreted with caution.'

Level of interest: An article of limited interest
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:
I declare that I have no competing interests