Author’s response to reviews

**Title:** Identification of losses to follow-up in a community-based antiretroviral therapy clinic in South Africa using a computerized pharmacy tracking system

**Authors:**

Mweete D Nglazi (Mweete.Nglazi@hiv-research.org.za)
Richard Kaplan (richard.kaplan@hiv-research.org.za)
Robin Wood (robin.wood@hiv-research.org.za)
Linda-Gail Bekker (linda-gail.bekker@hiv-research.org.za)
Stephen D Lawn (st Evelawn@yahoo.co.uk)

**Version:** 4  **Date:** 30 July 2010

**Author’s response to reviews:** see over
Dear Editor,

Thank you for your email of the 10th June 2010 and for forwarding the reviewers’ comments. We are pleased that our paper was found to have broad relevance to ART programmes in resource-limited settings and that the iDART computerized pharmacy tracking system was regarded as being potentially useful in identifying patients who had failed to attend ART clinics and who require tracing in the community.

We thank the reviewers for the useful comments that we have used to revise and strengthen the manuscript. We copy the comments below in italics and respond.

Reviewer 1:
We thank the reviewer for the useful and relevant comments and we respond below:-

This is an interesting operational research study looking at an important issue. I think the question has broad relevance to ART scale up and the iDART program might be useful in tracking patients.

Major compulsory points:

1. One possible issue with the study design and patients is that the relationship between pharmacy delay and LTFU status may differ according to time on ART. Given that many who are lost to follow-up are later found to be dead, and given the early timing of most first-year deaths after ART initiation in sub-Saharan Africa, being late by a few days may be more predictive of true LTFU when this occurs in the first month of ART compared to later months. Furthermore, I suspect that to the degree that LTFU represents deaths patients with low CD4 counts may have a different relationship between iDART delays and true LTFU status. So, the test characteristics of definitions for LTFU plausibly vary by time on ART, baseline CD4...
counts, and perhaps other characteristics not measured here (like employment). If this is true, then the point that simple monitoring strategies may have important subtleties in accuracy is an important point. This would suggest that automated strategies for flagging loss to follow-up, etc, need to be more complex. Perhaps the authors can either look at time on ART and baseline CD4 count as a possible modifiers of the test characteristics of delay for LTFU. If not, I think these limitations should be discussed.

We agree that multiple factors may be associated with LTFU and that in the long-term algorithms to identify patients who are potentially LTFU may be refined with a greater level of sophistication. However, our initial approach has been to develop a very pragmatic system that can be used to rapidly generate a list of patients that are potentially LTFU based simply on the time delay since their last clinic attendance. This initial study represents important proof of principle and explores one key variable (i.e. iDART pharmacy delays) which is likely to be of overarching importance. This system is now being used within the clinic on a routine basis and it works well. This represents a major step forward for our clinic and others in South Africa and the wider region that have started to use this system. The results of this initial simple cross-sectional survey have been used to devise a long-term prospective study in which the complexities of predicting LTFU can be further refined. This may enable more sophisticated algorithms to be developed in due course. This limitation has been included in the discussion on page 9 paragraph 3.

2. What was the rationale for starting the definitions at 6 weeks and not earlier? Delays clearly related to outcomes. If one aims to prevent rather than identify bad outcomes should not an earlier delay be examined?

The test definitions of delay were based on time since last attendance at the clinic and not the time since medication would have run out. Therefore, a six weeks delay effectively represents just a 2 week period without treatment. In our study, a delay of six weeks (2 weeks without medication) identified 22% (n=560) of the whole cohort which represents a colossal workload of patients that would potentially require tracing. This has not been found to be practical and shorter delays would have increased this number even further. We have included the rationale for starting the definition at 6 weeks in the discussion on paragraph 2 on page 8.

3. I would like more detail on the goals of this approach. Is the goal of this system to identify patients to trace who may be nearly lost but still recoverable? Or is it more to monitor program outcomes? This is addressed to some extent but I think this could be much clearer.

The goal of this approach is to identify the best performing definition of delay that can be used to rapidly generate a list of patients who need to be traced in the community to facilitate their return to care. We have discussed this in the introduction in paragraph 3 on page 3 and in the discussion on paragraph 1 on page 8. We have further clarified this on paragraph 1 on page 8.

4. It would seem to me that there could be a survivor bias in this cross sectional design. The patients who make it into the cross sectional design exclude many who have suffered early adverse outcomes. The rationale for use of the cross sectional design vs the cohort design would be helpful to know. Can the authors address
possible survivor bias that analytically using data at hand (eg through sensitivity analyses)? If not, this is a possible limitation that should be addressed in the discussion. This relates to my first point, above.

As discussed under point 1 above, this was an initial exploratory cross-sectional study as a key first step to developing iDART as a practical tool to generate a list of patients who were potentially LTFU, thereby triggering patient recall mechanisms. These data have been used to establish a prospective study that further explores how the function of this tool may change longitudinally with time. This point has been included in the revised discussion.

5. The optimal delay to use depends on the resources available and planned to be expended for tracing patients, and the extent to which a program feels it can attempt to prevent deaths among those who have pharmacy delays. So, while I appreciate the author's personal perspective that one threshold is "too nonspecific", I think the manuscript should address the fact that choosing between cut-points depends on many factors and there is unlikely to be a universally ideal threshold that fits for all clinics. In a setting where many patients have cell phones and a few minimally trained can use a land line to call patients all day, then setting a threshold earlier than was examined here (eg. 4 weeks), which might capture 100% of those bound for future LTFU, might be feasible.

We found that use of the 6 weeks delay identified 22% (n=560) of the entire cohort at that time-point and only 15% (n=85) of these were truly LTFU. The associated work-load is not a practical proposition in our clinic and we believe this is likely to be the case elsewhere where resources for tracing patients are greatly over-stretched. However, we agree entirely with the reviewer that what is optimal for one clinic may not be optimal for another as available resources may differ greatly. We have discussed this on page 10 paragraph 1.

Minor point:
1. I think the discussion could be considerably shorter as it seems repetitive of the introduction.
   We have tightened up the discussion to exclude the redundant material although we have also included other important points raised by the reviewers.

2. Jeff Stringer's group has a paper recently released from the American Journal of Public Health looking at this issue that the authors might want to evaluate in light of their findings.

   Thank you for highlighting this very useful reference. It is interesting to note that that the authors of this paper from a study in Zambia found that the best-performing LTFU definition was 56 days from the time medication ran out [1]. This is entirely consistent with our findings of a delay of >=12 weeks from the time since last attendance at the clinic (=54 days on average since medication ran out). This supports our findings and their generalisability to other clinics in the region. We have included this in the discussion on page 8 paragraph 3.

Reference:
Reviewer 2:
We thank the reviewer for the useful and relevant comments. We are pleased that this was found to be an interesting analysis and that the study objective was clear. We respond to the specific comments below:

The authors present an interesting analysis aimed at investigating the ability of an open-source electronic pharmacy system (the intelligent dispensing of ART – iDART) to identify patients who are likely to drop out or die in a community-based ART cohort in Cape Town, South Africa. Although the objective is clear and the analysis is potentially useful there are aspects of the nature of the data collection and of the analysis design which are unclear and which limit the interpretability of the results.

**Major compulsory revisions**

1. 583 patients who started ART after enrolment in the cohort and who did not take part of the survey at April 2008 were excluded from this analysis. Were iDART data available for these patients? If so, it is unclear why these were excluded as it seems crucial to know whether the delay in pick up of pharmacy prescription could predict the risk of these patients dropping out by April 1 2008. Indeed, I would focus the analysis on all 3,384 patients who started ART, if possible, to avoid selection bias.
   
   As stated in the original manuscript submitted on page 4 paragraph 1, the iDART system was introduced 5 years after the start of the clinic in September 2002. Therefore, longitudinal data were not available for the period suggested.
   
   As explained in our response to Reviewer 1 point 1, we agree that the cross-sectional survey design is unable to assess how the duration of iDART delay cut-offs changes over time with maturation of the cohort. However, our initial approach has been to develop a very pragmatic system that can be used to rapidly generate a list of patients that are potentially LTFU based simply on the time delay since their last clinic attendance. This initial study represents important proof of principle and explores one key variable (i.e. iDART delays) which is likely to be of overarching importance. The results of this initial study have been used to devise a long-term prospective study in which the complexities of predicting LTFU using iDART can be further refined. This limitation is included in the discussion on page 9 paragraph 3.

2. The definition of patients who failed to pick up ART based on iDART is unclear. Was this based on the time difference between the date of the survey and the time at which medication was last dispensed? If so, could the high sensitivity of the \( \geq 6 \) weeks and \( \geq 12 \) weeks definitions simply reflect the fact that it was too early for the 85 patients to come back to the clinic as they had sufficient drug supply? I suggest to define potential LTFU on the basis of the frequency of medication pick-up over, say, the first year of ART and to exclude patients who started ART within 12 weeks of the date of survey.

   This cross-sectional study utilized a pragmatic tool at a single time point to generate lists of patients with different durations of delays in pharmacy pick-ups. The time periods used represent the delay since medication was last dispensed as stated in the Methods paragraph on Study Design on page 4. We have defined true LTFU as confirmed failure to pick up medication for 3 months or more consistent with Rosen et al. and Chi et al. [1,2]. The high
sensitivity for $\geq 6$ weeks delay of course arises from the fact that a 3 months delay was the gold standard definition.

References:


3. For the given prevalence of LTFU, one key parameter here seems to be the positive predictive value (how many of those defined as failures on the basis of iDART were truly LTFU). These percentages seem to be reported on the first line of Table 1 (15%, 44%, 45% and 50%) although not defined as such. I suggest to remove Table 1 (and Figure 3 which seems to show the same results) and present a 2x2 table with the raw data, sensitivity, specificity, PPV and NPV for all four definitions of “predicted LTFU” used. These could be done separately for the April 2008 and April 2009 surveys. People who were dead as of the date of the survey should be removed from the analysis related to the survey in question. Why were people who had transferred to a clinic outside of the cohort-network not defined as LTFU? It seems to me that they should be, as long as their last clinical visit at the site in which they were originally enrolled was $>3$ months prior to the date of the survey? The only difference with the other patients defined as LTFU is that the reason for the drop-out is known.

We agree that the specificity, sensitivity, PPV and NPV or all four definitions of “predicted LTFU” used should be calculated and we have presented these in a new table (Table 2). This makes the analysis much clearer. We have deleted Figure 2 as these data are now included in the new Table 2. We have also removed figure 3 as suggested by the reviewer as this is indeed a reiteration of the data in Table 1.

In a public health programme patients who are LTFU have completely different implications to those who have transferred-out to another ART clinic. Those transferred-out to another clinic have been retained within the provincial/national health system and do not require tracing by community-based workers. In contrast, those LTFU do require tracing. Thus, it is extremely important to distinguish between these two very different outcomes.

4. It is unclear how many of the patients included in the survey of April 2008 could be assessed a year later (April 2009?) and how did the authors handle patients who were not included in the second survey? This comment applies also to the analysis with death as an endpoint. Also, how many patients of those who were defined as “truly LTFU” by April 2008 came back to the clinic between April 2008 and April 2009? I am not sure if it makes sense to call these as LTFU.

With regard to the data from 2009, these do not represent a further cross-sectional survey but rather show the outcomes one year later of the patients ($n=2548$) included in the cross-sectional survey in April 2008 as clearly stated in the methods and results sections. The point of this analysis is to provide some indicator of the longer-term outcomes of patients with and without
various lengths of pharmacy delays. The point being made is that detection of iDART delays is of prognostic value. Twenty-two (25.9%) patients who were defined as true LTFU subsequently returned to care and this has been included in the revised results section on page 7 paragraph 3. These patients were defined as true LTFU as they fulfilled the case definition despite later returning to care. This is a phenomenon we have reported in another study but further consideration of this does not fall within the aims and objectives of this study.

5. The methods used to ascertain deaths have not been explicitly stated. Given that this is likely to be a major reason for patients not coming back for a visit “notification from any source” seems too vague. Deaths refers to all-cause mortality notified from therapeutic counselors after home visits, relatives/family members and hospitals. We have previously reported on studies on mortality within this cohort and during these it was found that ascertainment of deaths by these means was good [1,2]. This point has been clarified on page 4 paragraph 3.

References:

6. There is not attempt to perform multivariable analyses. Were any of the other parameters collected in these patients, say, at the time of ART initiation predictive of LTFU? Could any of this (e.g. enrolment in the cohort when ART-naïve, gender, etc.) confound the association between iDART and risk of LTFU or death? Please see responses to point 1 Reviewer 1 and to point 1 above.

Minor essential revisions
1. The adopted definition of true LTFU needs to be stated in the abstract
   We have included this in the abstract on page 2.

2. Page 6 of Results. These patients had been receiving ART for a median of 1.9 years (IQR, 1.0-2.9). Does this sentence refer only to the 24% of patients who were enrolled when ART-experienced?
   This refers to all 2548 patients within this treatment cohort at the time of the cross-sectional survey. We have amended this sentence.

3. Page 6 of Results. In total, there were 85 patients who were true LTFU, representing 3.3% of the overall cohort. The real percent of patients LTFU in the cohort by April 2008 seems much higher: (85+334+249)/3384 =20% and more consistent with the rates observed in the resource limited settings.
As discussed in detail in the response to major point 3, true LTFU and transfers-out are completely different outcomes. This differentiation is important. So the actual LTFU rate is \((85+334)/3384=12.4\%\) over 5 years. The proportion of patients LTFU in this study is at a single cross-sectional time point.

4. Page 7 of Results. The proportions who were true LTFU ranged from 44-50\% and did not differ statistically between groups \((p=0.282)\). Test used is unclear. It is incorrect to test the difference between PPV of the 4 definitions using a standard chi-square because groups are not mutually exclusive. Authors should consider calculating areas under the ROC-curves and compare them using an appropriate t-test instead.

We thank the reviewer for raising this point. Thus, we have omitted this statistical comparison as it is essentially redundant and omission does not alter the findings and conclusions of the study.

5. Page 7 of Results. Similarly, higher proportions of those with pharmacy delays had died after one year (3.8\%, 7.7\%, 10.3\% and 8.8\% versus 0.02\%, respectively; \(P<0.02\) for comparisons of #12 weeks group and #18 weeks groups with no delay group). Suggest that raw data used to calculate these proportion are shown – see major point #3 above.

As discussed in the response to major point 4 above, the data from 2009 do not represent a further cross-sectional survey but rather show outcomes one year later of the patients \((n=2548)\) included in the cross-sectional survey. The proportions can be compared, in this instance, because the groups are mutually exclusive (patients with no iDART pharmacy delays versus those with pharmacy delays). Thus, we retained this sentence.

6. Page 7 of Discussion. Within each of these groups the proportions of patients who were true LTFU was similar (approximately one half). However, the #18 and #24 weeks cut-offs had substantially lower sensitivities for true LTFU (62\% and 47\%, respectively) compared to the #12. On this basis, we identified the #12 weeks delay as the optimal cut-off. Optimal cut-off is generally chosen on the basis of the trade-off between sensitivity and specificity and not sensitivity and PPV.

We have now presented the sensitivity, specificity, PPV and NPV for each of the iDART delay cut-offs in the new Table 2 in the revised manuscript. This now makes the data much clearer and it is easier to see how the optimum cut-off was derived using sensitivity and specificity. (Note: we did not calculate PPV in the original version of the paper).

7. Page 7 of Discussion. We conducted the cross-sectional study at a single time-point and we cannot be sure that results would be the same at other time-points. Time-points seem to be two, April 2008 and one year later?

Please see response to Reviewer 1 major point 4. As stated in the methods section, the cross-sectional survey was conducted at a single time point (April 2008) and was used to identify patients who missed scheduled pharmacy appointments using the iDART system at different durations of delay. Then we looked ahead a year later (April 2009) to determine the subsequent status of all patients identified at the time of the cross-sectional survey.

8. Table 1. There is no footnote referring to the “a” next to 95% CI – suggest to modify the table as described in major point #2.
We have removed the footnote and amended Table 1.

9. Suggest to replace Figures 1, 2 with ROC-curves and AUC, remove Figure 3 and replace Figure 4 with a cross-tabulation of survey results at April 2008 vs. April 2009.

Please see response to point 3 above regarding changes to the figures and tables in the revised manuscript. We would prefer to retain figure 4 because this is not a second cross-sectional survey (see response to point 7 above) but instead is a very clear way of illustrating for the readership the one year outcomes with different durations of delays ie their prognostic value.

Editorial request

We recommend that you copy-edit the paper to improve the style of written English. If this is not possible you need to use a professional copyediting service. Examples are those provided by the Manuscript Presentation Service (www.biomedes.co.uk), International Science Editing (http://www.internationalscienceediting.com/) and English Manager Science Editing (http://www.sciencemanager.com/). BioMed Central has no first-hand experience of these companies and can take no responsibility for the quality of their service.

Please also ensure that your revised manuscript conforms to the journal style (http://www.biomedcentral.com/info/ifora/medicine_journals). It is important that your files are correctly formatted.

We thank the Reviewers and Biomed Central Editorial Team for helping us to revise and strengthen this manuscript. We hope that this revised version is now acceptable for publication in BMC Infectious Diseases. We look forward to your feedback in due course.

Yours sincerely,

Mweete D. Nglazi, on behalf of the co-authors