Author’s response to reviews

Title: A low clustering rate with predominance of Haarlem strains in patients without known risk factors for multi-drug resistant tuberculosis in North-Eastern Lima, Peru

Authors:

Francesca Barletta (francescabarletta@yahoo.es)
Otero Larissa (larissaotero@gmail.com)
Jimena Collantes (jimena_collantes@hotmail.com)
Belisa Asto (baar_17@yahoo.es)
Bouke C de Jong (bdejong@itg.be)
Carlos Seas (cseas2@gmail.com)
Leen Rigouts (lrigouts@itg.be)

Version: 2 Date: 6 April 2013

Author’s response to reviews: see over
Cover Letter

April 05th, 2013

BMC Infectious Disease

To the Editor:

Thank you for the thoughtful review of the manuscript “A low clustering rate with predominance of Haarlem strains in patients without known risk factors for multi-drug resistant tuberculosis in North-Eastern Lima, Peru” (MS: 2758593168318009). We have incorporated all the suggested changes and answers in the revised manuscript.

Enclosed please find the documents “Answers to Reviewer 01”, “Answers to Reviewer 02” and the manuscript: “Predominance of Haarlem strains in Peru – Revised”.

Sincerely,

MSc. Francesca Barletta Solari
francescabarletta@yahoo.es
Laboratorio de Enfermedades Entericas y Nutricion
Unidad de Epidemiologia Molecular en TB
Instituto de Medicina Tropical Alexander von Humboldt
Universidad Peruana Cayetano Heredia
Answer to Reviewer #01:

Reviewer's report (Midori Kato-Maeda):

**Major Compulsory Revisions**

**Abstract:**

1. In the conclusions section please clarify what is the meaning of –early introduction-.

   We have rephrased the idea to make it more clear.

2. The current study can’t be used to conclude that a low proportion of TB within this study population was due to recent transmission. The 7 month study period is too short to study events of recent transmission with rapid evolution to active TB. Also, the study population was not community-based and the patients with risk of drug resistant TB that were excluded from the study may be in cluster or not with patients that do not have risk for drug resistant TB

   We agree with the reviewer that a seven-month period is too short to study recent transmission therefore we have suppressed this subject from objectives and conclusions.

**Introduction:**

1. Please include that one of the aims of the manuscript is to determine the factors associated with MDR and isoniazid resistant, which is detailed in your tables 2 and 3.

   Our objective was to identify predominant circulating TB lineages and to investigate the drug resistance patterns associated with these strains among patients without known risk factors living in a hyperendemic neighbourhood of Lima. Tables 2 and 3 have been modified and the information has been combined in only one table. The epidemiological factors associated with MDR and H-resistance were previously described by Otero et al., 2011.

2. Please describe what information is added in this study, that it was not reported in the published paper from Otero L, Krapp F, Tommatis C, Zamudio C, Matthys F, Gotuzzo E, Van der Stuyft P,
Seas C. High Prevalence of Primary Multidrug Resistant Tuberculosis in Persons with No Known Risk Factors. PLOS ONE 2011. 6(10):e26276.

_Otero et al._ found a high rate of primary MDR-TB in a general population with no identifiable risk factors for MDR-TB in a high burden area, but it remained to be determined which are the circulating strains and their possible association with drug resistance.

3. Please describe why describing the _M. tuberculosis_ diversity (distribution of lineages) in patients without risk factors for drug resistant tuberculosis is important.

There is evidence of a high rate of primary MDR-TB in this population with unknown risk factors, therefore we wanted to determine whether clonal spread indeed was an issue.

**Method**

1. Study setting and study population. In order to assess the burden of TB due to recent transmission in North Eastern district in the Lima Province in Peru, all (or almost all) patients with culture positive tuberculosis should be included, as transmission can occur between patients with and without risk factors for drug resistance. Therefore the study population chosen can’t be used to study the burden of recent transmission. Please reconsider the study population if you would like to assess the burden of TB due to recent transmission.

_We agree that not only the time period, but also the selected population were not adequate to study recent transmission among the general population therefore we have suppressed this subject from objectives and conclusions._

2. Data collection: Authors collected information about treatment outcome but it is not discussed in any part of the manuscript. Please describe.

_As this study population was meant to be the start of a prospective cohort study, collection of treatment outcome data was part of the set up. We agree, however, that for our part of the study this information is not relevant therefore we have suppressed it._
3. Fingerprint analysis: Please describe the goal of generating the dendrogram and describe the dendrogram in the result section.

The purpose of the dendrogram is to illustrate the clustering between the isolates. Given the fact that MIRU-VNTR did not reveal any clustering in our study, the dendrogram does not provide relevant information therefore we have decided to suppress it.

4. Statistical analysis.

a. Please describe the relationships that were tested with the statistical analysis.

b. Please clarify in which analysis the lineages the outcome variable.

Frequencies of lineages were calculated and this was the outcome variable. Categorical variables were analysed using chi square test (Epi Info v7; Georgia, USA). Risk ratios and 95% CI intervals were calculated for resistance and lineages.

5. Ethical considerations: Please describe if all the patients that were eligible consented. If not all eligible accepted participating in the study, please describe those that did not accept to determine the representation of the study population.

The patients included in the prospective cohort of new cases of sputum smear-positive pulmonary TB - in which this study is embedded - represented 78% of the eligible population. The 16% that were not included did not want to participate in the study or they were not reached.

Results

1. Please review the numbers.

a. 50 excluded from 314 (50/314) is not 24.8%. Also, 314-50 is not 254.

Math calculations have been corrected.

b. The result section of the manuscript describes that 147 were grouped in 22 spoligo-based clusters; however in the figure 1, there were just 146 isolates that were in cluster.

Math calculations have been corrected.
2. Please show the results related to the analysis of lineages and its association to age, sex, or residency, so the reader can understand the comments described in the discussion section.

The variable ‘residency’ does not provide relevant information therefore we have suppressed it. We have also added the information in the section of Results.

Discussion

1. The authors mention that the study illustrated the possible relatedness of the circulating genotypes with drug resistant patterns and patients characteristics. Please comment on the fact that none of these associations were statistically significant, which may be in part because of the small sample size.

A comment has been included.

2. The discussion regarding the molecular epidemiology of TB can’t be supported by this study which is not designed to determine the frequency of recent transmission. Although in the limitations the authors mentioned that clustering rates could be underestimated due to the relatively short study period, it is also because the study population is not appropriate: patients with risk factors for drug resistant TB may be in cluster with those that have no risk factors, which were excluded from the study.

Our study had some limitations. First, since this was a passive surveillance study only patients attending health care facilities at the public sector were included. Second, the selection of patients without high risk factors relied on patient’s report, which could have masked some risk factors, hence overestimating the MDRTB rates among low risk patients. Third, clustering rates could be underestimated due to the relatively short study period and the exclusion of patients with risk factors.
3. The comment related to the widespread occurrence of the Haarlem lineage in this setting could be attributed to its higher stability and/or relative superior fitness is very speculative, and out of context and should be deleted.

This comment has been suppressed.

4. The table 1 is mentioned when discussion the associations of lineage with age. However, table 1 describes the frequency of drug resistance. Please add the table, and discuss it in the results.

Correct referral was made in the text.

5. The authors comment that: “A high MDR rate (7.5%) was found among smear-positive pulmonary TB patients without previous treatment history, or unknown risk factors for MDR”.
   a. Smear status and previous treatment history were not described in the results or in tables, please add the information in results.
   b. Please describe what are the comparison groups to state that a high MDR rate was found among (patients) with unknown risk factors for MDR.

In methods we established that we were enrolling patients that had BK+ and that have not had TB. Regarding the comparison groups to state a high MDR rate, we included a comparison of our study group with data of the general population on a national level.

Summary:

1. Authors describe that their findings suggest ongoing transmission: however, the methodology of the study can’t support this observation.

2. The next sentence: “spread of MDR-TB occurs at low rates from multiple patients with Untreated MDR-TB” can’t be supported by this study.

3. Please elaborate and clarify why “Predominance of the Haarlem lineage suggests for the long presence of this lineage in Peru”

4. Please elaborate and define what is the ‘distinct introduction mechanism of the Harlem lineage to Peru. Figure 2 the dendogram related findings are not commented in the results no in the
discussion. Please explain what is in each of the yellow squares, what are the questions marks, where are the MIRU15 results (the text refers to figure 2 to see the MIRU). Table 2 and 3.

1. Please identify the referent groups.

2. The tables show the factors associated with MDR or INH resistant. This should be described in the aims.

Reference groups have been clarified in the tables.

Minor Essential Revisions

Abstract:

1. In the method section, please add how transmission was measured and how MTB diversity was defined.

We agree with the reviewer that study design is not the correct to study recent transmission therefore we have suppressed this subject from objectives and conclusions.

2. The results section says that spoligotype analysis identified clustering among 166 of 199 isolates. However, the result section of the manuscript says that 147 were grouped in 22 spoligo-based clusters. Please clarify.

We have corrected and clarified the numbers.

3. In the result section, please add the total number of cases included.

We have included this information.

Methods:

1. Spoligotyping: Please include the reference formatted as required by the journal.

We have corrected it according to the requirements of the journal.

2. Fingerprint analysis:

a. Please change subtitle to DNA fingerprint analysis.

b. Please clarify how demographic information was compared with the international spolDB4 Data base. If this was performed, please describe in the result section.
c. In the same sentence reads as MIRU VNTR plus software was used to compare with the international Spold DB4.0. Please rephrase. Please review and rephrase the sentence: “Identical spoligotypes and MIRU-VNTR patterns were considered to be genotypically clustered a cluster”.

**We have corrected the subtitle. For the analysis we have introduced spoligotypes, demographic data and DST results into MIRU-VNTRplus software. We have reviewed and rephrase the sentence cited.**

3. Statistical analysis

a. Please clarify what was subjected to the quality control described in this section.

b. Please provide the package used for the statistical analysis.

**This information has been included.**

4. Please add a section with definitions of the following terms:

a. TB due to recent transmission.

b. Alcohol and drug abuse.

**The topic of “recent transmission” has been suppressed of the manuscript.**

Patients were asked if they had a diagnosis of alcoholism, and if they had used illegal drugs.

**Results**

a. Please move the following sentence to the end of the paragraph as it refers to the spoligotypes and the sentences before and after are referring to the lineage:

“Among the identified strains, 147 (87.0%) were grouped into 22 spoligo clusters, with cluster sizes ranging from 2 to 34 (Figure 1)”.

b. The percentage of resistant to S is 8.6% not 8.5%.

c. Figure 1. There are just 24 ST in the LAM lineages and the text describes 25. Please clarify.

**Math calculations have been corrected.**

Discretionary Revisions
Methods

1. Spoligotyping: this is a standard method so the methodology can just be referred.

2. MIRU VNTR: this is a standard method so the methodology can just be referred.

We have referred both techniques as suggested.
Reviewer: M. Cristina Gutierrez

Reviewer's report:

A low clustering rate with predominance of Haarlem strains in patients without

known risk factors for multi-drug resistant tuberculosis in North-Eastern Lima,

Peru

BMC Infectious Diseases

This study focuses on a selected population of adults (>18 year-old) with new smear-positive

pulmonary TB and without known risk factors for MDR-TB, living in a geographical area with high

exposure to TB. The aim is to assess the burden of TB due to recent transmission, to identify

predominant circulating TB lineages and to investigate the drug resistance patterns associated with

these strains.

These are important questions to be solved for understanding the TB epidemiologic patterns in this

community and improving its control. The authors present original data in a properly structured and

well written manuscript. The study has efficiently determined the population structure of the

Mycobacterium tuberculosis isolated from this group of TB patients. However, this study design

fails to surely assess the burden of TB due to recent transmission. The accuracy of resistance testing
to rifampin remains to be confirmed.

Major Compulsory Revisions

1. My main concern is that the studied sample is not sufficient to infer the rate of recent

transmission of TB within this population in North-Eastern Lima. In its present format the study

could be leading to misinterpret that the rate of recent TB transmission for this population is low, in
spite of living in an area of high exposure to and transmission of TB. The authors should suppress this subject from the objectives and conclusions, carefully review the comments in the discussion, and focus the manuscript on the population structure and resistance of MTB population. Specifically, a seven-month period is too short for a molecular epidemiologic study of TB transmission in the general population; the sample is not exhaustive, representing only 16% of TB cases notified in San Juan de Lurigancho, the studied area of North-Eastern Lima (which notifies more than 2130 TB cases per year); the sample is biased, only MTB strains from a selected group of TB patients within the general population are compared, whereas transmission could have occurred between the selected population and the rest of TB patients in the area - the suspects to have a MDRTB-; <18 year-old patients are excluded, whereas TB in youngest people is an indicator of active transmission. The need of future studies to accurately estimate the rate of recent TB transmission could be mention in the summary, together the identification of risk factors.

We agree with the reviewer that seven-month period is too short to study recent transmission therefore we have suppressed this subject from objectives and conclusions.

2. The authors indicate that they used 40ug/ml of rifampin to test on Middlebrook 7H10 medium the susceptibility of the MTB strains. However the standard final concentration to test this drug on 7H10 is 1.0 ug/ml. Which was the real concentration that they tested? If 40ug/ml, the interpretation of the results for this drug and their conclusions are not valid, and the rate of rifampin resistance and MDR could be underestimated.

We have corrected the concentration of rifampicin.
3. Methods, Fingerprint analysis. Please review the first phrase, it is unclear the process followed to compare the patterns as only the spoligopatterns can be compared to SpolDB4.0. In the same line, it is unclear along the description of the results whether the authors determined the spoligotyping family for the strains using only MIRU-VNTRplus (as it is shown in Figure 2) or they also checked in SpolDB4.0 the strains that did not match the spoligotypes of the reference strains of MIRU VNTRplus. The process to assign strains to one family should be described in detail in Methods.

**For the analysis we have introduced spoligotypes, demographic data and DST results into MIRU-VNTRplus software. We have reviewed and rephrase the sentence cited.**

4. It is unclear whether MIRU-VNTR typing was performed with all the 199 strains or only with strains clustered by spoligotyping, this point should be clearly indicated in the text. Additionally, the manuscript provides scarce information on MIRU-VNTR results, which should have been more exploited. E.g., were all MIRU-VNTR profiles congruent with the family classification by spoligotyping? A situation that makes the analysis more robust and can be verified using MIRU-VNTRplus.

**We have rephrased in the M&M and Results sections to make this more clear.**

5. Percents in the text need revision. For instances, in Results page 6, 50 patients is not 24.8% of 314. I also count 44 H-resistant strains in Table 1, this corresponds to 22.1 % instead of 14.4% indicated in the abstract.

**Math calculations have been corrected.**

6. It is indicated that LAM strains were less likely to be H resistant than the other lineages (0.14 vs. 0.19 in the abstract; 0.84 RR, 95% CI 0.19-2.56 in the results). This data seems to have not
statistical significance. Table 3 shows lower RR for family U strains (0.43)? P values should be
added to Table 3, and the lack (or not) of significance mentioned in the text.

**We have combined tables 2 and 3 in one table to make it more clear and none of these associations were statistically significant.**

**We have also added the p-value.**

7. The manuscript is in some instances ambiguous regarding the covered population. Along the entire text, the authors should make clear that suspected MDRTB patients were excluded, and conclusions concern only TB patients without risk factors for MDRTB. Idem for the structure of the MTB population and the predominant circulating families, which cannot be extrapolated to the entire population, particularly because without additional data it cannot be excluded that MDRTB cases could be due to a predominant strain.

**We have re-emphasized this in the Results and Discussion section.**

**Minor Essential Revisions**

1. Methods, Culture and drug-susceptibility testing, please indicate MTB in full.

**We have corrected it.**

2. Methods, MIRU-VNTR. “MIRU-VNTR is a PCR-based typing method … to be polymorphic in MTB”. Please review the phrase, in the present format it refers to only 15-locus MIRU-VNTR typing, whereas the method is independent of the number of loci tested.

**We have rephrased it and written more clear.**
3. Methods, Fingerprint analysis. Please note that double alleles at 2 or more MIRU-VNTR loci can be considered not only mixed infections but also possible cross-contaminations. Additionally, in the phrase “Identical spoligotypes and MIRU-VNTR patterns … clustered a cluster” something is missing, “clustered in a cluster”? 

**We have rephrased it.**

4. Table 1, polyresistant (non-MDR) H+E is indicated twice.

**We have corrected.**

5. Table 2, there seems to be an error in the RR (95% CI) for either Sex or Residence as they are identical.

**The RR (95% CI) showed in table 2 corresponded to Sex.**

6. Legend to Figure 2 indicates that the dendrogram includes spoligotyping and MIRU15 patterns for the 199 isolates, however the tree has been constructed only based on spoligotyping data, and the MIRU15 patterns are not shown.

**The purpose of the dendogram is to illustrate the clustering between the isolates. Given the fact that MIRU-VNTR did not reveal any clustering in our study, the dendogram does not provide relevant information therefore we have decided to suppress it.**

7. Reference n. 24, the title is missing in Bibliography section.

**We have replaced it for the article published by the same research group after the conference.**

**Discretionary Revisions**
1. Table 2: Characteristics and MDR status. I would recommend to classify the population into 2 groups, 30 year-old or less, and more than 30 year-old, and reanalyze according to MDR status by chi square test as it looks to be a tendency to MDR (+) status among younger people.

**Tables 2 and 3 have been combined in only one table.**

2. Figure 2 adds little information to the already provided in Figure 1 that is descriptive enough. It could be more useful a figure focusing only on the orphan spoligotypes because specific ST patterns can be easily checked in SpolDB4.0.

**The purpose of the dendogram is to illustrate the clustering between the isolates. Given the fact that MIRU-VNTR did not reveal any clustering in our study, the dendogram does not provide relevant information therefore we have decided to suppress it.**