nature portfolio

Peer Review File

Mycobacterium tuberculosis Cough Aerosol Culture Status Associates with Host Characteristics and Inflammatory Profiles

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This is a very interesting paper, looking to identify markers of infectiousness of individuals with TB. The study is centred around the cough aerosol detection mechanism, and explores it's validity and potential markers.

While this is interesting data, which should be certainly be published and feed into the wider discourse around Mtb transmission, there are a number of issues which make me unsure if the claims in the title, abstract and certain section headings are backed up sufficiently by the data. Some of these are clarifications and could be fine, others have to do with the study design and analytical approach, which may be harder to overcome.

A major issue is lack of a control group of HH contacts, which would help to understand how the 53% IGRA+ in the CAC- population should be interpreted. This means that the findings can reflect on markers for 'higher' infectiousness, as suggested by CASS positivity.

A key gap in the current methods is a lack of description of how the shift from 142 participants with TB to including 55 index cases in the HH contact study. It is mentioned in the results section, but is not described at all. Similarly, n=58 index cases were included in the transcriptomics, but their selection is not described. That needs to be clarified (see more detailed comments below).

The analysis approach for the predictive models is a simple forward selection of variables. In the context of this well studied area, this seems too uninformed. The consequence is some strange findings which the authors do not pick up on. E.g. Why would MUAC be predictive but BMI would not? How should we interpret CASS predictive model mean where a dichotomised age variable is given equal importance as the sole bacteriological indicator (Xpert CT threshold)? While an uninformed approach to the cytokine/transcriptomics analysis seems OK given its novelty, this is not the case for associations for transmission or infectiousness, and this should be addressed, e.g. by discussing a conceptual framework, and exploring different predictive models with a-priori choices.

The more conceptual issue is what space of infectiousness this paper addresses. The title, abstract and body text shift between scope. For example, the title and abstract reflect on 'TB infectiousness', suggesting it is about individuals that are either infectious or not. As clarified later in the abstract, the authors actually focus on a differentiation between more or less infectious individuals, based on CASS. They show that CASS+ individuals are likely more infectious with a nested household contact study (which is not a new finding), but there is no control group, so no statement about a binomial infectious yes/no status. For the transcriptomics they then look in a subgroup (which is not defined, see above), and find a number of associations with CASS+. I would agree that this provides data to 'suggest host inflammatory signatures are associated with Mtb aerosolization independent of bacillary load and cavitary lung disease.', as stated in the discussion. But what it does not show is 'Tuberculosis Infectiousness is Associated with Distinct Clinical and Inflammatory Profiles', which is the paper title, and as such much too broad. I think the authors should be more clear, and e.g. change the title, and any

other statements regarding ‘infectiousness’ as a whole to reflect the more narrow scope.
-

I have provided extensive detailed comments below, which hopefully will help improve the paper when the data is published here or in another journal.

Methods:
While I understand why ignored trace+/culture- (given the low chance of being CASS pos), it would be good to state this explicitly, and state how the authors interpreted this group. (TB or not TB)?

There is some reporting of results for BMI vs MUAC in the methods, this should be moved to the results.

L314-318 is results I think, should move there.

It is not obvious from the methods that the HH contacts came from a subset (n=55, l352) of enrolled participants from the total population (n=142). I did not see it in l185-l189), nor is it mentioned in l142-l144. Firstly, this should be made clear more clear. Secondly the rationale and method for this selection should be described explicitly (e.g. funding restrictions and random draw?). There is a similar selection for the host transcriptomic work (going down to n=59 index patients), which needs to be similarly explained.

RESULTS
335: It seems odd that all of the participants were cough positive, as a percentage came from community screening. Given the difference in population and expected disease severity between ACF and PCF populations, these two populations should be described more in the text and supp materials. If the 100% cough positive holds, it is something to discuss, as it goes against most modern TB prevalence surveys where about 50% were found to be cough negative.

Table 1: The MUAC result is odd (i.e. a higher MUAC means more CASS+). It seems counterintuitive, and the lack of an association with BMI makes me concerned the MUAC results are less reliable. The BMI/MUAC discrepancy should be discussed.

Table 2: I am not clear what the OR and p-values are meant to show here. For the index case characteristics, those seem a near 1-2-1 repeat of table 1? I think those could be dropped, especially from the main body text. For e.g. the number of contacts per case – I think this is just a chi-square, but that is not the relevant finding. What we would like to know is whether the distribution of HH contacts is different by CAC+ and CAC-, which is currently unclear.

What were reasons for missing QFT (N goes from 143 to 127)? These should be reported.

Table 3: The results are quite stark, but not sure how superior the final model is. It makes sense that after one good measure of bacteriological burden is included in the model, the others are no longer adding enough value to be included. E.g. smear and cavitation seem to be quite useful explanatory/predictive variables by themselves. A simple forward stepwise modelling approach will not

highlight those limitations. Given what we already know about infectiousness, this type of analysis would be better served with a formal DAG or at least conceptual framework, and then working through the various options or models. I suggest the authors explore this wider, and thereby enrich the insights that can be gained from the data.

While the AUC for the scoring system for predicting CAC+/- is high, I don't quite understand the results from Sup table 1. It seems to suggest that age and Xpert CT are equally valuable (both 7 'points')? But what does the model/score mean if the sole bacteriological indicator is removed? Are the authors suggesting that we could decide on CAC+/- by age, CRP and MUAC alone? Could they provide an indication of how much value each individual characteristic adds, and how the model works if Xpert CT is not available. Could any other bacteriological indicator be switched in? I'd be surprised if the model functioned without a bacteriological indicator, and also not sure about the biological logic that would remain. It is also not clear if this predictive model is solely applicable for people who have bacteriologically positive TB? If so, that should be made clear.

The variation in association pattern of the different cytokines is a bit strange. Would one not expect them to be elevated across the various indicators, also given that e.g. CT threshold and cavitation were themselves associated with CAC? The lack of stability makes me a bit concerned. The results are what they are, but potential explanations (including random noise across many comparisons that could affect the threshold for interpretation of p-values).

DISCUSSION
L455-457: these are results, and should be moved there. On a broader note, it would be good to place those observations in the context of the ongoing discussions around subclinical TB (see previous comment re cough positivity).

L462: the statement about 'body mass' is strange, as BMI was not associated, but MUAC was. As mentioned before, this discrepancy needs discussion, and the statement should be adjusted accordingly.

L473: this only interprets one of the cytokine findings, but ignored the conflicting results in the actual section (see above), where some cytokines were associated for one indicator, but not the other. Do the authors assume only 1 indicator is valid? If so, that is a stretch. If not, the discrepancies need to be addressed.

L535: the suggestion that interventions (i.e. treatment) should target the most infectious is a bit strange. This approach was tried with the development of DOTS policies, where symptomatic smear positive individuals were targeted through PCF and 6m therapy. While saving lives, it has not interrupted transmission as expected, and it is not clear to me why this would work now? Shouldn't Identifying and treating all infectious individuals be the minimum target for all high burden countries, regardless of how infectious they are at the time of diagnosis?

Furthermore, infectiousness is not static, and will likely vary over time. Finally, given the high Number Needed to Test to identify any person with low or high infectious disease, the cost of treatment is really not the limiting factor. It would be helpful if the authors reflected a bit more on this (also in the

introduction).

Minor points
Ref 33, 34 and 58 are the same I believe. Would be good to recheck references and make sure in order.

Table 1: category ‘Enrolled thru prevalence survey’ – not appropriate language for scientific paper and not reflected in main text. Suggest something along lines of ‘ACF-enrolled’, as this is how it is described in methods.

Table 3: the unit for Age is not given. From the OR I suspect it is year of age, but this should be made clear. Also note this assumes a linear association across the entire age range, which is unlikely to be true given what we know about age-dependent population ARI. This should be discussed.

The section titles on l397 and l409 are phrased as interpretation rather than reporting of results. These should be adjusted. In addition, the last sentences in each of these sections is interpretation rather than result, and should be removed or put in discussion.

Reviewer #2 (Remarks to the Author):

The study "Tuberculosis Infectiousness is Associated with Distinct Clinical and Inflammatory Profiles" stands out for addressing a significant gap in tuberculosis research, focusing on individual infectiousness and biomarkers. It employs a prospective longitudinal study design, examining treatment-naïve pulmonary TB patients in Nairobi, thereby providing valuable insights in a high TB burden setting. Methodological innovations, such as using the Cough Aerosol Sampling System (CASS) for measuring Mycobacterium tuberculosis aerosolization, the evaluation of sputum cytokine, and assessing household contacts through IGRA testing, add important epidemiological and molecular dimensions and adds novel knowledge and tools to predict Mtb transmission. Overall, the manuscript was well written. The incorporation of suggested improvements would enhance the paper's quality and ensure that concerns have been adequately addressed.

#1 Introduction: The introduction section is well-written and effectively frames the context of TB infectiousness to its clinical and immunological aspects. It successfully outlines the importance of understanding individual variations in TB transmission and reviews relevant literature, thus situating the study within the broader field of TB research. I have no suggestions for improvement in this section.

#2 Methods: The methods section is well-structured and thorough. However, consider breaking down the dense text into smaller paragraphs to improve readability (i.e. break the paragraph in line 262. “We assessed....”. In addition:
#2.1 – Methods (Line 203-204): The study does not address the issue of drug-resistant TB. On lines 203-204, the study mentions classifying rifampicin resistance as "positive," "negative," or "indeterminate,"

but this is not further elaborated in the text. I recommend including these data in the results and tables, even if 100% were susceptible, to provide a complete picture of the resistance profile among the study participants.
#2.2 - Results (lines 222-228): It would be beneficial to provide a rationale for the decision to assess cytokines in sputum instead of blood; as well as the selection of the specific three cytokines studied. This is particularly important for readers who may have less experience in the field. Including one or two lines to explain these choices would greatly aid in comprehending the significance and implications of these methodological decisions.
#2.3 – Methods (Data analysis section): To enhance reproducibility, I recommend adding a table in the supplementary material section. This table should list the names of the R packages used, their respective versions, and references.

#3 Results: The results section provides a detailed and thorough presentation of the study's findings and analyses. The section appears to be well-organized and structured, with clear descriptions of the methods used and the outcomes observed. The use of tables, figures, and numerical values enhances the clarity of the presented results. However, there are a few areas where the clarity and objectivity could be further improved:
#3.1 Results (lines 350-368): Considering that age emerged as a significant variable in the mixed methods bivariate & multivariable analysis, and the low median age reported, it is essential to have information on how many HHCs were children or adolescents. Furthermore, this aspect of age, including the distinction between adult and non-adult contacts, should be highlighted in the discussion section.
#3.2 Results (lines 392-394): Please provide the sensitivity and specificity values of the ROC curve, with confident intervals.
#3.3. Results (lines 397-437): In lines 397-437, all paragraphs finish with a phrase with the same structure that starts with “Together, these data suggest”. I recommend rewriting some of them and evaluating if it is necessary. For example, I do not think that the phrase in lines 436-437 is fundamental for the text.
#3.4. Results (Figures): The figures in the manuscript, particularly Figures 1c and d, Figure 2, and Supplementary Figure 4, need to be improved for better comprehension. A key enhancement is to include the legend for colours and shapes directly on the figure itself, rather than solely in the written legend. The absence of this information on the figures themselves can hinder interpretation and understanding. Providing these details directly on the figures would facilitate a more immediate and clear comprehension of the data being presented, especially for readers who might be quickly scanning through the figure.

#4 Discussion: The discussion section demonstrates a good quality of scientific communication, with a thorough examination of the study's results and their implications. The section effectively contextualizes the findings within the existing literature and provides a balanced interpretation of the results. The limitations are acknowledged, and their potential impact on the study's conclusions is addressed. Minor comments:
#4.1. It would be beneficial to discuss in more detail the age of the Household Contacts (HHCs), as well as the rationale behind using sputum to assess biomarkers. These discussions could provide deeper insights into the study's findings and their implications.
#4.2 (lines 447-500): Breaking up the paragraphs could enhance the readability and clarity of

information, making it easier for readers to digest and understand the key points being conveyed.

Reviewer #3 (Remarks to the Author):

The study is very interesting and opens new ways for understanding TB pathogenesis and proposing clinical management. Based on the reported findings and prior studies, cough aerosol cultures are superior to sputum smear analysis in predicting Mtb transmission events and are likely the best estimators of TB infectiousness.
Few comments:
Figure 1A: is any possibility of evaluating the correlation of WHITE BLOOD CELLS components as Lymphocytes, monocytes, or neutrophils with CAC?. Indeed interesting correlations were found in the past using different approaches (La Manna, 2017, PMID: 28208160; Chedid, 2020, PMID: 32920230;). Moreover, it would be interesting to evaluate if there is an association between CAC score and platelets (La Manna, 2022 PMID: 35338775)

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This is a very interesting paper, looking to identify markers of infectiousness of individuals with TB. The study is centered around the cough aerosol detection mechanism and explores its validity and potential markers.

While this is interesting data, which should be certainly be published and feed into the wider discourse around Mtb transmission, there are a number of issues which make me unsure if the claims in the title, abstract and certain section headings are backed up sufficiently by the data. Some of these are clarifications and could be fine, others have to do with the study design and analytical approach, which may be harder to overcome.

Response: We thank the reviewer for these positive comments and attempt to address all concerns in the replies below.

****

A major issue is lack of a control group of HH contacts, which would help to understand how the 53% IGRA+ in the CAC- population should be interpreted. This means that the findings can reflect on markers for 'higher' infectiousness, as suggested by CASS positivity.

Response: We agree that providing a control group for the HH contacts is important to make population level claims regarding infectiousness. To address this concern, we include new data from two cohorts to estimate the frequency of QFT-positive tests among Nairobi residents without a known recent TB exposure (the "background" frequency of Mtb infection based on QFT result). For both cohorts, the rate of IGRA positivity was 51% and similar to rate of the household contacts of CASS-negative index cases (54%). All of these rates are lower than the household contacts of CASS-positive index cases (88%). These data suggest that CASS-positive pulmonary TB cases have a higher degree of infectiousness compared to those who are CASS-negative with evidence from multiple control groups.

The following language has been added to the Methods section:

"We estimated the frequency of positive interferon-gamma release assays in Nairobi residents who were not known to have active TB or recent close contact to a person with active pulmonary TB. Participants were enrolled through two studies. As part of a study of cough analysis for TB detection (38), we recruited adult outpatients (n=45) from the same TB and respiratory clinics as TBAIT participants with TB (described above) who were identified through passive case finding and presented with cough. All participants underwent sputum GeneXpert testing and chest X-rays to exclude pulmonary TB. Additionally, we report results from a study to evaluate treatment responses during isoniazid preventive therapy (39). From December 2019 to December 2020, we recruited adults with (n=121) and without (n=122) HIV from three HIV prevention and care centers in Nairobi. We assessed for TB using the WHO four-symptom screen, followed by sputum GeneXpert and chest X-rays among those with one or more symptoms. To calculate prevalence of QFT positive tests, we excluded indeterminate results."

We added the following to the Results section:

"To establish the background prevalence of QFT positive results in Nairobi residents, we evaluated QFT status in adult outpatients who presented for health care due to cough (n=45, median age 39 years (IQR, 34-45)) and to HIV prevention and care centers (n=121 PLWH; n=122 without HIV; overall median age 36 years (IQR, 28-44). After excluding indeterminate results (n=2 and n=34, respectively), we found that the prevalence of QFT-positive results was, 51% (n=22) and 51% (n=121), respectively. Among the 45 persons who presented to health care with a cough, QFT status did not differ by HIV status (PLWH = 17,

p-value 0.66). In participants recruited at HIV care centers, QFT results differed by HIV status: 43% of PLWH (48 of 111) and 60% of persons without HIV (59 of 98) (p-value 0.01) were QFT positive. The estimated Nairobi background prevalence of QFT test positivity did not differ from contacts of cough aerosol culture-negative participants (p-value 0.33) but was significantly lower than contacts of cough aerosol culture-positive participants (p-value <0.001)."

A key gap in the current methods is a lack of description of how the shift from 142 participants with TB to including 55 index cases in the HH contact study. It is mentioned in the results section, but is not described at all.

Response: Thank you for highlighting this concern. We added the following clarifying statements in the Results Section,

"We next evaluated IGRA results in household contacts to determine whether cough aerosol culture status is associated with infectiousness. Of the 142 participants with TB, 30 lived alone and were not eligible for the household contact study and 57 had household members who could not be contacted or the household members declined enrollment. From the remaining 55 index participants, we enrolled 143 household contacts and obtained QFT tests on 127 (16 declined phlebotomy or had unsuccessful phlebotomy) (Table 2)."

Similarly, n=58 index cases were included in the transcriptomics, but their selection is not described. That needs to be clarified (see more detailed comments below).

Response: To clarify the selection of individuals for transcriptomic studies, we added the following information to the Methods section:

"We selected 58 subjects for whole blood transcriptomic analysis which were a subgroup of the full cohort of 142. We initiated this experiment after 29 CAC+ and 29 CAC- subjects were available for analysis from the initial phase of enrollment (Supplementary Table 4)."

The analysis approach for the predictive models is a simple forward selection of variables. In the context of this well studied area, this seems too uninformed. The consequence is some strange findings which the authors do not pick up on. E.g. Why would MUAC be predictive but BMI would not?

Response: Thank you for this important question. While MUAC and BMI would seem to be two measurements of the same phenomenon (malnutrition) and therefore highly correlated, a body of findings from the literature demonstrates that there are differences between these biometrics. BMI measures the presence or loss of adipose tissue. MUAC assesses both adipose tissue and muscle mass. As such, MUAC reflects the presence of cachexia.

We added the following to the Discussion with new references:

"Our finding of a stronger association between cough aerosol status and MUAC than BMI may seem counterintuitive. However, MUAC and BMI measure related but different aspects of nutritional status: BMI reflects fat mass while MUAC measures fat and muscle mass. (51). Low MUAC was a better predictor than BMI of all-cause mortality in older adults (51-53), and in persons with TB. MUAC (unlike BMI) was independently associated with cavitary lung disease (54). Cachexia, severe weight loss that results from muscle atrophy and the loss of adipose tissue, may correlate with differential immune cell responses (55). For example, using parasitic infections in the murine model, mice deficient for CD4+ T cells experienced muscle and fat wasting as opposed to CD8+ T cell-deficient mice which experienced only fat wasting (56)."

How should we interpret CASS predictive model mean where a dichotomised age variable is given

equal importance as the sole bacteriological indicator (Xpert CT threshold)? While an uninformed approach to the cytokine/transcriptomics analysis seems OK given its novelty, this is not the case for associations for transmission or infectiousness, and this should be addressed, e.g. by discussing a conceptual framework, and exploring different predictive models with a-priori choices.

Response: Thank you for raising this important concern. We added a conceptual framework that is presented in Figure S1 as a Directed Acylic Graph (DAG) to evaluate pulmonary TB host characteristics (predictors) associated with cough aerosol culture positive status (outcome). We included predictors (listed in methods section below), potential confounders (bacillary burden, HIV, and sex), mediators (health care engagement) and effect modifiers (Mtb strain), strengthening the rationale for our models' variable selection. We provide further details in the Methods section and Figure legend below.

The following new language has been added to the following sections:

Methods:

"We assessed associations between host TB characteristics and cough aerosol culture-positive status using bivariate and multivariable logistic regression models in which we evaluated predictor variables based on our conceptual model presented as a directed acyclic graph (DAG, Supplementary Figure 2) (45). Based on published studies of TB transmission and cough aerosol status (18, 25, 29), we categorized as predictors characteristics that we hypothesized as determining the cough aerosol culture-positive phenotype including younger age, functional status (body mass index (BMI), mid-upper arm circumference (MUAC), TB symptoms, cough duration, Leicester Cough Questionnaire (LCQ) score, prior history of TB, and cough peak flow) (29) and systemic inflammation (CRP, hemoglobin, white blood cell count and differential, hemoglobin A1C percent, and sputum appearance). In the DAG we categorized HIV status, bacillary burden (GeneXpert semi-quantitative grade, GeneXpert Ct value, AFB-smear grade, time to detection (TTD) of Mtb growth in liquid culture, presence of chest X-ray cavitation, and number of radiographic quadrants with TB-related changes), and sex as potential confounders due to their shared associations with the exposure and outcome (46). Health care engagement is considered a mediator, a node that is on the causal pathway from exposure to outcome and should not be controlled in models. Pathogen-related factors, which are not included in this study, would be considered effect modifiers as they may contribute to the outcome and modify the effect of other causes of the outcome. GeneXpert Ct values were assigned using the smallest Ct value from any of the probes targeting the rpoB gene; participants with Xpert Ultra trace positive results (for whom rpoB probe Ct values were 0) were assigned a Ct value of 35, near the highest detectable Ct value for Xpert Ultra."

Results:
"We evaluated host characteristics associated with cough aerosol culture status based on a conceptual framework which included potential confounders (Table 1, Supplemental Figure 1)."

Figure Legends:
"**Supplementary Figure 1: Directed acyclic graph (DAG) depicting conceptual framework of factors that determine cough aerosol culture status in persons with pulmonary TB.** DAGs depict associations between an exposure (pulmonary tuberculosis), outcome (cough aerosols that are culture positive for M. tuberculosis), mediator (health care engagement), potential confounders (bacillary burden, HIV status, sex) and effect modifier (pathogen factors such as strain of M. tuberculosis),"

The more conceptual issue is what space of infectiousness this paper addresses. The title, abstract and body text shift between scope. For example, the title and abstract reflect on 'TB infectiousness', suggesting it is about individuals that are either infectious or not. As clarified later in the abstract, the authors actually focus on a differentiation between more or less infectious

individuals, based on CASS. They show that CASS+ individuals are likely more infectious with a nested household contact study (which is not a new finding), but there is no control group, so no statement about a binomial infectious yes/no status. For the transcriptomics they then look in a subgroup (which is not defined, see above), and find a number of associations with CASS+. I would agree that this provides data to 'suggest host inflammatory signatures are associated with Mtb aerosolization independent of bacillary load and cavitary lung disease.', as stated in the discussion. But what it does not show is 'Tuberculosis Infectiousness is Associated with Distinct Clinical and Inflammatory Profiles', which is the paper title, and as such much too broad. I think the authors should be more clear, and e.g. change the title, and any other statements regarding 'infectiousness' as a whole to reflect the more narrow scope.

Response: Thank you for these suggestions. We added two control groups to provide an estimated level of Mtb infection (based on QFT results) in adult residents of Nairobi who are without a known recent TB exposure (please see detailed response above). We changed the title to, "Mycobacterium tuberculosis Cough Aerosol Culture Status is Associated with Host Clinical Characteristics, Host Inflammatory Profiles, and Infection in Close Contacts". We have clarified the group of participants who underwent transcriptional profiling (please see explanation above).

Methods:

While I understand why ignored trace+/culture- (given the low chance of being CASS pos), it would be good to state this explicitly, and state how the authors interpreted this group. (TB or not TB)?

Response: We agree that more clarity is needed. The trace positive/culture negative individuals were identified through a prevalence survey where we used an active case finding method which is more likely to identify people with subclinical TB with minimal symptoms. We excluded individuals who were trace positive/culture negative as we do not consider them to have active TB. Participants in this category were identified only through active case-finding, were culture negative on four sputum samples that underwent MGIT culture and did not fit the eligibility of our CASS-study protocol. We added the following sentence to the Methods section,

"We did not include persons in this study identified through active case finding who were trace positive/culture-negative as we concluded that they did not have active pulmonary based on four negative sputum cultures."

There is some reporting of results for BMI vs MUAC in the methods, this should be moved to the results.

Response: We are unable to find this in the text.

L314-318 is results I think, should move there.

Response: We agree and moved the following from the Methods to Results:

"While a strong majority of the genes (78%) were best fit by models including cough aerosol culture status alone, 16% of genes were best fit with the inclusion of bacillary load with a smaller percentage best explained by more complex models with age and cavitary disease (Supplemental Figure 3A)."

It is not obvious from the methods that the HH contacts came from a subset (n=55, l352) of enrolled participants from the total population (n=142). I did not see it in l185-l189), nor is it mentioned in l142-l144. Firstly, this should be made more clear. Secondly the rationale and method for this selection should be described explicitly (e.g. funding restrictions and random draw?).

Response: Thank you. We added the following clarifying statements in the Results Section:
"We next evaluated IGRA results in household contacts to determine whether cough aerosol culture status is associated with infectiousness. Of the 142 participants with TB, 30 lived alone and were not eligible for the household contact study and 57 had household members who could not be contacted, or the household members declined enrollment. From the remaining 55 index participants, we enrolled 143 household contacts and obtained QFT tests on 127 (16 declined phlebotomy or had unsuccessful phlebotomy) (Table 2)."

There is a similar selection for the host transcriptomic work (going down to n=59 index patients), which needs to be similarly explained.

Response: Thank you. Please see detailed responses above.

RESULTS
335: It seems odd that all of the participants were cough positive, as a percentage came from community screening. Given the difference in population and expected disease severity between ACF and PCF populations, these two populations should be described more in the text and supp materials. If the 100% cough positive holds, it is something to discuss, as it goes against most modern TB prevalence surveys where about 50% were found to be cough negative.

Response: Thank you for this suggestion. We agree that more description of these two populations is important for clarity. All participants identified through active case-finding were CAC-negative. As persons who presented for health care contributed entirely to CAC-positive participants (thus, 100% with cough) and contributed the majority of CAC-negatives, we see the high frequency of cough. Of the 26 participants enrolled through active case finding (all of whom were CAC-negative), 5 denied cough, for a cough frequency of 81%. In the community-based prevalence survey, 42% of people with prevalent TB denied cough. We did not enroll all people with prevalent TB into our CASS study as some were GeneXpert negative and later diagnosed with TB based on culture results, declined enrollment or were not available for recruitment (e.g. were diagnosed on a Saturday). We are not suggesting that the ACF participants are representative of prevalent TB in the community and have added text for clarification. We added the following clarifying language to the Methods section:

"We did not include persons in this study identified through active case finding who were GeneXpert negative/culture-positive due to delays in their TB diagnosis, who were trace positive/culture-negative as we concluded that they did not have active pulmonary based on four negative sputum cultures, those who were diagnosed on weekends or those who declined enrollment."

We added the following to the Results:

"As compared to participants identified through passive case finding (n=116), those identified through active case finding (n=26) were older (mean 43 vs. 35 years, p-value 0.005), less likely to report a cough (81% vs. 98%, p-value <0.001), less likely to have cavitary disease on chest x-ray (38% vs 77%, p-value <0.001) and had higher mean GeneXpert Ct values (24.0 vs. 19.2, p-value <0.001); there was no difference in the frequency of women or PLWH."

Table 1: The MUAC result is odd (i.e. a higher MUAC means more CASS+). It seems counterintuitive, and the lack of an association with BMI makes me concerned the MUAC results are less reliable. The BMI/MUAC discrepancy should be discussed.

Response: Please see above response.

Table 2: I am not clear what the OR and p-values are meant to show here. For the index case

characteristics, those seem a near 1-2-1 repeat of table 1? I think those could be dropped, especially from the main body text. For e.g. the number of contacts per case – I think this is just a chi-square, but that is not the relevant finding. What we would like to know is whether the distribution of HH contacts is different by CAC+ and CAC-, which is currently unclear.

Response: Table 2 contains a sub-set (n= 55) of the index participants in Table 1 (n=142). Since some of the index participants lived alone or had household contacts that were not enrolled (could not be reached, declined participation), there are differences in the included index participants. The purpose of Table 2 is to show variables of interest that may associate with differences in transmission risk and QFT result in order to assess for contributors to the differences in household contact QFT status by cough aerosol status of the index participant. Ultimately, we are offering data to support our hypothesis that cough aerosol status is the largest driver (risk factor) for TB transmission to close contacts.

We have removed the odds ratio column from Table 2, clarified in Table 2 the number of index participants and their relation to household contacts, and removed the requested language from the main text. The p-value compares the size of households enrolled and shows a statistically significant difference. The number of contacts per household is larger for contacts of CAC-negative index participants: 36% of HHCs of CAC+ index participants were from households that enrolled 3 or more contacts compared to 48% of HHCs of CAC- index participants. We have clarified this by adding percents to the Table 2 cells.

What were reasons for missing QFT (N goes from 143 to 127)? These should be reported.

Response: The reasons for missing QFTs among enrolled HHCs was due to refusing phlebotomy or unsuccessful phlebotomy. We added the following text:

"From the remaining 55 index participants, we enrolled 143 household contacts and obtained QFT tests on 127 (16 declined or had unsuccessful phlebotomy) (Table 2)."

Table 3: The results are quite stark, but not sure how superior the final model is. It makes sense that after one good measure of bacteriological burden is included in the model, the others are no longer adding enough value to be included. E.g. smear and cavitation seem to be quite useful explanatory/predictive variables by themselves. A simple forward stepwise modelling approach will not highlight those limitations. Given what we already know about infectiousness, this type of analysis would be better served with a formal DAG or at least conceptual framework, and then working through the various options or models. I suggest the authors explore this wider, and thereby enrich the insights that can be gained from the data.

Response: Thank you for this suggestion. We added a DAG as Figure S1 which includes the exposure, outcome, predictors, potential confounders and effect modifiers. The DAG includes variables that are known to have impact on infectiousness and were used in our model development when data was available. We added new text to the Methods, Results, and Figure legend as summarized above.

While the AUC for the scoring system for predicting CAC+/- is high, I don't quite understand the results from Sup table 1. It seems to suggest that age and Xpert CT are equally valuable (both 7 'points')? But what does the model/score mean if the sole bacteriological indicator is removed? Are the authors suggesting that we could decide on CAC+/- by age, CRP and MUAC alone? Could they provide an indication of how much value each individual characteristic adds, and how the model works if Xpert CT is not available. Could any other bacteriological indicator be switched in? I'd be surprised if the model functioned without a bacteriological indicator, and also not sure about the biological logic that would remain. It is also not clear if this predictive model is solely applicable for people who have bacteriologically positive TB? If so, that should be made clear.

Response: Thank you for these comments. We adjusted the model to have 3 categories for GeneXpert

result based on semi-quantitative grade. The performance of this clinical prediction model is predicated on assessment of all the variables included in the tool. Thus, this tool should only be used for people who have bacteriologically positive TB. Supplementary Table 3 displays the changes in model sensitivity and specificity based on different cut-offs, which addresses aspects of the reviewer's questions regarding lack of access to data on model factors. Performance in other settings will need validation.

We added the following to the Results:
"We developed and evaluated the performance of a clinical prediction rule to predict cough aerosol culture-positive persons as a means to identify those who are likely to be highly infectiousness. We included predictors from our multivariable logistic regression model that are readily available to a clinician during the initial visit. For this reason, we removed time-to-detection of culture growth but kept CRP level as there are point-of-care versions of this test.(52) We also used Xpert Ultra semi-quantitative grade which is the result generally reported to clinicians and is derived from the Ct value. Using the coefficients obtained in the multivariable logistic regression analysis we derived the following prediction equation for cough aerosol culture-positive status where outcomes were coded according to weighted scores based on the beta coefficient: CRP>42.9 + Age < 43.8 years + MUAC > 22.8 cm + Xpert semi-quantitative grade (high, medium, or low/very low/trace). A total risk score is calculated by adding the risk points and has an optimum cut point of 15 points for predicting cough aerosol culture-positive status with estimated sensitivity of 0.86 (95% CI, 0.74 - 0.95) and specificity 0.65 (95% CI, 0.56 - 0.74). (Supplementary Tables 2 and 3) The prediction rule based on the calculated risk score had an area under the receiver operating curve (AUROC) of 0.84 (95% CI: 0.77-0.91) (Supplementary Figure 3). The model's Somers' D_{xy} index was 0.71; the equivalent in bootstrap validation was 0.66, with an optimism estimate of 0.051, indicating good stability of the model in internal validation."
."

We added the following to the Discussion:

"Like other clinical prediction rules, the model performance is predicated on having data for all of the characteristics that contribute to the risk score. While promising, the risk score that we present requires external validation and should not be used at this time for clinical (non-research) purposes."

The variation in association pattern of the different cytokines is a bit strange. Would one not expect them to be elevated across the various indicators, also given that e.g. CT threshold and cavitation were themselves associated with CAC? The lack of stability makes me a bit concerned. The results are what they are, but potential explanations (including random noise across many comparisons that could affect the threshold for interpretation of p-values).

Response: We agree that the sputum cytokine results contain variation in patterns of associations with different outcomes. Although these results are not consistent with a well-defined model of TB pathogenesis, each cytokine has distinct activation pathways and cellular sources that likely to contribute to the variation of results. For example, IL-1B secretion is inflammasome-dependent while CXCL8 and IL6 secretion are primarily regulated by Toll-like Receptor and NOD-like receptor signaling pathways. All three cytokines are produced by several myeloid cells, but to different degrees. Finally, each has different autocrine and paracrine effects that likely get amplified at different stages of disease. For example, CXCL8 is a chemokine for neutrophil recruitment with potential for amplification of its levels during disease stages which are neutrophil predominant. At the suggestion of Reviewer #3 below, we examined leukocyte subsets and found that peripheral blood neutrophil counts are associated with CAC status, an association which is consistent with the higher levels of CXCL8 in CAC+ versus CAC- sputum cytokine analyses.

To further address these cytokine observations, we collected more data and performed more analyses. First, we measured the same 4 cytokines in plasma to provide a parallel inflammatory profile in a different

compartment. Second, we used a multivariate linear regression model to examine whether any of the associations were independently associated with CAC status.

For the plasma cytokines, we were able to detect TNF and IL6, but not IL1B or CXCL8. In contrast, for our previous sputum cytokine results, we detected IL6, IL1B, and CXCL8 but not TNF. These results are consistent with compartmentalized immune responses and highlight the importance of analyzing inflammatory profiles in multiple sites. We then performed a multivariate logistic regression analysis to assess whether any of the sputum or plasma cytokines were associated with CAC independently of bacillary load. Interestingly, plasma TNF was the only cytokine associated with CAC in the adjusted analysis. These results provide additional evidence that inflammatory profiles are associated with CAC status and also highlight the importance of examining multiple sites of inflammation. However, we agree these cytokine observations are incomplete and not consistent with a fully developed or coherent model. Our RNASeq experiments were undertaken to attempt a more in-depth analysis of inflammatory pathways that are associated with aerosolization.

We revised the Methods section and Figure 1 to include the plasma cytokines, added Supplementary Table 4 for the multivariate analysis, and added the following text to the Methods and Results sections:

Methods:

“Four cytokines were examined in sputum and blood. TNF, IL6, IL1B, and CXCL8 were chosen based on their central roles in regulating inflammatory pathways and TB pathogenesis. (42) … Plasma was tested directly. Sputum supernatants and plasma were tested for CXCL8 (IL-8), Interleukin 1 beta (IL-1β), Interleukin 6 (1L-6), and TNF using sandwich ELISA according to the manufacturer’s instructions. (R&D Systems Inc. Minneapolis, USA.). In pilot sputum cytokine studies, TNF was not detectable at high levels and was not examined further. In pilot plasma cytokine studies, CXCL8 and IL1B were not detectable at high levels and were not examined further.”

Results:
“To further examine the association of inflammatory markers with cough aerosol culture-positive status, we analyzed sputum and plasma levels of CXCL8 (IL-8), IL-1β, TNF, and IL-6 (Figure 1B, Supplementary Table 5). TNF in sputum and IL-1β and CXCL8 in plasma were nearly undetectable in preliminary subgroup testing and were not examined further. The sputum concentration of CXCL8 was significantly higher (p-value 0.03) among the cough aerosol culture-positive compared to cough aerosol culture-negative participants. Sputum IL-1β (p-value 0.052) and IL-6 were not different (p-value 0.36) (Figure 1B). In addition, higher IL-6, CXCL8, and IL-1β levels were associated with greater bacillary burden (measured by GeneXpert cycle threshold, p-value 0.0009, 0.007 and 0.00007, respectively, Figure 1C). IL-1β, but not IL-6 or CXCL8, was positively associated with cavitary lung disease (p-value 0.0007) across cough aerosol culture status.

Plasma IL-6 was associated with bacillary burden (p=1.69 x 10-6) and cavitary disease (p=0.005) and was higher in CAC-positive compared CAC-negative participants (p=0.03). Plasma TNF was associated with cavitary disease (p=0.013), but not bacillary burden, and was lower in CAC-positive compared to CAC-negative participants (p-value 0.0007). For cytokines that had a significant association with cough aerosol status in bivariate models (sputum CXCL8, plasma TNF, plasma IL-6), we evaluated associations in multivariate logistic regression models in which we adjusted for age, sex, cavitary disease on chest x-ray and GeneXpert Ct value. We found that plasma TNF was independently associated with CAC status (adjusted p-value 0.007, Supplementary Table 4).”

Discussion (in the Limitations paragraph):

“Fourth, sputum collection methods are not standardized and have more technical heterogeneity in

comparison to other biologic samples. Despite this challenge, sputum provides a direct assessment of the primary site of TB disease where pulmonary immune responses are compartmentalized and differ from blood. In addition, sputum cytokines have been evaluated as biomarkers in the diagnosis of TB and treatment monitoring.(65-67) However, to our knowledge no prior studies have evaluated their association with infectiousness measured by culturable aerosols."

L455-457: these are results and should be moved there. On a broader note, it would be good to place those observations in the context of the ongoing discussions around subclinical TB (see previous comment re cough positivity).

Response: We agree that this sentence summarizes results, but the referenced statement ("In this study, a lower TB symptom score indicated a lower burden of findings attributable to TB disease as it is a summation of points for TB-related symptoms (cough, hemoptysis, dyspnea, chest pain, fever, night sweats), TB-related signs (anemia, tachycardia), lung auscultation findings, and malnutrition (low BMI, low MUAC).") refers to the study by Theron et al, not our findings. We have made it clear in the text that we are referring to findings from Theron's study and that this is in the context of our Discussion of the literature and findings suggestive of a CAC-positive phenotype.

L462: the statement about 'body mass' is strange, as BMI was not associated, but MUAC was. As mentioned before, this discrepancy needs discussion, and the statement should be adjusted accordingly.

Response: Thank you. We address this comment above. In addition, for line 462, we changed "body mass" to "muscle mass".

L473: this only interprets one of the cytokine findings, but ignored the conflicting results in the actual section (see above), where some cytokines were associated for one indicator, but not the other. Do the authors assume only 1 indicator is valid? If so, that is a stretch. If not, the discrepancies need to be addressed.

Response: Thank you for highlighting this issue which we partially address above including revision of the Results paragraph and addition of new data. Although the cytokine data has some level of association with CAC status, we do not claim to have a coherent model based on the limited scope of the cytokine analysis. We used the cytokine observations as a rationale for pursuing a more comprehensive analysis with whole blood RNAseq profiling. We considered how to revise Discussion section which includes line 473. Given the limited scope of the cytokine work, we would prefer to remove this sentence and focus the paragraph on discussion of the transcriptomic findings. We believe this modification adds more focus to the paragraph on the most important findings.

In summary, we removed the following sentence from the Discussion paragraph:

"Our results of significantly higher concentrations of sputum CXCL8 and 1L-1β, as well as higher serum CRP, in participants with culturable aerosols support the role of systemic inflammation in TB infectiousness."

L535: the suggestion that interventions (i.e. treatment) should target the most infectious is a bit strange. This approach was tried with the development of DOTS policies, where symptomatic smear positive individuals were targeted through PCF and 6m therapy. While saving lives, it has not interrupted transmission as expected, and it is not clear to me why this would work now? Shouldn't Identifying and treating all infectious individuals be the minimum target for all high burden countries, regardless of how infectious they are at the time of diagnosis?

Furthermore, infectiousness is not static, and will likely vary over time. Finally, given the high

Number Needed to Test to identify any person with low or high infectious disease, the cost of treatment is really not the limiting factor. It would be helpful if the authors reflected a bit more on this (also in the introduction).

Response: Thank you for these suggestions and perspective for which we fully concur. Smear-based diagnosis is a crude and insensitive method for diagnosis and our statements are not intended to evoke that approach. "Targeted treatment interventions" is not meant to imply that treatment should only be provided to those who are most infectious. Rather, in resource constrained settings where enhanced treatment support such as "true" DOT is not commonly practiced, the most infectious persons could also be supported with true DOT, early drug susceptibility testing to ensure that treatment is appropriate, or other treatment support interventions. Also, as mentioned, more extensive contact investigations could be pursued to ensure evaluations for active TB and TB preventive therapy in contacts who are likely to have been recently infected.

We edited the final Discussion paragraph to clarify our intended message,

"While persons with pulmonary TB who are cough aerosol culture-negative patients may transmit Mtb, identifying the most infectious persons would allow targeted interventions to support TB control efforts such as isolation, true direct observation of treatment, and drug-susceptibility testing to confirm that treatment is effective. TB control could also be supported through enhanced investigations in contacts of the most infectious persons to evaluate for active TB and provide TB preventive therapy given higher likelihood of recent transmission and progression to disease."

Minor points
Ref 33, 34 and 58 are the same I believe. Would be good to recheck references and make sure in order.

Response: Thank you. This has been corrected. All references were reviewed for accuracy and order.

Table 1: category 'Enrolled thru prevalence survey' – not appropriate language for scientific paper and not reflected in main text. Suggest something along lines of 'ACF-enrolled', as this is how it is described in methods.

Response: Thank you. We changed this category to "Enrolled through active case finding".

Table 3: the unit for Age is not given. From the OR I suspect it is year of age, but this should be made clear. Also note this assumes a linear association across the entire age range, which is unlikely to be true given what we know about age-dependent population ARI. This should be discussed.

Response: The units (years) has been added to the tables. is has been corrected.

We added the following to the Discussion:

"The annual risk of TB infection differs by age groups in high burden settings and notably, there was no difference in the ages of HHCs by CAC status. Among household contacts less than 10 years of age, who have a lower annual risk of infection compared to adolescents and adults, QFT positive results were also more common (p-value 0.01) among contacts of CAC-positive persons (9 of 10 participants, 90%) than contacts of CAC-negative persons (14 of 33 participants, 42%)."

The section titles on l397 and l409 are phrased as interpretation rather than reporting of results. These should be adjusted. In addition, the last sentences in each of these sections is interpretation rather than result and should be removed or put in discussion.

Response: Thank you for this suggestion. We edited the phrases and sentences to remove interpretation language from the Results section.

The new section titles are:
"Sputum cytokines and analysis of associations with cough aerosol culture positivity, bacterial load, cavitary lung disease, and extent of lung involvement."

And

"Whole blood transcriptomic signatures and analysis of *associations with cough aerosol culture positivity and sputum bacterial load."*

In addition, we removed the summary sentence of these sections to avoid over-interpreting in the Results section.

Reviewer #2 (Remarks to the Author):

The study "Tuberculosis Infectiousness is Associated with Distinct Clinical and Inflammatory Profiles" stands out for addressing a significant gap in tuberculosis research, focusing on individual infectiousness and biomarkers. It employs a prospective longitudinal study design, examining treatment-naïve pulmonary TB patients in Nairobi, thereby providing valuable insights in a high TB burden setting. Methodological innovations, such as using the Cough Aerosol Sampling System (CASS) for measuring Mycobacterium tuberculosis aerosolization, the evaluation of sputum cytokine, and assessing household contacts through IGRA testing, add important epidemiological and molecular dimensions and adds novel knowledge and tools to predict Mtb transmission. Overall, the manuscript was well written. The incorporation of suggested improvements would enhance the paper's quality and ensure that concerns have been adequately addressed.

#1 Introduction: The introduction section is well-written and effectively frames the context of TB infectiousness to its clinical and immunological aspects. It successfully outlines the importance of understanding individual variations in TB transmission and reviews relevant literature, thus situating the study within the broader field of TB research. I have no suggestions for improvement in this section.

Response: We appreciate your kind words.

#**2 Methods: The methods section is well-structured and thorough. However, consider breaking down the dense text into smaller paragraphs to improve readability (i.e. break the paragraph in line 262.)**

Response: Thank you, we made the suggested edits.

#2.1 – Methods (Line 203-204): The study does not address the issue of drug-resistant TB. On lines 203-204, the study mentions classifying rifampicin resistance as "positive", "negative", or "indeterminate", but this is not further elaborated in the text. I recommend including these data in the results and tables, even if 100% were susceptible, to provide a complete picture of the resistance profile among the study participants.

Response: Thank you for pointing this out; it was an oversight on our part. One subject had rifampin resistance. We added information on detection of rifampin resistance to Table 1.

#–.2 - Results (lines 222-228): It would be beneficial to provide a rationale for the decision to

assess cytokines in sputum instead of blood; as well as the selection of the specific three cytokines studied. This is particularly important for readers who may have less experience in the field. Including one or two lines to explain these choices would greatly aid in comprehending the significance and implications of these methodological decisions.

Response: We agree that adding more context around these issues will provide more clarity. We chose to analyze sputum cytokines due to our hypothesis that the inflammatory mediators and/or biomarkers of aerosolization are present in the lung and more likely to have a direct impact compared to the peripheral blood. In addition, the immune system is compartmentalized, and lung responses are often different than peripheral blood. However, there are technical challenges of collecting sputum in a standardized manner. In contrast, whole blood and plasma measurements are more standardized. To address this directly, we have now added cytokine measurements in plasma for comparison to sputum. We now have data for 4 cytokines in both sputum and blood. TNF was detectable at high levels in sputum and CXCL8 and IL1B were not present in plasma at high levels. These data support the compartmentalization of immune responses and the value of directly studying pulmonary immune responses when feasible.
The rationale for studying IL6, TNF, and IL1B is based on their known role in TB pathogenesis. We chose to study CXCL8, due to its role as a chemokine involved in neutrophil recruitment, an important cell in TB pathogenesis, and previous studies documenting its presence in sputum.

We added the following text with new references:

Methods:
"Four cytokines were examined in sputum and blood. TNF, IL6, IL1B, and CXCL8 were chosen based on their central roles in regulating inflammatory pathways and TB pathogenesis. (42)" … Plasma was tested directly. Sputum supernatants and plasma were tested for CXCL8 (IL-8), Interleukin 1 beta (IL-1β), Interleukin 6 (1L-6), and TNF using sandwich ELISA according to the manufacturer's instructions. (R&D Systems Inc. Minneapolis, USA.). In pilot sputum cytokine studies, TNF was not detectable at high levels and was not examined further. In pilot plasma cytokine studies, CXCL8 and IL1B were not detectable at high levels and were not examined further."

Results:
"To further examine the association of inflammatory markers with cough aerosol culture-positive status, we analyzed sputum and plasma levels of CXCL8 (IL-8), IL-1β, TNF, and IL-6 (Figure 1B, Supplementary Table 5). TNF in sputum and IL-1β and CXCL8 in plasma were nearly undetectable in preliminary subgroup testing and were not examined further. The sputum concentration of CXCL8 was significantly higher (p-value 0.03) among the cough aerosol culture-positive compared to cough aerosol culture-negative participants. Sputum IL-1β (p-value 0.052) and IL-6 were not different (p-value 0.36) (Figure 1B). In addition, higher IL-6, CXCL8, and IL-1β levels were associated with greater bacillary burden (measured by GeneXpert cycle threshold, p-value 0.0009, 0.007 and 0.00007, respectively, Figure 1C). IL-1β, but not IL-6 or CXCL8, was positively associated with cavitary lung disease (p-value 0.0007) across cough aerosol culture status.

Plasma IL-6 was associated with bacillary burden (p=1.69 x 10-6) and cavitary disease (p=0.005) and was higher in CAC-positive compared CAC-negative participants (p=0.03). Plasma TNF was associated with cavitary disease (p=0.013), but not bacillary burden, and was lower in CAC-positive compared to CAC-negative participants (p-value 0.0007). For cytokines that had a significant association with cough aerosol status in bivariate models (sputum CXCL8, plasma TNF, plasma IL-6), we evaluated associations in multivariate logistic regression models in which we adjusted for age, sex, cavitary disease on chest x-ray and GeneXpert Ct value. We found that plasma TNF was independently associated with CAC status (adjusted p-value 0.007, Supplementary Table 4)."

Discussion (in the Limitations paragraph):
"Fourth, sputum collection methods are not standardized and have more technical heterogeneity in

comparison to other biologic samples. Despite this challenge, sputum provides a direct assessment of the primary site of TB disease where pulmonary immune responses are compartmentalized and differ from blood. In addition, sputum cytokines have been evaluated as biomarkers in the diagnosis of TB and treatment monitoring.(65-67) However, to our knowledge no prior studies have evaluated their association with infectiousness measured by culturable aerosols."

#2.3 – Methods (Data analysis section): To enhance reproducibility, I recommend adding a table in the supplementary material section. This table should list the names of the R packages used, their respective versions, and references.
Response: Thank you for this suggestion. We added Supplementary Table 1 with the suggested information.

#3 Results: The results section provides a detailed and thorough presentation of the study's findings and analyses. The section appears to be well-organized and structured, with clear descriptions of the methods used and the outcomes observed. The use of tables, figures, and numerical values enhances the clarity of the presented results. However, there are a few areas where the clarity and objectivity could be further improved:

#3.1 Results (lines 350-368): Considering that age emerged as a significant variable in the mixed methods bivariate & multivariable analysis, and the low median age reported, it is essential to have information on how many HHCs were children or adolescents. Furthermore, this aspect of age, including the distinction between adult and non-adult contacts, should be highlighted in the discussion section.

Responses: Thank you. We have added age groups to Table 2 for young children (<5 years of age), children (5-10 years of age), and adolescents (10-15 years of age).

We added the following to the Results section:

"Among household contacts less than 10 years of age, QFT positive results were also more common (p-value 0.01) among contacts of CAC-positive persons (9 of 10 participants, 90%) than contacts of CAC-negative persons (14 of 33 participants, 42%)."

We added the following to the Discussion:

"Third, we demonstrated that cough aerosol culture-positive status was strongly associated with evidence of TB transmission in household contacts based on QFT results, including child HHCs for whom a positive QFT result is more likely to represent recent infection."

#3.2 Results (lines 392-394): Please provide the sensitivity and specificity values of the ROC curve, with confident intervals.

Responses: Thank you. We have added estimated sensitivities and specificities at different cut points in Supplementary Table 3A.

#3.3. Results (lines 397-437): In lines 397-437, all paragraphs finish with a phrase with the same structure that starts with "Together, these data suggest". I recommend rewriting some of them and evaluating if it is necessary. For example, I do not think that the phrase in lines 436-437 is fundamental for the text.

Response: Thank you. We have made the suggested edits and removed these summary sentences.

#3.4. Results (Figures): The figures in the manuscript, particularly Figures 1c and d, Figure 2, and

Supplementary Figure 4, need to be improved for better comprehension. A key enhancement is to include the legend for colours and shapes directly on the figure itself, rather than solely in the written legend. The absence of this information on the figures themselves can hinder interpretation and understanding. Providing these details directly on the figures would facilitate a more immediate and clear comprehension of the data being presented, especially for readers who might be quickly scanning through the figure

Response: Thank you for this suggestion. We made the suggested changes to Figures 1 and 2 and supplementary figure 4 (which is now supplementary figure 5.

#4 Discussion: The discussion section demonstrates a good quality of scientific communication, with a thorough examination of the study's results and their implications. The section effectively contextualizes the findings within the existing literature and provides a balanced interpretation of the results. The limitations are acknowledged, and their potential impact on the study's conclusions is addressed. Minor comments:

Response: Thank you, we appreciate the kind words.

#4.1. It would be beneficial to discuss in more detail the age of the Household Contacts (HHCs)

Response: Please see above response.

as well as the rationale behind using sputum to assess biomarkers. These discussions could provide deeper insights into the study's findings and their implications

Response: Thank you for this suggestion. As discussed above, our interest in studying sputum is to assess whether the compartmentalized immune response of the lung will generate novel insights compared to blood, especially with regards to the question of Mtb aerosolization. However, we fully acknowledge the methodologic challenges of working with sputum and the lack of standardization. We added more discussion of this in the limitations paragraph since we believe it is both a strength and a weakness due to these methodologic challenges.

We added the following text to the Discussion:

“Fourth, sputum collection methods are not standardized and have more technical heterogeneity in comparison to other biologic samples. Despite this challenge, sputum provides a direct assessment of the primary site of TB disease where pulmonary immune responses are compartmentalized and differ from blood. In addition, sputum cytokines have been evaluated as biomarkers in the diagnosis of TB and treatment monitoring.(57-59) However, to our knowledge no prior studies have evaluated their association with infectiousness measured by culturable aerosols.”

#4.2 (lines 447-500): Breaking up the paragraphs could enhance the readability and clarity of information, making it easier for readers to digest and understand the key points being conveyed.

Response: Thank you. We made the suggested edits.

Reviewer #3 (Remarks to the Author):

The study is very interesting and opens new ways for understanding TB pathogenesis and proposing clinical management. Based on the reported findings and prior studies, cough aerosol cultures are superior to sputum smear analysis in predicting Mtb transmission events and are likely the best estimators of TB infectiousness.

Figure 1A: is any possibility of evaluating the correlation of WHITE BLOOD CELLS components as Lymphocytes, monocytes, or neutrophils with CAC?. Indeed interesting correlations were found in the past using different approaches (La Manna, 2017, PMID: 28208160; Chedid, 2020, PMID: 32920230;). Moreover, it would be interesting to evaluate if there is an association between CAC score and platelets (La Manna, 2022 PMID: 35338775)

Response: Thank you for this suggestion. We evaluated associations between granulocyte, lymphocyte and monocyte counts and CAC status. While there was no association between the latter two white blood cell types and CAC status, the granulocyte count was significantly associated with CAC status (Table 1). We do not have platelet counts on study participants. The granulocyte odds ratio is very similar to the WBC count odds ratio; this is not unexpected as granulocytes are the largest number of WBCs and the two values (granulocyte count and WBC count) are highly correlated.

We edited the following sentence in Results to incorporate this data:

“Higher WBC counts (total and granulocyte count) and CRP levels were also associated with cough aerosol culture-positive status (Figure 1A).”

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I thank the authors for their responses and the changes to the paper, which have addressed most of my major concerns.

I remain sceptical of the sharp contrast between BMI and MUAC, and the suggested difference. We see the effects of nutrition on TB v clearly in BMI, as the recent RATIONS papers (Bhargava Lancet 2023, Bhargava Lancet GH 2023) and RePORT (Sinha preprint) papers have shown in terms of TB mortality and progression. For BMI to suddenly not matter in aerosol status seems more than a little odd. Also, MUAC is not solely driven by muscle mass, depending on the BMI, there is fat tissue as well. I suggest the text should be more open to the BMI vs MUAC findings reflecting artefact rather than fact and/or that the explanation is an hypothesis, rather than fact.

The other point that needs more clarification is how these findings in almost solely cough reporting individuals contrast with the findings of other aerosol studies, such as the face masks and RASC groups. Simply mentioning them in l558-570 does not address this. It seems strange to suggest that only cough reporting individuals are infectious, where e.g. face mask positivity was not correlated with cough (Williams Lancet ID 2020) but was correlated with HH infections (Williams CID 2023). It's possible that both are true, but then restricting to cough-positive individuals in this study poses a limitation.

I have some minor issues still, which the authors should also resolve in a final revision.

- The path from 142 to 55 participants is useful. It would be good if the paper reflected whether this could affect the findings. In particular whether HHs that refused to participate are different in a way that is associated with QFT positivity
- The explanation for how the 29 CAC+ and 29 CAC- is not sufficient. Is it a convenience sample (e.g. the first ones through the door). Again, anything that isn't random should be described and considered for analysis.
- It should be made clear what age range was studied, and the age range the authors are comfortable for their prediction model to be valid for.
- L298: With the 3 Xpert categories, variables are no longer all dichotomised, suggest to revise.

Reviewer #2 (Remarks to the Author):

The authors addressed all my questions and also the questions of the other reviewers. I am satisfied with the review and recommend the publication.

Reviewer #3 (Remarks to the Author):

the manuscript is interesting and great value scientifically and clinically. the authors answered the questions raised.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I thank the authors for their responses and the changes to the paper, which have addressed most of my major concerns.

I remain skeptical of the sharp contrast between BMI and MUAC, and the suggested difference. We see the effects of nutrition on TB v clearly in BMI, as the recent RATIONS papers (Bhargava Lancet 2023, Bhargava Lancet GH 2023) and RePORT (Sinha preprint) papers have shown in terms of TB mortality and progression. For BMI to suddenly not matter in aerosol status seems more than a little odd. Also, MUAC is not solely driven by muscle mass, depending on the BMI, there is fat tissue as well. I suggest the text should be more open to the BMI vs MUAC findings reflecting artefact rather than fact and/or that the explanation is an hypothesis, rather than fact.

Response: Thank you for highlighting this concern and the citations. We agree that the association of MUAC, but not BMI, with aerosolization may be a statistical artefact and that our discussion of the finding is speculative. We also agree that MUAC represents both fat and muscle.

We added more text to the Discussion to clarify that the BMI/MUAC differences may be a statistical artefact and that MUAC measures both fat and muscle mass:

“Our finding of an association between cough aerosol status and MUAC, but not BMI, may seem counterintuitive and may represent a statistical artefact. Alternatively, we hypothesize that MUAC and BMI measure related but different aspects of nutritional status: BMI reflects fat mass while MUAC measures fat and muscle mass.”

The other point that needs more clarification is how these findings in almost solely cough reporting individuals contrast with the findings of other aerosol studies, such as the face masks and RASC groups. Simply mentioning them in l558-570 does not address this. It seems strange to suggest that only cough reporting individuals are infectious, where e.g. face mask positivity was not correlated with cough (Williams Lancet ID 2020) but was correlated with HH infections (Williams CID 2023). It’s possible that both are true, but then restricting to cough-positive individuals in this study poses a limitation.

Response: Thank you for highlighting this concern. We agree that there is exciting new information that is potentially paradigm shifting to the field of TB. We also believe that the findings from the face mask and RASC groups requires further study to understand the relationship to infectiousness.

We added the following to limitations in the Discussion:

“Second, 95% of participants reported cough and our findings may not apply to persons with TB without cough. Recent studies that have detected TB in exhaled breath call into question whether cough is essential and/or a primary driver of TB transmission. (33, 34, 67, 68).”

33. Williams CM, Abdulwhhab M, Birring SS, De Kock E, Garton NJ, Townsend E, et al. Exhaled Mycobacterium tuberculosis output and detection of subclinical disease by face-mask sampling: prospective observational studies. Lancet Infect Dis. 2020;20(5):607-17.
34. Dinkele R, Gessner S, McKerry A, Leonard B, Leukes J, Seldon R, et al. Aerosolization of Mycobacterium tuberculosis by Tidal Breathing. Am J Respir Crit Care Med. 2022;206(2):206-16.
67. Williams CM, Muhammad AK, Sambou B, Bojang A, Jobe A, Daffeh GK, et al. Exhaled Mycobacterium tuberculosis Predicts Incident Infection in Household Contacts. Clin Infect Dis. 2023;76(3):e957-e64.
68. Patterson B, Dinkele R, Gessner S, Koch A, Hoosen Z, January V, et al. Aerosolization of viable Mycobacterium tuberculosis bacilli by tuberculosis clinic attendees independent of sputum-Xpert Ultra status. Proc Natl Acad Sci U S A. 2024;121(12):e2314813121.

I have some minor issues still, which the authors should also resolve in a final revision.

- The path from 142 to 55 participants is useful. It would be good if the paper reflected whether this could affect the findings. In particular whether HHs that refused to participate are different in a way that is associated with QFT positivity.

Response: We agree that differences between those index participants and contacts who did or did not participate in the household contact analysis needs to be carefully evaluated to determine whether they reflect the characteristics of the entire cohort and are generalizable. To assess the characteristics of our subgroup analysis, we compared index participants with (N=48) and without (N=94) household members. These comparisons are presented in (new) **Supplementary Table 5**. We found 3 variables that differed between index participants with and without enrolled household contacts (sex, Xpert Ct, and cavitary CXR). Each of these variables was tested in the multivariate model. These results suggest that our subgroup had some differences from the full cohort (possibly related to female index participants being less likely to live alone than male index participants) and could lead to selection bias and/or decrease generalizability. Concerns around selection bias were addressed through use of random effects models clustered on index participant. (Bell A. Fixed and random effects models: making an informed choice". Qual Quant 53, 2019). We have added these concerns to the limitations section.

We added the following to the Results:

"Index participants (N = 48) with enrolled household members compared to index participants (N = 94) without enrolled household members did not differ by age (p-value 0.12), HIV status (p-value 0.74), MUAC (p-value 0.60), or CASS status (0.17). (Supplementary Table 5) However, there were differences in the proportion of women (42% vs. 20%, p-value 0.007), Xpert Ct values (20.5 vs. 18.1, p-value 0.01), and the frequency of cavitations on chest X-ray (58% vs. 76%, p-value 0.04) between those with and without enrolled household members, respectively. ."

We added the following to the limitations paragraph in the Discussion:

"In addition, the index participants who contributed to our evaluations of risk factors for HHC IGRA results differed from index participants without enrolled HHCs which introduce selection bias and limit generalizability. Although our random effects models should address concerns around selection bias, generalizability of the HHC IGRA status remains a limitation."

In the original submission, we included household contacts of both pulmonary TB cases and those investigated for TB who were determined to not have TB. In the revision, we removed the contacts (n=14) of subjects who were determined to not have pulmonary TB after investigation. The number of total pulmonary TB index participants (n=142) is unchanged. The number of HHCs decreased from 143 to 129. Our study findings are the same whether these household contacts are included or not. The revised numbers are in updated versions of Tables 2 and 3.

- The explanation for how the 29 CAC+ and 29 CAC- is not sufficient. Is it a convenience sample (e.g. the first ones through the door). Again, anything that isn't random should be described and considered for analysis.

Response: Thank you for asking for further clarification. The 29 CAC+ and 29 CAC- were a convenience sample due to being the first available participants for transcriptomic analysis. To assess whether these 58 individuals were representative of the full cohort, we compared the variables of CAC+ and CAC- groups in the transcriptomic subgroup analysis (new Supplementary Table 6, N=29 CAC+ and 29 CAC-) with the full cohort (Table 1, N=43 CAC+ vs 99 CAC-). We found no differences in the participant characteristics of

the CAC+ and CAC- groups when comparing age, sex, HIV status, Xpert Ct, Cavitary CXR, and degree of lung involvement (CXR quadrants). These analyses suggest that the subgroup used for the transcriptomic profiling is representative of the full cohort.

We also included a new Supplementary Table 6 with a summary of these bivariate analyses (entitled "**Comparison of participant characteristics of transcriptomic subgroup and full cohort**")

We added the following text to the Results section:

"We found no differences in clinical and biologic variables (age, sex, HIV status, Xpert Ct, Cavitary CXR, and degree of lung involvement (CXR quadrants) when comparing participants included in the transcriptomic subgroups (CAC+ and CAC-) with the remaining cohort (Supplementary Tables 6 and 7)."

- It should be made clear what age range was studied, and the age range the authors are comfortable for their prediction model to be valid for.

Response: For the entire study, we did not exclude any age groups and had a range of 18 to 100 years. For the predictive model, we also did not exclude any ages and included the same age range.

We edited the Results to read as follows:

"To examine the biology of Mtb aerosolization and transmission, we enrolled 142 individuals with microbiologically confirmed pulmonary TB in Nairobi, Kenya. The median age of participants was 35 years (interquartile range (IQR), 27-44) and ranged from 18 to 100 years; 27% were women (Table 1)."

In regards to questions about age and use of the prediction tool, we are not recommending its use as it requires external validation. As we have noted in the Discussion:

"While promising, the risk score that we present requires external validation and should not be used at this time for clinical (non-research) purposes."

- L298: With the 3 Xpert categories, variables are no longer all dichotomised, suggest to revise.

Response: Thank you for this suggestion. We corrected the sentence to read as follows:

"Included variables were categorized and the optimal cut-points were determined using Youden index (*J*) method,"

Reviewer #2 (Remarks to the Author):

The authors addressed all my questions and also the questions of the other reviewers. I am satisfied with the review and recommend the publication.

Response: We thank the reviewer for these positive comments.

Reviewer #3 (Remarks to the Author):

the manuscript is interesting and great value scientifically and clinically. the authors answered the questions raised.

Response: We thank the reviewer for these positive comments.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

I thank the authors for working through the comments. It is an interesting study with results that should go into the wider scientific discourse, and I think has all the information that engaged readers need to judge the finding.

Reviewer #4 (Remarks to the Author):

I was asked only to review the X-ray component of this study. The team obtained posterior anterior chest x-rays from all participants (plus lateral for children <10), which were scored by one member of the study team who is a pulmonologist, according to: presence/absence of cavitations and number of quadrants with abnormalities.

With regards to this, they found:
-participants with TB identified through passive case finding were less likely to have cavitary TB (expected)
-cough aerosol culture positive patients were more likely to have cavitary TB
-TB-related abnormalities in more chest X-ray quandrants was associated with cough aerosol culture positivity

These findings are expected and are clearly described. I have no major concerns with this. The only question is whether the reader was blind to the study results when assessing the chest x-rays, which should be the case. If so, they could clarify this in the methods.

Unrelated to x-ray:
Line 359-360 “GeneXpert cycle threshold and semi-quantitative grade were significantly higher among cough aerosol-culture positive compared with cough aerosol culture-negative participants.”
-this should be GeneXpert cycle threshold was lower and semi-quantitative grade was higher..” Lower PCR cycle thresholds correspond to higher semi-quantitative grade. They also on line 363 show that the Ct relationship was cough aerosol status was negative as expected as lower Ct values correspond to higher bacillary load.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I thank the authors for working through the comments. It is an interesting study with results that should go into the wider scientific discourse, and I think has all the information that engaged readers need to judge the finding.

Response: Thank you for your time, knowledge and insights in raising the quality of our manuscript.

Reviewer #4 (Remarks to the Author):

I was asked only to review the X-ray component of this study. The team obtained posterior anterior chest x-rays from all participants (plus lateral for children <10), which were scored by one member of the study team who is a pulmonologist, according to: presence/absence of cavitations and number of quadrants with abnormalities.

With regards to this, they found:
-participants with TB identified through passive case finding were less likely to have cavitary TB (expected)
-cough aerosol culture positive patients were more likely to have cavitary TB
-TB-related abnormalities in more chest X-ray quandrants was associated with cough aerosol culture positivity

These findings are expected and are clearly described. I have no major concerns with this. The only question is whether the reader was blind to the study results when assessing the chest x-rays, which should be the case. If so, they could clarify this in the methods.

Response: Thank you for your review. We have edited the manuscript to state,

"DJH was blinded to CASS results in the review of chest x-rays."

Unrelated to x-ray:
Line 359-360 "GeneXpert cycle threshold and semi-quantitative grade were significantly higher among cough aerosol-culture positive compared with cough aerosol culture-negative participants."
-this should be GeneXpert cycle threshold was lower and semi-quantitative grade was higher.."
Lower PCR cycle thresholds correspond to higher semi-quantitative grade. They also on line 363 show that the Ct relationship was cough aerosol status was negative as expected as lower Ct values correspond to higher bacillary load.

Response: Thank you for identifying this error. We have edited the sentence to read,

"GeneXpert cycle threshold was significantly lower and semi-quantitative grade significantly higher among cough aerosol culture-positive compared to cough aerosol culture-negative participants."