Author's response to reviews

Title: Early detection of tuberculosis through community-based active case finding in Cambodia

Authors:

Mao Tan Eang (mao@online.com.kh)
Peou Satha (peousatha@yahoo.com)
Rajendra Prasad Yadav (yadavr@wpro.who.int)
Fukushi Morishita (sketchbook1006@hotmail.com)
Nobuyuki Nishikiori (nobu.nishikiori@gmail.com)
Pieter Van Maaren (vanmaarenp@wpro.who.int)
Catharina Van Weezenbeek (vanweezenbeekc@wpro.who.int)

Version: 2 Date: 23 April 2012

Author's response to reviews: see over
Response to Reviewers Comments:

We appreciate extensive comments provided by two reviewers on our manuscript and thankful for the time and effort given by them. We strove to incorporate that feedback and we believe our manuscript has improved to fulfil the quality required for the journal publication.

Please find below point-to-point response to the reviewers’ comments.

Reviewer #1:

<table>
<thead>
<tr>
<th>Discretionary Revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
Our intention by using early case finding is not related so called “Enhanced Case Finding” elsewhere. By “early” case finding, we simply meant early detection of cases. We changed the text accordingly to avoid confusion. We also changed the phrase “vulnerable groups” to “a vulnerable age group” which, we believe, can be easily assumed from our argument in the text.

These data indicate that (slightly) older TB patients are found through ACF, can you really conclude that therefore “patients from vulnerable populations and age groups” are found? This seems to be a slight stretch from a single finding to multiple groups.

We agree with the referee’s comment and have changed the phrase to “a vulnerable age group” in the conclusion sections of both abstract and main body.

To me it is noteworthy that there were not more differences found— gender, children, etc. (did you look only a children to see if they were picked up more with ACF?). I am curious as to why patients under 15 were excluded from the study? Please elaborate on why more differences were not found.

The patients under 15 were excluded from the study simply because these ACF sessions specifically targeted adults. As mentioned above, we have a different type of ACF approach with TST focusing on childhood TB but this paper did not examine that part. Hopefully we can conduct a similar study to evaluate ACF among children.

The authors do not say why there have been “general recommendations against community-ACF” in the past, and the reader is left wondering why WHO would have opposed such a logical step.

We agreed that this is an important logical step, and addressed this by adding a short statement. (introduction/3rd paragraph)

Is this correct that that there was a single CXR reader for all cases? Is it correct to assume that PCF patients did not get a CXR? If so, it may be clarifying to readers to state this explicitly. It is not clear to this reader how PCF patients were diagnosed.

Yes, this is correct. The CXR reader is a leader of the CENAT ACF team, directing the radiography unit at CENAT. For PCF cases, diagnosis of smear-negative TB is according to the national guidelines. Some were self-referral patients presenting to the health centre where only a sputum smear test is available. After negative AFB, those who are still suspected with smear-negative TB should be referred to district level hospital for CXR and diagnosis. In short, PCF cases had access to CXR at referral level but some suspects might have dropped during the diagnostic process.

Figure 2 would be easier to appreciate if it were collapsed into one graph (like Fig 3) because the reader can compare within age stratum, which is the intended message of the image.

We appreciate your feedback on the figures. We agree that there may be some moment needed for readers to compare three stacked graphs (Fig 2). We have tried several ways in line with reviewers’ suggestion but we felt that they were not necessarily advantageous over the original one. The quantity the graphs represent is a proportion of cases within the group. Age distribution patterns (shapes of distribution) are more important than comparing percentage figures within age group.
The conclusion is that further research is required to determine if benefits outweigh the costs. If that is so, then how can you simultaneously suggest that cost effectiveness is established? I would suggest not to call for more cost effectiveness research, but rather to propose more research on the impact of ECF/ACF on reducing on-going transmission by detecting and treating early silent transmitters.

a. How do the authors know that the initial defaulter rate is already low in Cambodia? The complexity of initial defaulters is that they are rarely recorded in the TB treatment register and so it is practically impossible to gauge how large or small the magnitude of the problem is by looking in the TB tx register unless one compares a TB suspect register (does Cambodia maintain a TB suspect register?) against a lab register. The citation given as evidence that Cambodia’s initial defaulter problem is small is the regular WHO annual report, which does not report such special studies.

We fully understood the referee’s assertion, and the text was a little modified based on the suggestion (abstract and conclusion).

[Note: The reason behind why we mentioned “benefits outweigh the costs” was that we wanted to call for more rigorous cost-effective/benefit analysis of ACF which clarifies all the monetary benefits brought about by ACF. This may be possible, in the future, by calculating, for example, a logistics cost saved, an opportunity cost saved and a cost per DALY averted among index cases identified through ACF as well as these costs among possible secondary cases averted as a result of ACF. In that sense, the proposed conclusion that “more research on the impact of ACF on reducing on-going transmission by detecting and treating early silent transmitters” was one of our unwritten messages, but was difficult to be picked up.]

What we meant to emphasise was that the defaulter rate in a routine programme (PCF) setting at national level was very low in Cambodia therefore the WHO annual report was cited. Since, as the referee mentioned, it was very difficult to obtain a comparable indicator (initial defaulter rate) for PCF in this study, we did not look at the initial defaulter rate but the defaulter rate at national level. As we wrote in the discussion section, high initial defaulter rates in ACF were reported in India and South Africa. The defaulter rates at national level in these two countries were also very high according to the WHO report, which could support our argument. Furthermore a defaulter rate is available almost everywhere even at provincial and district level. Thus we believe that it is more feasible to use a defaulter rate (and even other programme indicators) as a benchmark to estimate the magnitude of initial default in ACF when planning ACF activities.

Why is HIV status not included in Table 1 – too low numbers? 

HIV status is included in the last row in Table 1.

Why are women not given as a category in Table 1? Why are the PCF-post ECF patients not listed in Table 1 to allow for comparison?

A binary variable such sex and HIV status takes on only two values. Therefore only one of the two was included in the table so that another can be simply calculated. This is one of the conventional techniques to draw up a table in order to maximize the space of a scientific paper.

We initially considered including the PCF-after group in the table. After deep deliberation, however, we decided not to include it to make the table simple so that readers can easily interpret what is presented.

Would the authors consider producing a comparative table of socio-demographic characteristics of TB suspects identified by the 2 methods, similar to Table 1? This could be instructive to readers to help to perceive some of the types of persons who are not seeking care.
In this study, we used the existing programme records to retrospectively review the registered TB cases, and the information taken from them was quite limited. We no longer keep track of the reviewed TB cases because their treatment outcome had already been confirmed. Therefore it was not feasible to obtain the additional information on their socio-demographic characteristics on an individual basis, which requires much more time and resources. Yet we strongly agree that this is an important area that requires further investigation.

The authors say that ethical clearance was not required and therefore it was not done. Wouldn’t it be slightly more appropriate to also clarify that because no identifying information was collected there was also no risk of deductive disclosure or social or physical harm to individuals that is why no ethical clearance was sought?

We agree with the referee’s suggestion and employed the proposed text (quantitative data/2nd paragraph).

A little additional detail on the statistical methods would be illuminating for the reader. Much of the data are ordinal and not normally distributed. Would chi-square be the best test in this instance?

We applied chi-squared test as a test for independence. In general, assumptions to meet for the test include: (a) simple random selection, (b) overall sample size, and (c) not too small expected cell count. We believe our data presented in Table 1, fulfilled these assumptions, except (c) for some variables. For those breaching (c), we applied Fisher’s exact test. In response to the suggestion from the other reviewer, we have also conducted non-parametric test to compare age distribution between groups (Man-Whitney test). We have included these additional details in the text.

### Minor Essential Revisions

1. The definition of ACF given in the intro “systematically looks for cases of TB, rather than waiting for people to develop...” is vague and appears to be an adapted version of the Royal College of Physicians standard definition. More problematically it overlaps with the definition of “enhanced case-finding”. To me authentic ACF implies two related pre-conditions:
   - No health seeking whatsoever on the part of the patient.
   - Sample collection on the spot – i.e. at the home.

   A prevalence survey is ACF. The intervention described in this paper does meet part of the 2nd criterion because of the mobile CXR, but not the first. In my opinion, the intervention described in the manuscript is rather an elegant form of contact investigation, i.e. systematic mobilization of TB contacts to promote health seeking, via traditional facility-based diagnosis. or perhaps ECF: “enhanced case finding” but not ACF per se. When the authors use the term “campaign” in the last Para of the discussion, I think they are correctly describing the intervention and its social mobilization dimension.

After long deliberation and discussion, we left the term ACF as it is because:
   - We believe that the definition of ECF has not been established yet;
   - Cambodia has been conducting this activity since 2005 and addressing this as ACF inside and outside countries including partners and donor communities. We do not want to confuse our partners and donors;
   - The ACF session in this Cambodian approach is quite distinctive from routine health centre function. It required a dedicated expert team to visit the site, setting up the venue, systematically handling participants. The invited TB suspects attended a special ACF session knowing it would be an TB screening session. They are not visiting a
health centre for routine health service as part of health seeking behavior.

We clarify in the text that ACF sessions were dedicated special session, not part of routine health centre activities.

2 Passive and active case finding are not dichotomous as implicitly suggested in the manuscript but rather points on a continuum of patient effort. The diagnostic mobilization contact investigation intervention described in the manuscript falls somewhere in the middle of the effort spectrum, because symptomatic persons were screened locally but still needed to be motivated enough to travel to the facility and go through the traditional systems for diagnosis. I recommend to re-title the paper and to use the term “Enhanced Case Finding” or “contact investigation” throughout to describe the intervention that was conducted.

As described above, we would like to keep our paper to be presented as an evaluation of an ACF approach that has been conducted in Cambodia for more than 7 years.

By the way, the comment above was not precise by saying “symptomatic persons were screened locally but still needed to be motivated enough to travel to the facility and go through the traditional systems for diagnosis”. The participants did not go through the traditional system, rather going through a dedicated, special ACF session for intensified/early diagnosis.

3 I think this ECF study describes an approach that is very novel and important and should not be subsumed in the ACF category, because it is very likely that in systematic lit review of ACF studies, we will find that ACF is NOT cost effective for most NTPs. However, a very strategic, middle-ground approach, such as the one described in this manuscript, is clearly feasible, has high yield, and is probably cost effective. Lumping this innovative approach in with other miscellaneous forms of ACF will make it harder to convince policy makers of its value because it will tar with the same brush of cost ineffectiveness.

We highly appreciate referee’s positive interpretation and suggestion to promote our unique approach.

4 Please tell readers what were the proportions of refusal and lost to follow up at each step on the diagnostic pathway if possible.

The registration of TB suspects for ACF screening was processed at respective health centres at the timing of the individual consultation before X-ray screening. Therefore the number of TB suspects who were advised to present to health centres was not recorded. Among those registered for ACF screening, no suspects refused to take X-ray screening. Among those with X-ray abnormal results, no suspects refused to take sputum-smear tests. All 21 patients (initial defaulters) dropped out after the clinical decision was made and before patient registration for treatment. We briefly explained this in the text (result/treatment outcome).

5 Please clarify what is meant by “were involved” the first sentence of the result section. You call these 33,631 TB suspects, but how many were symptomatic?

TB suspects presenting to the health centres include TB contacts (symptomatic and asymptomatic) and TB symptomatics from neighbours. Unfortunately, ACF register did not record whether subjects were symptomatic or not as all subjects went through the same procedure during the ACF sessions.

All TB suspects presenting to the health centres were registered for ACF screening, as explained above., which means 33,631 suspects were registered for ACF screening. They all underwent individual consultation and X-ray screening, and then sputum smear tests according
to their X-ray results. We clarified this point in the manuscript (result/NNS).

6 Please recalculate the NNS according to first screening step i.e. symptom screen. As I understand it,
a. The first screening step was a symptoms screen (n=?). b. Then CXR (n=33631) c. Then sputum microscopy was done (n=5844) d. Then 885 cases were found
e. So the real NNS would be = [885/# of symptomatics]. Not 38 (=885/33631) as stated.
1. Please also report the NNS for the particular study population also. If this unknown, perhaps give the number of villages, number or days, or number of households as some sort of proxy.

We understand the reviewer’s concern and suggestion to make NNS calculation clearer. However, we have some problems in following the steps suggested. First, the participants presented to the ACF session were combination of TB contacts and TB symptomatics. For TB contacts, all family members who had contacted with smear-positive TB cases were invited to the ACF sessions. During the same process of identifying family contacts, symptomatic neighbours were also invited to ACF sessions. Ideally, we should have a number of individuals contacted/interviewed by community workers during the first step in the community. However, we did not record it.

Our NNS calculation starts from those enrolled in ACF sessions in the health centres.

7 Since CXR seems to be the main screening tool, it would be essential to indicate what were the “predefined criteria” used to classify the CXR.

Based on the radiographic image shown in the film, the radiologist initially classified the findings into five categories that are ‘normal’, ‘active’, ‘suspect’, ‘healed’ and ‘others’, estimating their likelihood of being related to TB or non-TB conditions. While those in the ‘active’ category moved straight onto sputum-smear microscopy, those in the categories of ‘suspect’, ‘healed’ and ‘others’ were further evaluated based on the individual risk of infection including clinical history to see if sputum-smear microscopy is required for further clinical investigations. We clarified this in the manuscript (method/programmatic information/1st paragraph).

8 The term “ACF session” is new to some of us and should be defined. Is a’ session’ a particular program effort? Or is it time bound? How does it differ from the term “13 ACF activities were conducted”? Is this equivalent to “the massive ACF campaign” referred to in the last Para of the discussion?

Thanks for pointing this out. After the discussion on our side, we decided to use the term "session" consistently throughout the manuscript except the one describing generic activities so that we can avoid confusion.

9 “Initial defaulters” should be defined. Do the authors refer to persons diagnosed who never start treatment (i.e. standard definition) or those who began treatment and then desisted?

Yes, we followed the standard definition. Initial defaulters are those who were on the list of TB patients as a result of ACF screening but never start treatment. We clarified this in the manuscript (result/treatment outcome).

10 In addition to the p values given in the abstract, it would be instructive for readers to have the 95% Confidence Intervals for the estimates for the main comparisons. This additional information helps readers to determine for themselves if the differences are clinically and/or programmatically meaningful or not.
Agree with referee on the usefulness of confidence intervals for the main comparators. We assume the referee is addressing **main outcome variables** by main comparators (e.g. percentage with disease, etc). In our study, as in Table 1, the table is raid out to compare exposure (explanatory) variables in two groups (somehow, similar to a case-control study). In this type of table, confidence intervals of each category of exposure variables can be included but will make table complex (e.g. CI of all percentage in each age group category). We would like to opt for a simple expression of the table as in the original manuscript.

11 Rather than say “high treatment success rates”, which takes up a lot of text characters in the abstract and is not optimally informative, just give the figures- 94.6% and 95.2%.

Thank you for the feedback on this. We changed the text as you suggested (abstract/result).

12 The only study limitations mentioned are issues that I do not find to be limitations at all. For example, the issue of quality of smear. Since it applies to both PCF and ECF equally, I fail to see why it would be a limitation in a comparative study. I suggest striking this text and instead focusing the discussion on the other limitations.

To my mind, the potential limitations of this study are as follows:

a. I suspect that no accurate records were kept on the number of TB suspects identified by symptom screen at the household level, so the universe and the refusal rates are unknown. This is the major limitation and should be acknowledged because it means that yield cannot really be calculated.

b. The differences in the sensitivity and specificity of the 2 different diagnostic algorithms applied to ECF and PCF suspects are a potential confounder, but this is onlyopaquely referred to in the limitations section.

c. Day-time home-based mobilization will naturally find fewer persons in the economically active age group, and this study is no exception. This selection bias slightly undermines the argument about ECF detecting older age groups. Night-time home based mobilization might find just as many younger TB patients to void the difference detected.

d. This ECF effort still may have missed TB cases among symptomatic suspects who could not travel or would not travel to the facility. Therefore the true benefit of ECF may be underestimated by the logistical requirements.

e. The exclusion of children probably diluted the beneficial effect of ECF as there is good evidence from the Cambodia DHS that families do not routinely bring children with cough and ARI symptoms to care in a timely manner.

We agree with the referee on the point that this may not be a study limitation from the methodological aspect because it applies both case finding methods. But, the reason why we mentioned the limitation of various factors influencing smear grades was not to clarify possible confounders between ACF and PCF, rather to support and even protect one of our important argument and logic linking lower smear grades to early case detection. Otherwise some might say that such factors may greatly vary on an individual patient/slide/lab basis and smear grades cannot be a proxy indicator to determine the timing of diagnosis. Hence we believe that it is essential to explain this explicitly.

We appreciate the list of referee’s suggestions on possible important limitation. We incorporated some of them into our manuscript (discussion/4th paragraph). Although the limitation regarding the suggestion (c) was, we believe, already mentioned in the different paragraph (discussion/2nd paragraph), we added a quick explanation to make it clear that the economically active age group might not be detected.

13 One of the conventional arguments against ACF has been that the treatment completion will be
low due to low motivation, lesser severity, etc. etc. This study seems to provide powerful evidence against that criticism. This could be better highlighted more explicitly in the discussion and the abstract.

a. Tell readers explicitly what were the treatment outcomes of the sub-group of initial defaulters in particular. If they were high, then it is recommended to state this, and to strike the whole ¶ that begins with “on a negative note” if study results were not negative.

In accordance with the referee’s suggestion, we highlighted the low initial defaulter rates in the abstract. We also changed the phrase of “on a negative note” (discussion/5th paragraph) as it might give a negative impression for the whole paragraph including the Cambodia’s ACF.

Because the initial defaulters in this ACF activity never started treatment, there is no treatment outcome available.

| 14 | Please also consider to include an estimate of the cost of detecting a case through the routine passive system if possible, calculated by the same method. Include this in the abstract. This would then justify the use of the term “cost effective” in the conclusion. Include a foot note to indicate that a rigorous measure of cost effectiveness of ECF vs. PCF would be calculated in DALYs. Such a task is currently impossible because it would have to take into account the likely impact on reduced transmission, which cannot be measured cross-sectionally. |
| 15 | Since this excellent work may wish to be replicated elsewhere, it would be good to have supplemental document describing the programmatic details of the intervention.

a. Were community leaders involved?

b. Would transport reimbursement have worked better than rounding up suspects?

c. Was it one CXR per patient? Which view(s) were taken?

d. Were CXR read on the spot and then again later?

| 16 | The authors have done a solid job reporting, but can do even better to “market” the positive results to program implementers by adding a paragraph discussing the feasibility of this exercise.

a. How many staff did it take?

b. How was the performance of the mobile CXR units?

c. Did it cause any problems for the lab or could the lab handle the extra work?

d. How did the community respond? What about stigma? |
e. When I first read mobile CXR, I assumed it was at the household, but reading it again, I see maybe it was at the facility? where was the mobile CXR parked?

We appreciate suggestions from the referee. Although we have been conducting the ACF with the approach described in the paper for many years, this is the first attempt to document and evaluate our activities in a peer-reviewed journal. We have just upgraded our ACF strategy with Xpert under a TB REACH Grant project. We will be glad to describe and further disseminate our approaches and operational details when we obtain more data in the current upgraded project.

### Minor issues not for publication

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 1 | The authors do not use ‘the Oxford comma’, which sometimes makes the sentences more complex to interpret than necessary. Suggest adding throughout when listing 3 or more items.  
   | Following the referee’s suggestion, we added the Oxford comma. |
| 2 | Intro ¶ 2- substitute the word “symbols” for the word “symptoms”.  
   | Thanks for the correction. We corrected. |
| 3 | The Intro is very standard, and has no journalistic “hook” to engage the reader and pull them into the text. This is a burning question in our field but the tone is clinical and could be a bit more enticing.  
   | Thanks for the feedback on this. We emphasised that the approach used in Cambodia was an innovative approach that promotes retrospective contact investigations combined with symptom screening (introduction/3rd paragraph). We hope this could attract readers’ attention. |
| 4 | Intro ¶ 4- substitute “could contribute to” for ’would present’- less presumptuous of a positive outcome.  
   | Thanks for the suggestion. We modified accordingly. |
Reviewer #2:

<table>
<thead>
<tr>
<th><strong>Major Compulsory Revisions</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methods</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
internal or external quality control? And if so, what were the outcomes of this? Did laboratory smear readers know if slides came from PCF or ACF suspects?

We clarified that external quality assurance systems were in place, just same as routine conditions (discussion/4th paragraph).

The reasoning in the last paragraph of the discussion on ACF having a sustained impact on smear-positivity should be worked out more. The authors are probably trying to say that the ACF activities might have increased awareness in the community (or at least among people in contact with TB cases) who might then be more likely to present to the clinic earlier if symptoms of TB develop. However, the current reasoning that the reduction of the overall patient pool in the community leads to lower smear grades seems not to make a lot of sense.

The implication of our finding, showing low smear grades observed in PCF-after, has two different directions in its logics.

First, as the referee mentioned, likelihood to increase community awareness and promote early health seeking among symptomatic suspects who were supposed to be the case of delayed diagnosis if ACF activities were not implemented. We stated this in the latter part of the paragraph.

Second, due to its nature, ACF is intended to diagnose a substantial proportion of existing TB cases in the community including those who do not recognize their symptoms. Asymptomatic TB cases would not have been detected under the current routine programme conditions as TB programmes normally define TB suspect by their symptoms. For this reason, we believe that ACF session will certainly reduce a pool of TB cases in the community and subsequently developed TB cases could have fewer bacterial load as indicated in the PCF-After group in our study. This type of phenomena can be called ‘harvesting effect’ in epidemiology. We tried to clarity these in the text accordingly.

**Minor Essential Revisions**

**Introduction**

Page 1 – 2nd paragraph: “....... rather than waiting for people to develop systems.....” should be symptoms

Thanks for the correction. We corrected.

**Discretionary Revisions**

**Methods**

What were the predefined criteria for the X-ray? What were considered abnormal results? More detail is needed here.

These are pointed out by the first referee, and we clarified accordingly (method/programmatic information).

Was ZN or FM smear microscopy done? I assume ZN but this should be specified.

Yes. ZN AFB microscopy was done as in the routine programme setting. This was clarified in the text.

The PCF group after is a lot smaller than the other two groups. How long after the ACF activities were completed were patients enrolled in the PCF-after group? In other words, is
this for a very short period immediately after the ACF activities, or for a longer duration? Did the patients in the PCF-after group come from all the health centres, or only a few of the 39 health centres included.

In this study, patients who had registered during 4 months after the ACF sessions (but until the end of each year) were reviewed for PCF-after cases. Since the ACF sessions were performed in the latter half of each year and they detected a considerable number of TB cases in each health centre, the number of patients detected after the sessions was quite limited therefore the PCF-after group became much smaller than the other two groups.

All health centres were included for our review without exceptions, but some did not report any cases.

It is mentioned that ethical clearance was not required. Perhaps the authors can state whether the data were anonymized, or some other information on how personal information was protected?

This is also suggested by the first referee and we explained a bit more detail (method/quantitative data/2nd paragraph).

It might be useful to provide a table with all the costing aspects that were considered in the cost-effectiveness analysis. From the abstract it is clear that mobile radiography units were used. Are the costs of running this mobile unit (e.g. a car or bus, driver, petrol) also include in the unit cost of $1.20 of an X-ray?

As you can assume from our description “Using the overall project cost of US$ 16 917…”, the cost calculation was made retrospectively without breakdown for each expense item. Hence we are unable to provide a table that contains much useful information for readers.

The costs of running the mobile unit (van, driver, petrol) were included in the operational cost, not in the unit cost of X-ray. We believe this is clear from our description “including the costs of logistic support and human resources”.

Was the difference in median age also significant? An inter-quartile range around the medians and a p-value for the non-parametric test of this comparison should be presented.

Thanks for very important suggestion. We have incorporated IQR and non-parametric test results in the text. Below is for your easy reference.

Median (IQR; inter-quartile range) for age are:
PCF before group: 48 (IQR 36-59)
ACF group: 55 (IQR 47-64)
A non-parametric test (Man-Whitney test) revealed the two groups were significantly different (p<0.001)

Is X-ray screening part of the diagnosis for PCF? Can the difference between the proportion of smear-positive and smear-negative patients between the two methods, be due to the screening and diagnostic method rather than being a true difference between the two patient groups? The X-ray screening used in the ACF group might detect a lot of TB suspects with X-ray abnormalities, of whom then few are confirmed by smear, while in the PCF group an X-ray might not be part of the standard diagnostic process, and diagnosis is predominantly based on smear microscopy. This could also be discussed.
This is also mentioned by the first referee. We clarified that X-ray was used also for PCF groups, but required patient referral to higher level facilities, according to the national guidelines for smear-negative diagnosis. (method/programmatic information).

Yes, the different screening and diagnostic methods might be one of the reasons to explain the difference in the proportion of smear-positive and smear-negative patients between the two methods. This can be assumed from our statement of over-diagnosis of smear-negative TB in the study limitation, but we further clarified this.

Can a p-value be added for the comparison of 56.9% lower smear grades in the ACF group compared to 42.7% in the PCF group?

After grouping scanty and 1+ into lower smear group, and 2+ and 3+ into higher smear group, the calculated p-value was 0.010, which was added in the text.

Is transfer-out really an unfavourable outcome? Perhaps this should be excluded from the unfavourable outcomes, since it is not really a bad outcome, compared to default, dead, and failure.

Definition of transfer-out is patients transferred to another treatment site and the original facility was not able to trace the final treatment outcome. Because health facilities were unable to keep track of the patients’ treatment progress and its outcomes, ‘transfer-out’ is commonly categorised as an unfavourable outcome.

P-values should be added to comparisons presented in the last paragraph of the results, where the ACF cases are compared with the PCF-after cases.

We have reported p-values in some of the comparisons in the text. We also conducted non-parametric test for age distribution of PCF-After group.

Discussion

What is meant by a delay-study? Perhaps the purpose and method of such study should be described, because every reader might have a different understanding of what such a study might involve.

Delay study is a study that quantifies patient delay and health system delay by examining date of onset of symptoms, date of treatment seeking, date of diagnosis, and date of treatment initiation. We briefly explained this in a bracket (discussion/4th paragraph).

Why do the authors conclude that further investigations are needed to clarify the benefits of ACF in early case finding and the costs associated, if this is what they themselves tried to show in this article? I thought that they concluded that the extra costs are worth it?

This is also stated by the first referee and we modified the text accordingly.