Reviewer's report

Title: Improving data quality and supervision of antiretroviral therapy treatment sites in Malawi: an application of Lot Quality Assurance Sampling

Version: 1 Date: 21 February 2012

Reviewer: Fabienne Shumbusho

Reviewer's report:

General comment:
The topic of this paper raises an interesting discussion, and it makes a contribution to improving ART programs overall quality and management and in particular with organization of efficient supervisions that help identify sites which do not meet threshold requirements.

I believe the text can be improved in the results presentation in particular to make sure that the elements of the two models that are being compared are consistently described (effectiveness; time of implementation, etc. of both approaches). Some edits are needed to avoid some ambiguity found in some sections of the paper. Specific comments provided below concern details on comprehension and some questions of clarification.

A comment about LQAS is that for those who do not know what it is, there is no description of what it is which limits the audience of this paper.

- Major Compulsory Revisions

Abstract objectives
The authors mention that “This study quantifies the current levels of data quality”: if this was one of the objectives, then could the authors describe how this was done, as it looks like the authors used the results of the supervision as reference but no other assessment was done? The suggestion is to be clear about whether the.

- Minor Essential Revisions

Abstract, Methods
"A study team joined supervision teams at 19 purposely selected sites… ". PLEASE PRECISE THE METHODOLOGY; THIS SAME EXPRESSION ("purposely") IS USED LATER IN THE TEXT AS WELL.
"76 records were then randomly extracted the time required for sampling and documenting secondary outcomes recorded" NOT CLEAR – CHECK SENTENCE

Abstract, Findings:
IT IS NOT CLEAR WHETHER THE FINDINGS ABOUT DISCREPANCIES WERE SIMILAR USING SUPERVISION APPROACH AND LQAS

Abstract, Conclusion:
YOU MENTION THAT " data cleaning required significant amounts of time with little effect on program estimates "; SEE COMMENT ABOVE AS THE FINDINGS DESCRIBED ABOVE DON'T SAY THAT THE TWO APPROACHES OF DATA QUALITY MONITORING ARE EQUALY EFFECTIVE.

Background, paragraph 3:
"Further, this full audit of patient records by the supervision team may be unnecessary at some sites, since clinic staff have correctly updated treatment cards and clinic registers in preparation for the supervision visit": IS THE AUDIT GOAL NOT TO VERIFY THAT THE SITES HAVE DATA CORRECT? I SUGGEST A REVISION OF THE REASON NOT NEEDING TO AUDIT SOME SITES.

Method, Paragraph 1
"The study team randomly sampled from the list of all patients enrolled at the site and pulled the corresponding treatment cards to determine time required for sampling and record secondary outcomes on patients still alive and on treatment (side effects and adherence, defined as eight or fewer pills remaining since last visit)." REVISE FORMULATION – THE USE OF “to” SEEMS CONFUSING

"…adherence, defined as eight or fewer pills remaining since last visit" REVISE FORMULATION AS NECESSARY: THE TEXT SUGGESTS THAT WHAT IS NOT COMPRISED IN THIS DEFINITION IS "NON ADHERENCE" AND THAT "eight or fewer pills remaining since last visit" IS ASHERENCE.

Results, paragraph 1:
THE TOTAL % IS 101%; # PLEASE CORRECT

Results, paragraph 7
YOU MENTION RANGES BUT NOT CONSISTENTLY THE MEAN OR MEDIAN VALUES. WHAT DOES "AGGREGATED" REFERS TO?

Discussion/Conclusion, paragraph 3
"we expect that this would lead to conservatively overclassifying sites as poor data quality": I SUGGEST YOU RECONSIDER WHAT YOU CALL "OVER CLASSIFYING" BECAUSE THE DISCREPANCIES REFERRED TO CAN RESULT, FOR EXAMPLE, IN ERRORS IN PROCUREMENT, AND RELATED BUDGETING.

TABLE 1:
- TITLE "Site Characteristics"; REPLACE WITH "Sites Characteristics"
- REVISE THE TOTAL %: IT IS 101% INSTEAD OF 100%
- Length of time providing treatment (private sites excluded from National ART Program numbers): THE TEXT IN BRAKETS IS NOT CLEAR, AND NOT REFERRED TO IN THE TEXT INTRODUCING THE TABLE.

TABLE 2:
- PLEASE REVIEW TITLE OF THE TABLE TO BETTER REFLECT THE CONTENT;
- REVIEW TITLE OF COLUMN "Without Exhaustive Review": IS THIS "CUMULATIVE VALUE"?

FIGURE 1:
FIGURE 1b IS NOT WHAT IS DESCRIBED; YOU DON'T NEED TO HAVE THE TWO FIGURES; 1b IS ENOUGH, UNLESS THERE IS A REASON OF PRESENTING BOTH.

- Discretionary Revisions

Title:
Check harmonization with rest of the text: "antiretroviral therapy treatment sites"; In Abstract objectives, you refer to "antiretroviral therapy sites" # KEEP SAME FORMULATION ACROSS THE PAPER.

Abstract, Objectives
“Quarterly supervision of these antiretroviral therapy (ART) sites ensures high quality care…": IS THERE A REFERENCE AS TO THE HIGH QUALITY CARE ENSURED BY QUARTERLY SUPERVISION?

Background, paragraph 2
"The teams also review the patient cohort analyses produced by the ART staff and clean patient data and aggregate quarterly and cumulative reports on 44 data elements." - YOU MENTION ABOVE THAT THE SUPERVISION TAKES "UP TO ONE DAY"; DOES THE TEAM DO DATA CLEANING DURING THE SAME VISIT? IF YES THEN OK, BECAUSEE THAT'S WHAT THE TEXT SUGGESTS OR SAYS.

Method, Paragraph 1
"A study team joined supervision teams for the 4th Quarter (Q4) 2009 supervision (11–22 January 2010)"; THE TEXT MIGHT CREATE CONFUSION SO PLEASE CONSIDER THIS SUGGESTION: "From 11 to 22 January 2010, a study team joined supervision teams..."

Results, paragraph 2:
"... identified few discrepancies in outcomes and treatment regimens comparing the treatment card and register ...": DO YOU MEAN "WHEN comparing"?

" Of the 210 discrepancies, a majority (58%) were incorrect ...": DO YOU MEAN " Of the 210 RECORDS WITH discrepancies..."? IF SO CORRECT THE REST OF THE SENTENCE AS WELL.

Results, paragraph 3:
"For the 11,034 patients alive and on treatment": DO YOU MEAN "patients RECORDS of patients alive..."?

"For the 2,605 patients who initiated treatment in Q4 2009, 378 (15%) had errors in the registration data": SAME COMMENT AS ABOVE; ALSO 378 IS > 210 REFERRED TO IN THE PREVIOUS PARAGRAPH - NEED TO MAKE SURE THERE IS NO CONFUSION BETWEEN "RECORDS WITH DISCREPANCIES" AND "DISCREPANCIES" AND ELSEWHERE DIFFERENTIATE WITH "ERRORS".

Results, paragraph 5
WHAT DO YOU DO WITH SITES QUALIFYING BETWEEN "HIGH QUALITY" AND "LOW QUALITY"?

Discussion/Conclusion, paragraph 3
"formally registered patients ": DO YOU MEAN "FORMERLY registered patients "]?

TABLE 2:
- PLEASE REVIEW TITLE OF THE TABLE TO BETTER REFLECT THE CONTENT ;
- REVIEW TITLE OF COLUMN "Without Exhaustive Review": IS THIS "CUMULATIVE VALUE"?

FIGURE 1:
FIGURE 1b IS NOT WHAT IS DESCRIBED; YOU DON'T NEED TO HAVE THE TWO FIGURES; 1b IS ENOUGH, UNLESS THERE IS A REASON OF PRESENTING BOTH.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:
I declare that I have no competing interests