PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Introduction and Evaluation of a 'pre ART care' service in Swaziland: An Operational Research Study
AUTHORS	David Burtle, William Welfare, Susan Elden, Canaan Mamvura, Joris Vandelanotte, Emily Petherick, John Walley and John Wright

VERSION 1 - REVIEW

REVIEWER	Kathryn Church Research Fellow Population Studies Department London School of Hygiene & Tropical Medicine UK
REVIEW RETURNED	13/06/2011

THE STUDY	While it is encouraging to read this pragmatic evaluation of an important intervention using routine service data, I have raised some concerns below about the study design and the report. I feel many could be addressed through revisions to the paper. The results presented are interesting and important. Research question(s): While the aim of the intervention is described, the research aims are not clearly defined. Outcomes described at the end of the 'Background' section could be reformulated as research aims (since outcomes are repeated in Methods anyway). Study design: - It is unclear why the authors chose to use a sample of pre-ART clients as their study population, instead of VCT clients. If one of the main aims of pre-ART is to increase timely access to ART for those testing HIV positive, VCT clients should have really formed the study population, in order to gain a more valid estimate of programme effectiveness. This could be highlighted in the discussion as a usual piece of further research to demonstrate the effectiveness of pre ART care. -There is also no rationale described for new testers being excluded; shouldn't these form an important target group of a pre-ART intervention? Also, pregnant women and those with TB co-morbidities are also likely to get lost in the system, and another important target group? - I'm not sure the 'baseline' presented forms an adequate 'pre-test' population since it seems that these clients received the intervention as well (or if not, further clarity is needed). The causal mechanism by which the differences between groups were found is therefore difficult to understand, since everybody seemed to receive the 'pre-ART package'.

	- If there were significant differences in the 'time since HIV test' between baseline and follow-up, these should have at least been noted in the background characteristics. - It is slightly strange that the second group "Cohort 1" was recruited/sampled immediately after the first cohort; nevertheless, there do seem to be important trend markers in the results. - Lastly, it is slightly misleading to term the groups 'cohorts', which implies the same population was followed over time; I think I understand that ART initation was documented (after a certain number of months?) for each group, hence why you called them cohorts (i.e. you had to follow up to measure the outcomes for each group? (see comments on methods below)); however, it is not a cohort study, and thus a more appropriate term should be used (e.g. Group 1, Group 2, Group 3). - As noted by the authors, the study design also had critical limitations, in particular the introduction of a lower ART initiation threshold during the follow-up period, which could had significantly biased one of the key outcomes reported. Participants: The description of how the 2nd and 3rd groups (cohort 1 and 2) were selected is not sufficient; is the 3rd cohort the first 200 clients enrolling in Feb 2010? This should be noted. It is unclear why the 2nd cohort was recruited for so much longer, and thus is so much larger. Methods: -The description of the intervention has been placed within the methods, but I feel this should be a separate section, or part of the background. - Also, information is missing about the follow-up period and methodology for the 'cohorts'. Did all the clients sampled initiate treatment? Or when was the study conducted/what is the cut off point? Is the sample size of the 'mean time to initiation' outcome the same as the total sample size for each cohort, or does it only include those who did initiate? - The methodology should also make clear the source of the data used (patient registers? Or records?), how clients were followed up? - The Priority interventions outcome is not clearly defined; what are these interventions? Abstract/summary/key messages/limitations: The abstract seems OK. The key limitations could be elaborated (i.e. no VCT client sample to estimate full effect). A few specific comments: - 'Pre ART care' is "constituted" of a package of interventions (not "provide"). - I'm not sure you have measured impact on quality; see my point in the discussion. Continuity of care and service uptake may be more appropriate outcomes. - Throughout the paper many hyphens seem to be missing e.g. evidence-based; follow-up; HIV-related mortality; task-shifting, etc.

	Statistical methods: - Is the data skewed? Especially the mean time to ART initiation? If so, it would be important to also report the median values. Also, if the data is skewed (and I imagine it is), then you might need to consider non-parametric tests instead of ANOVA (e.g. Mann-Whitney U test). - I think it would be useful to include the SD along with the means. - P values should be reported in the table, as well as the F statistic for ANOVA. - You might want to consider doing tests for trend instead of basic tests; for the categorical data, you could do Chi2 test for trend; for the continuous data, you could do a regression, and produce a p value for the coefficient. (though do consider the skewness first). - Instead of the ANOVA, you could also consider giving a separate p value for group 2 vs group 1; and group 3 vs group 1 (e.g. Z test; or Mann Whitney; or do a post-estimation test on the ANOVA, e.g. Tukey Kramer, which gives tests for differences between the individual groups. - See further comments in next section on table. References: There does not seem to be any discussion of other studies on pre-ART care, see point below in next section. Background: A few additional comments on the background section: - As noted earlier, there is a need to discuss existing studies on the effectiveness of Pre-ART interventions as important scientific background to the study (at least briefly) - The references for components of pre ART services are WHO guidelines; are there any stipulated or agreed packages within Swaziland? - You state on line 41 that HIV care was episodic and unmonitored; did you mean Pre-ART care (I'm not sure that the statement applies to ART services?) -HTC needs to be defined in the text since it appears later undefined (line 46). - I think CD4 levels not usually reported as copies/ml? (please check) Regarding the intervention description (figures on pages 16 and 17, not labelled in my PDF copy): - What is the difference between regular review and clinical staging? Does a regular review not include clinical staging? - Does Pre ART care not include any other routine diagnostic tests? - What about condom provision on counselling on sexual behaviour/sexual health? Does this not form an important aspect of Pre ART care? Or advice on nutrition, etc? - The second diagram is missing 'up' in the last box.
RESULTS & CONCLUSIONS	Results presentation: -Regarding Table 2-A, I find it hard to understand who the 'N' is: the results present a 'proportion being assessed for ART' – is this a proportion of those testing? Or is it a proportion of those given a Pre-ART file? If they are given a pre-ART file, why are so few assessed for ART at Round 1? - The sample size for each group and each indicator should be noted in the table, since I assume it is different. - The tense in the tables seems incorrect, should be past tense e.g. "proportion assessed for ART"? -The table should contain the p values, rather than just reporting them in the text. You could use "*" system to denote the significance

	level of each round compared to baseline. - Regarding the bar chart (no figure no.), the sample size for each group should be included on the graph. Also, this data is not presented as a table, and I think it would help to see the proportions, and certainly it would help to see a statistical test on each outcome reported. At a minimum, you could include "*" above each bar to show statistical difference to Group 1. - Also, how is "HIV clinical staging" different to "Assessment of eligibility for ART"; if the client receives clinical staging, does that not mean that they are assessed for ART initiation? Interpretation/conclusion: The discussion section could be reorganised and refocused: - There is no discussion of the results in light of other studies on pre-ART packages and their impacts. Overall the balance between the study discussion and the 'lessons learned' seems uneven; there is too much on lessons learned, and not enough about the research presented. - It is not really appropriate to refer to results table in discussion. The first paragraph also just seems to repeat the previous section in the results and could be written in a more reflective style. - You discuss a 'reduction in waiting times' due to the intervention, but the time lag to ART initiation is not a function of waiting times (which implies service waiting lists). - You discuss cotrim results not improving, but was it not statistically improved between round 1 and 3? (no p values presented, as mentioned earlier) Why were the cotrim results still relatively low (<60%) after Round 3? - The point about CD4 thresholds changing is critical, and should be discussed next to your point about causation, i.e. earlier in the discussion. The discussion about the causation could also include a reflection on any other policies or programmes were ongoing in Swaziland at that time, which could have contributed to the observed changes (I imagine not many, which would then give greater weight to the causation argument), but it is important to note it. The fact that you saw important improvements in key outcomes despite not including recently tested patients should also be highlighted, since this shows that the intervention is still powerful, even among those who were already at more advance disease stages. - You call the study a before/after evaluation, which it doesn't necessarily seem to be since Round 1 received the intervention (unless I have got it wrong and you can explain it another way?) - I wouldn't have worried about the weight of the different intervention elements; you clearly stated that it was a package of care composed of several components, or clearly related to each other, and I don't find this undermines the outcomes reported in any way. - The term 'early service cohorts' on line 54 p10 is confusing to the reader; it would also be important to know what proportion of entries had missing data. - I think it would be useful to come up with a hypothesis about why there were more clients with advanced disease in the later sample, as otherwise it certainly undermines the study findings. As noted earlier, this may influence the key outcomes, or is even an outcome itself? - You discuss implementing a pathway across 'multiple programmes' but the description of the intervention and Figure 2 really cover only HIV care and treatment. If greater coordination was achieved with other health programmes, this could be explained more in the text.

	- The discussion of vertical programmes could be made clearer. As could the last para on p.11, the last sentence of which becomes hard to follow. - I find Figure 4 not particularly useful; it is unclear if the 'actual impact' is what you implemented in your intervention, or whether this is what you wish you could achieve? - I was not able to review Figure 5, it was not included (did you mean 4B?). - You discuss task-shifting, and also mention it as part of the intervention, but don't provide any assessment of that component of the study. Your conclusion that "task shifting can address critical bottle necks in paritnet pathway" then seems overstated; as a reader we are unable to assess what the task-shifting actually achieved compared to other study components. - I find the final conclusion that you improved "quality of care" problematic; I know the HIV community has a tendency to measure quality in terms of outcomes (e.g. good adherence must mean good quality), but many other fields of medicine prefer to use process measures to examine quality of care, more focused on adherence to clinical algorithms, or on inter-personal aspects of care. I think you can say that you improved continuity of care, as one aspect of quality of care. Or at least perhaps acknowledge that you only tackled one aspect of quality?
GENERAL COMMENTS	One small comment regarding ethics: I would have questioned the ethics of the intervention component which led to medical staff visiting people's houses on motorbikes; this could seriously breech client confidentiality in the Swazi context where unusual visitors to homesteads could lead to involuntary disclosure of HIV status.

REVIEWER	Bruce Larson Associate Professor Department of International Health Boston University School of Public Health No competing interests
REVIEW RETURNED	28/06/2011

GENERAL COMMENTS	Referee Report: Improving the quality and timeliness of HIV care in Swaziland: Evaluation of a "pre-ART care" service General This study identifies an important problem, loss to follow up in pre-ART care after HIV testing, and evaluates an intervention package or strategy to improve linkage to care and treatment after HIV testing. Using data for three patient cohorts and a pre/post study design, study results suggest that the intervention increased the proportion of HIV-infected patients who were 'assessed for ART', the proportion of ART-eligible patients who initiated ART (they also started ART more quickly). This is a useful study with important implications for supporting improvements in HIV/AIDS programs. The paper as currently drafted would benefit from some revisions to clarify the patient populations, the statistical analysis, and the intervention package. A list of comments and suggests is provided below. These

comments are roughly in the order they appear in the paper.

1. In the pre-ART care discussion on page one, the point needs to be made that the “assessment” time is short because: (1) they waited too long before HIV testing; and (2) they tested in the past, did nothing, waited, re-tested again, and so on.

2. On page 1 (really figure 1), the first line of Figure 1 needs to be edited (the CD4 eligibility criteria is off.

3. On page 1, you mentioned that “Patients with unknown status were tested in the HIV testing and counseling centre”. It would help to be more specific here: what portion were patients already in the ward because they were sick (are these hospital patients) and what portion were walk-ins for counseling and testing (basically the four groups in Figure 2). The current analysis pools data for four patients groups, which I think are rather different (at least the walk-in VCT compared to ward patients) into one analysis. The analysis would be stronger if the analysis was stratified by patient group.

Also, if they were found to be HIV positive, it seems more correct to say that a blood sample from taken for CD4 testing (instead of a CD4 count was taken).

4. The “task shifting” discussion is not clear to me and could be revised to clarify what tasks where to whom.

5. On page 7 in the “sample” section, more information on patients is needed.

It seems you have four populations that have different situations that are relevant for evaluating the results. For the three study group, it would be good to know the proportion in each category identified in Figure 2.

After the cohort was defined (e.g. baseline in Feb/March 2009), how long after this period was the follow up period (e.g. 6 months after enrollment)?

Was this follow up period the same for all study groups?
This analysis would be much cleaner if the sample was only VCT patients newly diagnosed with HIV. Mixing up the four group into one clouds the interpretation of results, especially since the populations seems to be different across time (e.g. the proportion with CD4 < 200 increases substantially between baseline and cohort 2).

Many studies are reporting improvements in the rates of ART eligible patients who initiate ART (and in shorter times). Dealing with the patients whose CD4 counts are high seem to be the more complicated group to try to retain in regular pre-ART care perhaps for two years before they become eligible. Thus, it might be better to report results for patients ART eligible and those who are not (e.g. in Table 2-A, no information is currently provided for patients who are not ART eligible).

6. Table 1: The proportions of patients in each group (from Figure 2 for example) would help to understand if the cohorts were similar over time.

Also, did all patients sampled provide blood for a CD4 test? In our

	reviews elsewhere, even through all patients were offered the opportunity to provide a blood sample for CD4 testing, some number left after HIV testing without providing a sample. 7. Table 2-A: Please explain somewhere what “assessed for ART means” (somewhere earlier eligibility should also be defined). I assume CD4 count results and/or exam for staging, but it would help readers on familiar with Swaziland. 8. I’d suggest the authors think carefully about their statistical analysis. I’m not sure ANOVA is adequate for what you are doing because of the underlying normality assumption and the implications of outliers. It would be reasonable to me to compare baseline to cohort 1 and baseline to cohort 2. I’d at least like to see an absolute risk or relevant risk analysis that included patient categories as covariates. I’d rather see the analysis stratified by patient group (the four in Figure 2) to see if patients eligible for ART are driving the results. 9. Figure 3 does not provide adequate information. The table(s) with the actual results and numbers would be better. Again, it would be better to report such information separately for the ART eligible and those not eligible. 10. Because the intervention was a package of interventions, it is unclear if the whole package is important or if certain components drove the results. For example, having information on the % of patients who could not be contacted by phone for each group and knowing how many adherence officer visits were conducted for each group would help have perspective on the relative importance of different components of the intervention. While I understand it is not possible to attribute the impacts of individual components, having information on amount of activities in each component (as well as the cost of the component – e.g. the active follow up) provides useful perspective. 11. In the strengths and limitation section, you mention ‘gaps in data’ but no where in the analysis did you mention the issue. Where there gaps in data used in the analysis not reported? 12. The issue of known versus new diagnosis makes sense. While not a solution to this issue, it would make sense to separate the analysis into patient groups that are testing through different situations: walk-ins for VCT, provider initiated in OPD, and provide initiated in the Ward. Patients in the ward because they are sick with HIV/AIDS related conditions should already be considered “failures” for evaluation of pre-ART care and excluded from the analysis. See attached file

VERSION 1 – AUTHOR RESPONSE

Reviewers Comments NOTES

Research question(s): While the aim of the intervention is described, the research aims are not clearly defined. Outcomes described at the end of the 'Background' section could be reformulated as research aims (since outcomes are repeated in Methods anyway). DONE

- It is unclear why the authors chose to use a sample of pre-ART clients as their study population, instead of VCT clients. If one of the main aims of pre-ART is to increase timely access to ART for those testing HIV positive, VCT clients should have really formed the study population, in order to

gain a more valid estimate of programme effectiveness. This could be highlighted in the discussion as a usual piece of further research to demonstrate the effectiveness of pre ART care.
UNFORTUNATELY WE WERE UNABLE TO LINK TESTING DATA TO PRE ART DATA. WE HAVE MADE REFERENCE TO THIS IN THE DISCUSSION.
-There is also no rationale described for new testers being excluded; shouldn't these form an important target group of a pre-ART intervention? Also, pregnant women and those with TB co-morbidities are also likely to get lost in the system, and another important target group? WE RECOGNISE THAT THE INCLUSION OF ALL THOSE WHO HAVE AN HIV DIAGNOSIS WOULD HAVE BEEN PREFERABLE. THE EXCLUSION OF NEW TESTERS AND PREGNANT WOMEN AND THOSE WITH TB IS IS DUE TO THE ARRANGEMENTS FOR PRE ART WITHIN THE HOSPITAL.
- I'm not sure the 'baseline' presented forms an adequate 'pre-test' population since it seems that these clients received the intervention as well (or if not, further clarity is needed). The causal mechanism by which the differences between groups were found is therefore difficult to understand, since everybody seemed to receive the 'pre-ART package'. THE BASELINE IS NOT AN IDEAL PRE TEST POPULATION, BUT WE DID NOT HAVE SUCH A POPULATION AVAILABLE AS SYSTEMIC DATA COLLECTION WAS PART OF THE INTERVENTION. WE ARE UNCLEAR WHY THE GROUPS WERE DIFFERENT.
- If there were significant differences in the 'time since HIV test' between baseline and follow-up, these should have at least been noted in the background characteristics. WE DO NOT HAVE GOOD DATA ON THIS.
- It is slightly strange that the second group "Cohort 1" was recruited/sampled immediately after the first cohort; nevertheless, there do seem to be important trend markers in the results. THIS WAS A PRAGMATIC DECISION BASED ON DATA THAT HAD PREVIOUSLY BEEN ENTERED FOR A PILOT EVALUATION.
- Lastly, it is slightly misleading to term the groups 'cohorts', which implies the same population was followed over time; I think I understand that ART initation was documented (after a certain number of months?) for each group, hence why you called them cohorts (i.e. you had to follow up to measure the outcomes for each group? (see comments on methods below)); however, it is not a cohort study, and thus a more appropriate term should be used (e.g. Group 1, Group 2, Group 3). CORRECTED.
- As noted by the authors, the study design also had critical limitations, in particular the introduction of a lower ART initiation threshold during the follow-up period, which could had significantly biased one of the key outcomes reported. AGREED
Participants: The description of how the 2nd and 3rd groups (cohort 1 and 2) were selected is not sufficient; is the 3rd cohort the first 200 clients enrolling in Feb 2010? This should be noted. It is unclear why the 2nd cohort was recruited for so much longer, and thus is so much larger. PRAGMATIC DECISION.
Methods:
-The description of the intervention has been placed within the methods, but I feel this should be a separate section, or part of the background. ALTERED
- Also, information is missing about the follow-up period and methodology for the 'cohorts'. Did all the clients sampled initiate treatment? Or when was the study conducted/what is the cut off point? Is the sample size of the 'mean time to initiation' outcome the same as the total sample size for each cohort, or does it only include those who did initiate? NOT ALL INITIATED TREATMENT AS THEY DID NOT MEET ART CRITERIA
- The methodology should also make clear the source of the data used (patient registers? Or records?), how clients were followed up? PATIENT REGISTERS, ADDED.
- The Priority interventions outcome is not clearly defined; what are these interventions? ADDED.
Abstract/summary/key messages/limitations: The abstract seems OK. The key limitations could be elaborated (i.e. no VCT client sample to estimate full effect). FURTHER INFORMATION ADDED
A few specific comments:
- 'Pre ART care' is "constituted" of a package of interventions (not "provide"). ALTERED

- I'm not sure you have measured impact on quality; see my point in the discussion. Continuity of care and service uptake may be more appropriate outcomes. ALTERED
- Throughout the paper many hyphens seem to be missing e.g. evidence-based; follow-up; HIV-related mortality; task-shifting, etc. STYLE ISSUE. WE HAVE TRIED TO ADDRESS THIS.

Statistical methods:
- Is the data skewed? Especially the mean time to ART initiation? If so, it would be important to also report the median values. Also, if the data is skewed (and I imagine it is), then you might need to consider non-parametric tests instead of ANOVA (e.g. Mann-Whitney U test). DATA REANALYSED
- I think it would be useful to include the SD along with the means. DONE
- P values should be reported in the table, as well as the F statistic for ANOVA. DATA REANALYSED
- You might want to consider doing tests for trend instead of basic tests; for the categorical data, you could do Chi2 test for trend; for the continuous data, you could do a regression, and produce a p value for the coefficient. (though do consider the skewness first). DATA REANALYSED
- Instead of the ANOVA, you could also consider giving a separate p value for group 2 vs group 1; and group 3 vs group 1 (e.g. Z test; or Mann Whitney; or do a post-estimation test on the ANOVA, e.g. Tukey Kramer, which gives tests for differences between the individual groups. DATA REANALYSED

References: There does not seem to be any discussion of other studies on pre-ART care, see point below in next section. ALTERED

-Background: A few additional comments on the background section:

As noted earlier, there is a need to discuss existing studies on the effectiveness of Pre-ART interventions as important scientific background to the study (at least briefly) ALTERED
- The references for components of pre ART services are WHO guidelines; are there any stipulated or agreed packages within Swaziland? NOT AT TIME OF IMPLEMENTATION. THERE ARE NOW
- You state on line 41 that HIV care was episodic and unmonitored; did you mean Pre-ART care (I'm not sure that the statement applies to ART services?) ALTERED

-HTC needs to be defined in the text since it appears later undefined (line 46). ALTERED
- I think CD4 levels not usually reported as copies/ml? (please check) ALTERED TO CELLS/MM3

Regarding the intervention description (figures on pages 16 and 17, not labelled in my PDF copy):-

What is the difference between regular review and clinical staging? Does a regular review not include clinical staging? CLINICAL STAGING IS PART OF REVIEW
- Does Pre ART care not include any other routine diagnostic tests? WE ARE UNCLEAR WHAT THIS MEANS
- What about condom provision on counselling on sexual behaviour/sexual health? Does this not form an important aspect of Pre ART care? Or advice on nutrition, etc? COUNSELLING IS PROVIDED (ADDED TO TEXT). DUE TO RESTRICTIONS OF THE CATHOLIC CHURCH, CONDOMS ARE NOT PROVIDED AT THE HOSPITAL
- The second diagram is missing 'up' in the last box.

Results presentation:

-Regarding Table 2-A, I find it hard to understand who the 'N' is: the results present a 'proportion being assessed for ART' – is this a proportion of those testing? Or is it a proportion of those given a Pre-ART file? If they are given a pre-ART file, why are so few assessed for ART at Round 1? ALTERED
- The sample size for each group and each indicator should be noted in the table, since I assume it is different. DONE
- The tense in the tables seems incorrect, should be past tense e.g. "proportion assessed for ART"? DONE

-The table should contain the p values, rather than just reporting them in the text. You could use "*" system to denote the significance level of each round compared to baseline. DONE
- Regarding the bar chart (no figure no.), the sample size for each group should be included on the graph. Also, this data is not presented as a table, and I think it would help to see the proportions, and certainly it would help to see a statistical test on each outcome reported. At a minimum, you could

include "*" above each bar to show statistical difference to Group 1. DONE
- Also, how is "HIV clinical staging" different to "Assessment of eligibility for ART"; if the client receives clinical staging, does that not mean that they are assessed for ART initiation? ASSESSMENT FOR ELIGIBILITY INCLUDES CD4, BIOCHEMISTRY AND CONTRA INDICATIONS
Interpretation/conclusion: The discussion section could be reorganised and refocused:
- There is no discussion of the results in light of other studies on pre-ART packages and their impacts. Overall the balance between the study discussion and the 'lessons learned' seems uneven; there is too much on lessons learned, and not enough about the research presented. ALTERED
- It is not really appropriate to refer to results table in discussion. The first paragraph also just seems to repeat the previous section in the results and could be written in a more reflective style. ALTERED
- You discuss a 'reduction in waiting times' due to the intervention, but the time lag to ART initiation is not a function of waiting times (which implies service waiting lists). THERE WERE SIGNIFICANT SERVICE WAITING LISTS WHICH HAVE IMPROVED
- You discuss cotrim results not improving, but was it not statistically improved between round 1 and 3? (no p values presented, as mentioned earlier) Why were the cotrim results still relatively low (<60%) after Round 3? ADDED
- The point about CD4 thresholds changing is critical, and should be discussed next to your point about causation, i.e. earlier in the discussion. The discussion about the causation could also include a reflection on any other policies or programmes were ongoing in Swaziland at that time, which could have contributed to the observed changes (I imagine not many, which would then give greater weight to the causation argument), but it is important to note it. The fact that you saw important improvements in key outcomes despite not including recently tested patients should also be highlighted, since this shows that the intervention is still powerful, even among those who were already at more advance disease stages.
- You call the study a before/after evaluation, which it doesn't necessarily seem to be since Round 1 received the intervention (unless I have got it wrong and you can explain it another way?) THERE WAS NO CLEAR BASELINE. THE DESCRIPTION " BEFORE AND AFTER " IS AN OVER SIMPLIFICATION BUT WAS THE CLOSEST MATCH OF THE LIST GIVEN
- I wouldn't have worried about the weight of the different intervention elements; you clearly stated that it was a package of care composed of several components, or clearly related to each other, and I don't find this undermines the outcomes reported in any way. NOTED, THANK YOU
- The term 'early service cohorts' on line 54 p10 is confusing to the reader; it would also be important to know what proportion of entries had missing data. ALTERED
- I think it would be useful to come up with a hypothesis about why there were more clients with advanced disease in the later sample, as otherwise it certainly undermines the study findings. As noted earlier, this may influence the key outcomes, or is even an outcome itself? WE ARE NOT CLEAR ON WHY THIS HAPPENED.
- You discuss implementing a pathway across 'multiple programmes' but the description of the intervention and Figure 2 really cover only HIV care and treatment. If greater coordination was achieved with other health programmes, this could be explained more in the text. CLARIFIED
- The discussion of vertical programmes could be made clearer. As could the last para on p.11, the last sentence of which becomes hard to follow. CLARIFIED
- I find Figure 4 not particularly useful; it is unclear if the 'actual impact' is what you implemented in your intervention, or whether this is what you wish you could achieve? NOTED
- I was not able to review Figure 5, it was not included (did you mean 4B?). APOLOGIES FOR THE CONFUSION
- You discuss task-shifting, and also mention it as part of the intervention, but don't provide any assessment of that component of the study. Your conclusion that "task shifting can address critical bottle necks in paritnet pathway" then seems overstated; as a reader we are unable to assess what the task-shifting actually achieved compared to other study components. AGREED.
- I find the final conclusion that you improved "quality of care" problematic; I know the HIV community has a tendency to measure quality in terms of outcomes (e.g. good adherence must mean good

quality), but many other fields of medicine prefer to use process measures to examine quality of care, more focused on adherence to clinical algorithms, or on inter-personal aspects of care. I think you can say that you improved continuity of care, as one aspect of quality of care. Or at least perhaps acknowledge that you only tackled one aspect of quality? CONCLUSION ADAPTED
One small comment regarding ethics: I would have questioned the ethics of the intervention component which led to medical staff visiting people's houses on motorbikes; this could seriously breech client confidentiality in the Swazi context where unusual visitors to homesteads could lead to involuntary disclosure of HIV status. THIS IS A WELL ESTABLISHED SERVICE FOLLOWING UP A RANGE OF DEFAULTERS (INCLUDING ART, TB, EPILEPSY, PMTCT AND OTHERS). IT WAS NOT INTRODUCED AS PART OF THIS INTERVENTION
Reviewer 2
1. In the pre-ART care discussion on page one, the point needs to be made that the “assessment” time is short because: (1) they waited too long before HIV testing; and (2) they tested in the past, did nothing, waited, re-tested again, and so on. ADAPTED
2. On page 1 (really figure 1), the first line of Figure 1 needs to be edited (the CD4 eligibility criteria is off. DONE
3. On page 1, you mentioned that “Patients with unknown status were tested in the HIV testing and counseling centre”. It would help to be more specific here: what portion were patients already in the ward because they were sick (are these hospital patients) and what portion were walk-ins for counseling and testing (basically the four groups in Figure 2). The current analysis pools data for four patients groups, which I think are rather different (at least the walk-in VCT compared to ward patients) into one analysis. The analysis would be stronger if the analysis was stratified by patient group. WE AGREE, BUT UNFORTUNATELY THIS DATA WAS NOT COLLECTED. THESE WERE NOT NEW TESTERS BUT THOSE WITH A KNOWN STATUS.
Also, if they were found to be HIV positive, it seems more correct to say that a blood sample from taken for CD4 testing (instead of a CD4 count was taken). ALTERED
4. The “task shifting” discussion is not clear to me and could be revised to clarify what tasks where to whom. ALTERED
5. On page 7 in the “sample” section, more information on patients is needed. ALTERED
It seems you have four populations that have different situations that are relevant for evaluating the results. For the three study group, it would be good to know the proportion in each category identified in Figure 2. WE ONLY EVALUATED THOSE WITH A KNOWN STATUS
After the cohort was defined (e.g. baseline in Feb/March 2009), how long after this period was the follow up period (e.g. 6 months after enrollment)? VARIABLE. AT LEAST 3 MONTHS. ADDED TO TEXT
Was this follow up period the same for all study groups? NO
This analysis would be much cleaner if the sample was only VCT patients newly diagnosed with HIV. Mixing up the four group into one clouds the interpretation of results, especially since the populations seems to be different across time (e.g. the proportion with CD4 < 200 increases substantially between baseline and cohort 2). AGREED, BUT DATA LIMITED. ONLY USED KNOWN STATUS.
Many studies are reporting improvements in the rates of ART eligible patients who initiate ART (and in shorter times). Dealing with the patients whose CD4 counts are high seem to be the more complicated group to try to retain in regular pre-ART care perhaps for two years before they become eligible. Thus, it might be better to report results for patients ART eligible and those who are not (e.g. in Table 2-A, no information is currently provided for patients who are not ART eligible).
6. Table 1: The proportions of patients in each group (from Figure 2 for example) would help to understand if the cohorts were similar over time. ALTERED
Also, did all patients sampled provide blood for a CD4 test? In our reviews elsewhere, even through all patients were offered the opportunity to provide a blood sample for CD4 testing, some number left after HIV testing without providing a sample. ALL THOSE IN PRE ART CARE PROVIDED CD4 TEST, BUT MANY WHO TESTED DID NOT ENTER ART CARE.
7. Table 2-A: Please explain somewhere what “assessed for ART means” (somewhere earlier

eligibility should also be defined). I assume CD4 count results and/or exam for staging, but it would help readers on familiar with Swaziland. ALTERED
8. I'd suggest the authors think carefully about their statistical analysis. I'm not sure ANOVA is adequate for what you are doing because of the underlying normality assumption and the implications of outliers. It would be reasonable to me to compare baseline to cohort 1 and baseline to cohort 2. I'd at least like to see an absolute risk or relevant risk analysis that included patient categories as covariates. I'd rather see the analysis stratified by patient group (the four in Figure 2) to see if patients eligible for ART are driving the results. RE ANALYSED
9. Figure 3 does not provide adequate information. The table(s) with the actual results and numbers would be better. Again, it would be better to report such information separately for the ART eligible and those not eligible. DONE
10. Because the intervention was a package of interventions, it is unclear if the whole package is important or if certain components drove the results. For example, having information on the % of patients who could not be contacted by phone for each group and knowing how many adherence officer visits were conducted for each group would help have perspective on the relative importance of different components of the intervention. DATA NOT READILY AVAILABLE
While I understand it is not possible to attribute the impacts of individual components, having information on amount of activities in each component (as well as the cost of the component – e.g. the active follow up) provides useful perspective. INFORMATION NOT READILY AVAILABLE
11. In the strengths and limitation section, you mention 'gaps in data' but no where in the analysis did you mention the issue. Where there gaps in data used in the analysis not reported?
12. The issue of known versus new diagnosis makes sense. While not a solution to this issue, it would make sense to separate the analysis into patient groups that are testing through different situations: walk-ins for VCT, provider initiated in OPD, and provide initiated in the Ward. Patients in the ward because they are sick with HIV/AIDS related conditions should already be considered "failures" for evaluation of pre-ART care and excluded from the analysis. AGREED BUT REGISTERS DID NOT RECORD SOURCE OF PATIENT

VERSION 2 – REVIEW

REVIEWER	Kathryn Church Research Fellow London School of Hygiene & Tropical Medicine.
REVIEW RETURNED	21/12/2011

THE STUDY	Please see my detailed comments. There are still some problems with the description of statistical methodology and various problems with the reporting in the paper.
RESULTS & CONCLUSIONS	Key figures should be cited in the text. Cell contents of tables not correctly defined. P values still missing. A small amount of further work needed on discussion section. Findings not discussed in light of previous evidence.
GENERAL COMMENTS	The article has been improved with the recent revisions. However, I feel it still needs some work before being accepted for publication. Here are my specific comments: Abstract: 1. Counselling could be listed in the Pre ART interventions here. 2. You cite a p value for difference between Baseline and group 2, but do not present any p value for differences between 2 groups in the results, only differences across the 3 groups (or otherwise the p values columns in tables are unclear?). There is no p value for the second result presented in the abstract. The last p value is

	incorrectly written (p<=0.001 should be p<0.001). 3. The first 2 points in the Weaknesses section could be combined as you repeat the same point twice in different ways. Background: 5. I think listing 18 references on line 17 (p.4) to support one point made is excessive. Some of these references do not even seem to be supportive of the point being made. Refs 13-15 seem the most relevant. I also feel you could include a bit more detail about what the relevant studies do say - what are the key problems in retention of patients after testing and/or enrolment for ART? 6. Line 20 (p.4), you say "did not provide data for this cohort of patients" - which cohort are you talking about? 7. Line 21 (p.4) – perhaps should be "there is emerging evidence on the relative success of different approaches"...; also should be "These include...". 8. Line 22 (p.4)- there is a notation on "CD4" but no footnote? 9. Line 45 (p.4)– should be "prevent onward viral transmission" 10. The section on 'components of pre-ART care' seems like a generic description of pre-ART, in which case I would strongly encourage the authors to include 'condom promotion' in this list alongside 'counseling'. While I understand it is not available at your hospital for religious reasons, it is available at most other ART clinics – or otherwise, please clarify that this list pertains to your own intervention and hospital only. 11. Line 31 p. 5, remove "but". 12. Unclear why you start with Figure 2 (cited on line 45, p. 5), why is there no Figure 1? Figure 1 is referred to on page 6 (line 33) but there are no figure headings in my copy of the file so I'm not sure if I'm missing one again? Perhaps I am missing either Figure 1 or 2 again? The acronyms in the first Figure presented need defining (they don't seem to be in the text? E.g. OPD, HTC?) Also, 'up' is missing from the last box in the first figure presented. I also wonder whether ANC should be listed in your boxes at the top, since it is one of the main locations where women test positive? Why does a person go from VCT to pre-ART HTC? (Is this the ideal, or the current?) 13. I feel the 'service gaps' in the care pathway (as identified and quantified by staff, line 46 p.5) could also be shown in the figure (again, perhaps this is because I don't have the correct Figure 1?); and this would be very informative for the reader – do you have these data? 14. There seems to be repetition around the presentation of a Pre ART service; the components are listed on page 4, and then again as part of the intervention on page 6. Have you exceeded the word limit? The first one could be cut, and all components fully described in your intervention section. 15. The active follow up is listed twice; once on line 12 (p.6) and again on line 41. One could be cut. 16. The title of Figure 1 in the text (line 45 p. 6) refers to the "current care pathway" – do you mean the new one or the old one? Methods: 17. In your analysis section, you say you conducted a Kruskal Wallis test (please note correct spelling) to look at differences in time between ART initiation AND WHAT? Half the sentence is missing. Do you mean "to look at time between identification of treatment eligibility and ART initiation over time" (or similar?) 18. The first outcome would be better phrased as "The proportion of patients with a documented assessment of eligibility for ART

	initiation (including XYZ), among patients recruited to the Pre ART service...". The second outcome would be better phrased as "The proportion of patients started on drug treatment among those eligible for ART initiation". Is the third outcome also among a sub-sample of patients, in which case this should also be noted in the same way? Also you should list your outcomes specifically for this third outcome, not just say 'including co-trim' – which ones did you measure, so it is clear? Results: 19. The tables should note the correct cell contents (i.e. N(%)), instead of saying "proportion" only. 20. Was age distribution skewed? In which case, it may be better to present median. (you have high SD scores). Why is SD not presented for Group 3? 20. P values should also be included in Table 1. 21. Should be "...adult patients with a CD4 count measurement" instead of 'value'? 22. Range of CD4 not presented for Group 3? 23. I'm still worried about the statistical tests in Table 2; ideally this should be a test for trend, OR as both reviewers noted previously, the ideal would be to compare each group to Baseline, rather than just one overall P value. I don't think you have conducted a Chi2 test for variance (see line 33) but rather a Pearson's chi2 test. 24. It would be better to highlight the key figures or at least the % increase in the text, rather than just referring to the table and chart. I don't think the figures should be reported together with the graph, but it is better to report the key figures in the text when referring to the chart. 25. There are no p values reported for the variables in the 'priority interventions' graph; these should also be added. Discussion: 26. I don't feel it is appropriate to refer to tables or graphs in the discussion, but instead to discuss the results more generally. 27. Line 49, p. 8– a 'however' or similar is needed to make the sentences flow here. 28. You talk about 'ongoing work' being needed to improve the drop out (line 50, p.8) but don't mention what this is, and it sounds rather ambiguous. 29. You still talk about a 'before and after' evaluation in line 19, but you did not collect data before the intervention, so I still feel it should be reworded. 30. Further restructuring is needed in the 'strengths and limitations section'. I feel the new point made on lines 29-31 should be better linked to the limitations on lines 36-40 as you are talking about the same problem here. The implication of not linking the testing data could be mentioned (i.e. you may have underestimated the magnitude of effect seen?). 31. You discuss differing 'cohort sizes' on line 40 (p.9) but this terminology has been replaced by groups; you rationale for differing sizes is still unclear; what is the implication of having a much larger group 2? 32. Your statement that increasing numbers of patients with advanced immunodeficiency should not have affected key outcomes seems optimistic (line 44 p9); is a busy nurse not more likely to assess a patient for ART who is very sick, than one who seems very healthy? I agree that it shouldn't have affected it, but the reality in clinics means that it may have affected it? 33. Line 12, p. 10: better to phrase as "separate silos of information

	within the HIV program (such as....)..." 34. There are 2 parts to Figure 3; these should be better signposted from the text since you refer to the sub-figures separately. i.e. line 28, p10 please refer to Figure 3a, and then to 3b on Line 29. 35. You could clarify that Lubombo is a region again on line 30. Other points 36. The title of the study still refers to a 'quality improvement study' but you have now removed most references to quality; perhaps something along lines of "An applied (or OR?) study to improve patient management and retention in care"...? 37. Please make capitalisation of 'Pre ART' consistent throughout and decide whether you want to hyphenate it or not as currently inconsistent still. 38. Data are plural.

REVIEWER	Bruce Larson Associate Professor Boston University School of Public Health USA
REVIEW RETURNED	21/12/2011

GENERAL COMMENTS	Please edit sentence in pre-ART care section on lines 29-31. The "ie." clause can be eliminated and the point simply made.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1 Comment
Abstract:
1. Counselling could be listed in the Pre ART interventions here.
Done
2. You cite a p value for difference between Baseline and group 2, but do not present any p value for differences between 2 groups in the results, only differences across the 3 groups (or otherwise the p values columns in tables are unclear?). There is no p value for the second result presented in the abstract. The last p value is incorrectly written ($p<=0.001$ should be $p<0.001$).
The results presented in the abstract are p values for differences across all three groups. For succinctness we only presented the summary proportions for the baseline and group 2 results. To clarify the intent of this analysis results of group 1 proportions are also included in the abstract.
3. The first 2 points in the Weaknesses section could be combined as you repeat the same point twice in different ways.
These are different points. One is about the inability to follow up cohorts, the other about the sample included

Background:
5. I think listing 18 references on line 17 (p.4) to support one point made is excessive. Some of these references do not even seem to be supportive of the point being made. Refs 13-15 seem the most relevant. I also feel you could include a bit more detail about what the relevant studies do say - what are the key problems in retention of patients after testing and/or enrolment for ART?
The references were used to provide the summary described below.
6. Line 20 (p.4), you say "did not provide data for this cohort of patients" - which cohort are you talking about?
Clarified
7. Line 21 (p.4) – perhaps should be "there is emerging evidence on the relative success of different

approaches"...; also should be “These include...”.
Amended
8. Line 22 (p.4)- there is a notation on “CD4” but no footnote?
This read CD4+, correct but not that common. Amended to read CD4
9. Line 45 (p.4)– should be “prevent onward viral transmission”
amended
10. The section on ‘components of pre-ART care’ seems like a generic description of pre-ART, in which case I would strongly encourage the authors to include ‘condom promotion’ in this list alongside ‘counseling’. While I understand it is not available at your hospital for religious reasons, it is available at most other ART clinics – or otherwise, please clarify that this list pertains to your own intervention and hospital only.
amended
11. Line 31 p. 5, remove “but”.
amended
12. Unclear why you start with Figure 2 (cited on line 45, p. 5), why is there no Figure 1? Figure 1 is referred to on page 6 (line 33) but there are no figure headings in my copy of the file so I’m not sure if I’m missing one again? Perhaps I am missing either Figure 1 or 2 again? The acronyms in the first Figure presented need defining (they don’t seem to be in the text? E.g. OPD, HTC?) Also, ‘up’ is missing from the last box in the first figure presented. I also wonder whether ANC should be listed in your boxes at the top, since it is one of the main locations where women test positive? Why does a person go from VCT to pre-ART HTC? (Is this the ideal, or the current?)
Figure references changed and acronyms added
13. I feel the ‘service gaps’ in the care pathway (as identified and quantified by staff, line 46 p.5) could also be shown in the figure (again, perhaps this is because I don’t have the correct Figure 1?); and this would be very informative for the reader – do you have these data?
Data not readily available
14. There seems to be repetition around the presentation of a Pre ART service; the components are listed on page 4, and then again as part of the intervention on page 6. Have you exceeded the word limit? The first one could be cut, and all components fully described in your intervention section.
The introduction focuses on the theory while the methods outlines the intervention as implemented
15. The active follow up is listed twice; once on line 12 (p.6) and again on line 41. One could be cut.
Amended
16. The title of Figure 1 in the text (line 45 p. 6) refers to the “current care pathway” – do you mean the new one or the old one?
clarified

Methods:
17. In your analysis section, you say you conducted a Kruskal Wallis test (please note correct spelling) to look at differences in time between ART initiation AND WHAT? Half the sentence is missing. Do you mean “to look at time between identification of treatment eligibility and ART initiation over time” (or similar?)
This refers to differences between the groups
18. The first outcome would be better phrased as “The proportion of patients with a documented assessment of eligibility for ART initiation (including XYZ), among patients recruited to the Pre ART service...”. The second outcome would be better phrased as “The proportion of patients started on drug treatment among those eligible for ART initiation”. Is the third outcome also among a sub-sample of patients, in which case this should also be noted in the same way? Also you should list your outcomes specifically for this third outcome, not just say ‘including co-trim’ – which ones did you measure, so it is clear?
Rephrased

Results:

19. The tables should note the correct cell contents (i.e. N(%)), instead of saying “proportion” only.
amended
20. Was age distribution skewed? In which case, it may be better to present median. (you have high SD scores). Why is SD not presented for Group 3?
The reviewers are correct that the SD is large however the data is normally distributed despite the large variance.
20. P values should also be included in Table 1.
Updated
21. Should be “...adult patients with a CD4 count measurement” instead of ‘value’?
amended
22. Range of CD4 not presented for Group 3?
Amended
23. I’m still worried about the statistical tests in Table 2; ideally this should be a test for trend, OR as both reviewers noted previously, the ideal would be to compare each group to Baseline, rather than just one overall P value. I don’t think you have conducted a Chi2 test for variance (see line 33) but rather a Pearson’s chi2 test.
A Pearsons Chi test was undertaken as the objective of this analysis was to compare the difference in proportions of people receiving assessment and treatment for ART. As these are comparisons of proportions of categorical values comparison of variances is not appropriate for categorical data in this table.
24. It would be better to highlight the key figures or at least the % increase in the text, rather than just referring to the table and chart. I don’t think the figures should be reported together with the graph, but it is better to report the key figures in the text when referring to the chart.
25. There are no p values reported for the variables in the ‘priority interventions’ graph; these should also be added.
Added

Discussion:
26. I don’t feel it is appropriate to refer to tables or graphs in the discussion, but instead to discuss the results more generally.
amended
27. Line 49, p. 8– a ‘however’ or similar is needed to make the sentences flow here.
amended
28. You talk about ‘ongoing work’ being needed to improve the drop out (line 50, p.8) but don’t mention what this is, and it sounds rather ambiguous.
removed
29. You still talk about a ‘before and after’ evaluation in line 19, but you did not collect data before the intervention, so I still feel it should be reworded.
amended
30. Further restructuring is needed in the ‘strengths and limitations section’. I feel the new point made on lines 29-31 should be better linked to the limitations on lines 36-40 as you are talking about the same problem here. The implication of not linking the testing data could be mentioned (i.e. you may have underestimated the magnitude of effect seen?).
amended
31. You discuss differing ‘cohort sizes’ on line 40 (p.9) but this terminology has been replaced by groups; you rationale for differing sizes is still unclear; what is the implication of having a much larger group 2?
Amended and clarified
32. Your statement that increasing numbers of patients with advanced immunodeficiency should not have affected key outcomes seems optimistic (line 44 p9); is a busy nurse not more likely to assess a patient for ART who is very sick, than one who seems very healthy? I agree that it shouldn’t have affected it, but the reality in clinics means that it may have affected it?

Assessment was introduced as a routine part of the service. Clinical staging is part of this.
33. Line 12, p. 10: better to phrase as “separate silos of information within the HIV program (such as....)...”
amended
34. There are 2 parts to Figure 3; these should be better signposted from the text since you refer to the sub-figures separately. i.e. line 28, p10 please refer to Figure 3a, and then to 3b on Line 29.
amended
35. You could clarify that Lubombo is a region again on line 30.
amended

Other points
36. The title of the study still refers to a ‘quality improvement study’ but you have now removed most references to quality; perhaps something along lines of “An applied (or OR?) study to improve patient management and retention in care”...?
amended
37. Please make capitalisation of ‘Pre ART’ consistent throughout and decide whether you want to hyphenate it or not as currently inconsistent still.
amended
38. Data are plural.
amended

Reviewer: Bruce Larson
Please edit sentence in pre-ART care section on lines 29-31. The "ie." clause can be eliminated and the point simply made.
amended