Author’s response to reviews

Title: Evaluation of the Performance of Routine Information System Management Framework (PRISM): Evidence from Uganda

Authors:

   David R Hotchkiss (hotchkis@tulane.edu)
   Anwer Aqil (anwer_aqil@isi.com)
   Theo Lippeveld (theo_lippeveld@isi.com)
   Edward Makooyo (emukooyo@yahoo.com)

Version: 2 Date: 15 January 2010

Author’s response to reviews: see over
Dear editors,

We very much appreciate the comments we received on our paper. We thank the reviewers. The comments were very useful in helping us strengthen the manuscript. Please see below point-by-point responses to the comments.

We look forward to your response.

Best regards,

David Hotchkiss, PhD
Professor

Reviewer 1

This is an interesting research project that is conducted in a developing country. The authors have strived to address an important, yet unknown issue: the determinants of routine health information system performance and the effectiveness of routine health information system strengthening interventions. The question posed by the authors is a big and well defined.

We very much appreciate these comments. Although the reviewer comments that the research questions are well-defined, Reviewer 3 suggests that the objectives of the paper could be further clarified. We agree with Reviewer 3. As a result, we have revised the abstract and introduction. In the revised manuscript, we now state that the objectives of our study are: a) to assess the reliability and validity of the PRISM instruments that capture the various components of the PRISM framework, and b) to assess the validity of the PRISM framework by providing empirical evidence of the hypothesized relationships.

Are the methods appropriate and well described? The data collection methods are relevant, implementable and can lead to answers to the research questions. Cronbach’s alpha is normally above 0.7, suggesting good internal consistency of the measurement statements for each variable; however, as the loading value for each indicator is not known (in fact, if they are presented as the ‘factor score’ presented in Appendix I, they may be too low to suggest that each group of indicators are measuring the same variable. This may threat the validity of the PRISM data collection tools. Therefore, clarification has to be made to ensure the validity of the measurement instrument.

We are pleased to get these comments on the merits of our study. We assume the last part of this comment, on loading values and validity, refers to the principal components analysis. Just to clarify, we applied principal components analysis only for the purpose of constructing a composite index of the use of routine health information data at the facility or higher levels, which is an important part of our methodological approach to assess the validity of the instruments measuring PRISM inputs and of the overall PRISM framework.
As we mention in the paper, the use of RHIS information is operationalized by a series of dichotomous indicators, including: whether RHIS information was discussed, whether RHIS information was used to help make decisions; whether RHIS information was used to help take follow-up actions or to refer issues for action; and whether updated information on various topics was displayed. Our view is that these are unambiguously valid measures of the use of information, as they are based on facility observation and record review, the "gold standard". As recommended by the PRISM analysis guidelines (Aqil et al. 2005), these indicators were aggregated to generate a composite indicator of the use of RHIS information. This approach gives equal weight to each of the indicators used in the index. We tested whether this assumption makes a difference in the analysis by applying principal components analysis (PCA) to generate the index. As mentioned in the paper, PCA is a well-established method to create summary indices using weighted sums (Jolliffe 2002). PCA generates the weights that maximize the variance of the resulting composite index. In generating an index of RHIS data use, the advantage of the PCA approach over the simple addition approach is that it imposes fewer restrictions – the PCA approach generates weights while the simple aggregation approach is just a weighted sum where all weights are restricted to have the value of one. For both the bivariate and multivariate analysis based on indicators of the use of RHIS information, we used both approaches described above, and our results are very similar. Thus, where sophisticated statistical technical skills and software are unavailable, the aggregation approach serves the purpose of creating an overall use of RHIS information score. In the manuscript, we chose to present the results based on the index created through the PCA approach, but our results are robust to the method used.

The issues described above are now mentioned in the revised version of manuscript. We hope we have addressed the reviewer's comments, but please let us know if further clarification is needed.

Major compulsory revisions

The authors have difficulty to reach the adequate sample size for principle component analysis and structural equation modeling to quantitatively validate the construct validity of the PRISM framework. As such, the construct validity of the framework is not adequately validated. A much bigger size of sample may be required to fully validate the complex evaluation framework PRISM. As most likely the participants in the 2007 survey were different from those in 2004, possibly the survey responses from these two years can be amalgamated for statistical data analysis.

The reviewer raises an important point regarding sample size and power to conduct multivariate modeling in order to validate the PRISM framework. Due to resources constraints, we did not have the ability to increase the sample size of facilities. However, the sample size of 110 in each of the two years of the study yielded a total of 220 facilities, which are used in pooled analyses presented in the manuscript. This sample size is large enough to find a correlation coefficient of 0.2 between two variables of interest with power of 0.81 and a significance level of 0.05 (Cohen & Cohen 1983). Similarly, in a simple multivariate regression analysis of ten independent variables, a sample size of 210 would suffice to yield a power of 0.8. However, as pointed out by Cohen & Cohen (1983), many independent variables are a set of variables. For example, a multi-dimensional construct might have an overall score but have many subscales of various dimensions. Thus, it is better to consider this overall score as a set of variables and for the power calculation to be based on that. To explain our point, we present below power calculations based on alternative model specifications. Table 1 shows that with a small effect size of .06 or (5% variance explained), five independent variables with a sample size of 200 will yield a power of .95. On the other hand, including 24 independent variables in the model will reduce the power to .79. However, if the effect size is increased to .12 (13% variance explained) then despite 24 independent variables the power still remains above convention (.87). Please note that 24 independent variables could be a single or a set of independent variable. The point is that our analysis does provide sufficient power for conducting
multivariate regression analysis. We have revised the manuscript by including a sentence with the results of our power calculations.

Table 1: Power calculation* based on the sample size, independent variables**, effect size*** and the Noncentrality parameter $\lambda$

<table>
<thead>
<tr>
<th>Total sample size (N)</th>
<th>Number of independent variables (U)</th>
<th>$(V) = (N-u-1)$</th>
<th>Effect size ($f^2$)</th>
<th>$\lambda$</th>
<th>Power –based on $\alpha=0.5, \lambda, u$ and $v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>5</td>
<td>194</td>
<td>0.06</td>
<td>12</td>
<td>0.95</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>189</td>
<td>0.06</td>
<td>12</td>
<td>0.79</td>
</tr>
<tr>
<td>200</td>
<td>15</td>
<td>184</td>
<td>0.06</td>
<td>12</td>
<td>0.73</td>
</tr>
<tr>
<td>200</td>
<td>24</td>
<td>175</td>
<td>0.12</td>
<td>24</td>
<td>0.87</td>
</tr>
</tbody>
</table>


**independent variable could be single or composed of set of variable  
***effect size $f^2$ could be converted into partial $R^2$ by formula of $f^2/(1+ f^2)$. Thus 0.06 is equal to 5% $R^2$

Although the sample size is large enough for current analyses presented in the paper, we agree with the reviewer that our sample is not large enough to include all variables of the PRISM framework to conduct structural equation modeling. However, we do believe that it is not necessary and that validation can be accomplished by criterion-related validity analysis and construct-validity analysis among different components of the framework with sufficient sample size and power. For example, we assessed bivariate relationships among a number of organizational and behavioral factors described in the PRISM framework, including a culture of information and motivation, RHIS tasks self-efficacy and RHIS tasks competence. Statistically significant relationships were found to exist between RHIS task self-efficacy and RHIS use of information as well as between the availability of RHIS staff and use of information. Thus, we are of the view that the study results support the validation of the PRISM framework’s hypothesized relationships and conclude that PRISM framework is a useful approach for elucidating the relationships among various components of the RHIS and provides useful information on the strengths and weakness of the system, creating opportunities for interventions to fill the gaps and running the system in a harmonious and efficient manner.

The big number of statements in the questionnaire survey instrument suggests that the measurement instrument is not well developed and still needs several rounds of iteration to reach its maturity.

The PRISM tools consist of several construct measures and each construct measure has a maximum of three to five items. For example, in measuring a culture of information, there are several dimensions, and each dimension is measured by three to five statements. For each scale and sub-scale, we describe in the manuscript the specific statements that were used as well as the results of Cronbach’s alpha analysis. Thus, we are not sure what the reviewer means by commenting that there are a big number of statements in the questionnaire. We do agree that the PRISM instruments can be further improved, and believe that
knowledge from future applications of PRISM will help provide insights on how to further improve the tools.

P18. Appendix 1. The results of PCA are not properly presented, which makes it difficult to judge the validity of the measurement instrument. It is not clear whether ‘factor score’ is ‘loading factor’ of each indicator. If so, it suggests that the measurement statements for one variable are not only measure this variable, but several variables. For example, there are eight statements measuring use of information. As none of them has loading factor that is bigger than the conventional cut-off value of 0.6. This suggests that these eight statements are not just measure one variable, but several variables. The author needs to conduct more detailed analysis and present the correct information if the results of PCA are shown to the reader.

To construct the composite indices of use of RHIS information, we used the “PCA” command in STATA Release 10 (2007), and to present the results of the Principal Components Analysis, we followed a well-accepted approach used in the peer-reviewed literature. As examples, please see the following references that apply PCA to construct composite indices of socio-economic status:


Also, in reading the comments, we wonder whether the reviewer is referring to another type of analysis and not PCA, as PCA does not yield loading factors. According to the references in the manuscript and in Vyas and Kumaranayake (2006), the method generates factor scores or weights, which we present in the appendix.

Inconsistency in use of terms. In the ‘Abstract’, ‘RHIS task self-efficacy’ is mentioned. This is presented as ‘Confidence levels for HIS Tasks’ in Figure 1. To keep consistency, ‘Tasks’ should be ‘tasks’.

As suggested, we have replaced “Tasks” with “tasks” throughout the paper.

Poor grammar: ‘a promotion of culture of information’ should be ‘promotion of culture of information’

As suggested by the reviewer, we have made this correction. In a few instances, we are of the opinion that “the promotion of a culture of information” is correct, and have used this wording.

Table 1. Unclear. You should separate the presentation of the number of facilities and the number of staff members in each facility that participated in the study.

As recommended by Reviewer 3, we have dropped Tables 1-3, and have placed most but not all of the information in the text.

The last sentence in the first paragraph on Page 5 should be put into ‘Methods’ section.
We did not make this change, as we were unclear about the rationale for the comment. The sentence referred to above states how we hope the results of our study will contribute to future RHIS assessment and evaluation efforts. We believe the introduction is a more appropriate section for this type of sentence than the methods section. Please let us know if you have any objections.

Page 6 First dot point. ‘Data demand’ may be better placed under ‘diagnostic tool’ category. ‘the merit system’ is poorly structured and not understandable. The last sentence ‘RHIS performance as a criterion for annual appraisal, and supportive management’ is not clear.

To respond to this comment, we have revised the description of the Organizational and Behavioral Assessment Tool. The terms ‘data demand’, ‘merit system’, ‘RHIS performance as a criterion for annual appraisal’ and ‘supportive management’ have been dropped from the section, as questions on these issues were not included in the version of the instrument administered for the study.

Page 6. Para.3, ‘The decision of which districts to include’, should be ‘The decision of which districts to be included’

The sentence has been revised, as suggested above.

Page 6 Last para. The last sentence is not supported by Table 1.

The sentence referred to above was incorrect, and has been fixed. Also, as mentioned above, Table 1 has been dropped from the paper.

Page 7. Para 2. The first sentence is not understandable.

The sentence – and the paragraph – has been revised. We hope this is now understandable.

Page 8. Para. 2. Have the authors used valid scale to measure motivation in this research? There is no reference to the validation issue.

We believe there is a misunderstanding on how we assessed the validation of the motivation scale. Construct validation is about whether the hypothesized relationships among variables hold in the real world. As hypothesized in the PRISM conceptual framework, motivation affects RHIS task self-efficacy, RHIS task competence and RHIS performance. Our results provide empirical support for these hypothesized relationships (see Table 6 in the paper), indicating that the motivation scale is valid.

Motivation is considered a validity criterion for a culture of information (Pedhazur 1991). Again, our results suggest that the culture of information scale is valid as both are shown to be statistically associated in our study.


Page 9 Para 3. For the first time, ‘Job satisfaction’ is mentioned as a construct here. Where has it been placed in the PRISM Evaluation Framework?
Job satisfaction is not part of the PRISM framework. However, it is included as a validation construct for motivation, promoting a culture of information and other organizational variables based on previous studies of organizational culture and organizational communication (references available upon request).

*Page 9 last sentence is not readable.*

We have now revised this sentence. We hope this more clearly describes the OLS models that we estimated.

*Page 10. The last sentence in the first paragraph is not understandable.*

This sentence has now been revised to describe more clearly that the response codes for the question on educational attainment were different in the 2004 and 2007 surveys, which prevented us from comparing the results on educational attainment.

*Page 10. The comparison of number of male and female participants is not important as the statistical difference is not calculated for this variable.*

In response to the comment above, we calculated the statistical difference for the variable in both 2004 and 2007. For each survey year, the difference was not found to be statistically significant at the 0.10 level of significance. We note this finding in the revised manuscript.

*Page 11. Para1. The median data is discussed. Whereas the data presented in Table 3 is mean.*

Thank you for pointing out this error. We have fixed this problem by replacing ‘median’ with ‘mean’. Also, as mentioned above, Table 3 has been dropped from the paper.

*Page 11. Heading ‘Internal consistency…’ need to be re-written.*

The sub-heading has been corrected – thank you.

*Table 4. Eight items do not have scores of Cronbach’s alpha.*

We reviewed table 4 and verified that all items do have scores for Cronbach’s alpha. The eight items at the end of the table, on page 35, should be grouped with the last two on the previous page (page 34). All ten indicators were used in the composite indicator measuring ‘overall promotion of a culture of information’.

*Page 12. Para. 2. The description is not clear. The statistical analysis is dubious.*

We revised the paragraph to better describe the Cronbach’s alpha analysis of the composite indicators of the promotion of a culture of information. We hope the revised description is clear.
Page 13. *The method for testing test-retest reliability is not specific. The results are not mentioned in the paper.*

In response to the first comment, we have revised the description of the method used for the test-retest analysis, and have moved this description from the results section to the methods section. Regarding the second comment, the results of the test-retest analysis were mentioned in the first version of the manuscript (as part of our presentation of Table 5), but the presentation lacked clarity due to our presenting results on RHIS data quality before mentioning Table 5. The paragraph at issue has now been revised for clarity and flow, and by adding a sentence that presents the results of new Chi-square tests that support the reliability of the instruments for assessing RHIS data quality. By making these changes, we hope that we have addressed the reviewer’s two comments above.

**P15. para. 2. The first sentence, ‘the mean level of the promotion of a culture of information was significantly higher’, is poorly written.**

In the revised version, we fixed this problem by replacing the sentence with two properly-worded sentences describing the results of the test-retest analysis. We hope the new sentences are clear.

**How do you measure that the changes are internally consistent?**

No specific tests were conducted to assess whether the changes from 2004 to 2007 are internally consistent. However, in the revised manuscript, we draw attention to the results that assess the relationships found among various variables of interest as hypothesized by PRISM framework. Please note that the results section did include a number of sentences that provide explanations on how the changes in variables from 2004 to 2007 are interrelated, coherent and logically consistent. These sentences have been edited and moved from the results section to the discussion section, as suggested by Reviewer 3.

**P16. The last sentence in Para. 2 is poorly written.**

The sentence has now been revised. It now reads: “Regarding completeness, as measured by the proportion of facilities submitting reports to the district, 50 percent or more of the district health offices included in the study showed, through a record review, that 80 percent or more of the health facilities under their authority submitted timely reports in 2007 as compared to 75 percent in 2004, indicating that more facilities were reporting in 2007.”

**P17. Para 2. The second sentence the use of word ‘allocated’ does not make sense here.**

We fixed this sentence by replacing ‘allocated’ with ‘associated’. We appreciate the reviewer identifying this mistake.

**It is helpful if a detailed proof read and editorial services could be conducted on the paper.**

The paper has now been carefully proof-read. We also revised the style of the references according to the journal guidelines. We very much appreciate the reviewer’s comments.
Reviewer 2

There is little doubt that we need to know more about the performance of health information systems—overall, they perform badly in terms of providing relevant information of adequate quality and timeliness; and the information they contain is little used in health system management in many, if not most, low income countries. Staff in many developing countries must spend significant amounts of time filling in the routine data forms and transmitting them to higher levels of the system knowing full well the information will frequently not be used. At the same time managers of higher levels of the health system are unable or unwilling to revise the system which is the result of additions (but seldom, if ever, deletions) over a long period of time; the content continues to grow irrespective of use. There is little doubt that tools to help evaluate the performance of routine information systems are needed.

Major Compulsory Revisions.

As they review the paper I suggest that the authors address the following points. HIS evaluation requires that, at a minimum, there is consideration of the content of the system, the extent to which the information could be used at the various levels of the system (i.e. is there the decision space within the individual health system for actual use of the information), whether the information provided is relevant to the actual decision space that exists, whether the information provided is of adequate quality to allow decisions to be made, whether the information arrives on time for a decision that makes a difference to be made, and whether the information is actually used. In many health systems the information on acute infectious diseases is so out of date when it is collated that it cannot be used to intervene in a meaningful way because the epidemic has already peaked and passed; at the same time information on chronic non-communicable diseases such as diabetes and heart disease, often the dominant cause of morbidity and mortality is not available. The problem is that the PRISM framework at the center of this paper concentrates only on the process and does not provide any evaluation of content, its appropriateness at each level of the system, and whether the information can actually be used in a meaningful way. Further, although it is not made explicit in the paper, this system seems to deal with information from the public part of the health system and to ignore the private sector, an increasingly important component of the health system, especially for ambulatory care.

See our response to the above comments after the next comment.

Thus, this paper would benefit greatly from a succinct description of the health system in Uganda, the actual information needed to manage the health system at each level, actual information the health system is providing (and not providing) at each level, the timeliness of the information, the extent to which health system managers and staff at the various levels actually have the scope and resources to make decisions based on the HMIS that they can implement quickly, the quality of the data, and so on. Provision of this context would allow the PRISM approach to be assessed and the performance of the information system to be evaluated.

We thank the reviewer for these well-taken and very useful comments. In the background section of the revised manuscript, we now include a number of new paragraphs that describe the overall country context of the study. The revisions include a description of the health system, an overview of health system strengthening initiatives introduced by the government in recent years, and a description of the routine health information system (termed the HMIS in Uganda). We believe the decentralization reforms are particularly noteworthy in Uganda, as they have increased managerial decision-space at the local government- and facility-level.
In addition, we wish to point out that PRISM is not about process only, but actually assesses the RHIS structure and functions and its relevance to health system management. It takes a normative approach that RHIS should have minimum set of standards, knowledge, skill and competences (content) at different levels for its effective functioning. Thus, by assessing strengths and weaknesses of the information system, the PRISM framework provides direction on how to fill the gaps. It assumes that decision space is available at different levels (which is in fact the case in Uganda given the decentralization reforms described above) as RHIS is developed to address the existing information needs at different levels for making planning and managerial decisions. PRISM indirectly assesses decision space by exploring problem solving skills and perception of problem solving in the health department under promotion of a culture of information. In working with the Ministry of Health on this study, the Ugandan Government was more interested in knowing about efficiency and effectiveness of the system as noted by level of data accuracy, data completeness, and timeliness and whether information is used and different level and what factors these system outcomes. No emphasis was made to address relevance of collected information, as the government views that it is collecting relevant routine information, except for services related to HIV/AIDS, which was omitted in 2004 and then added in 2006. We also would like to point out that a PRISM tool is available which measures relevance of information but that tool was not used in the present study. Thus, the paper is more about whether PRISM framework and its tools were able to portray the big picture of the Uganda HMIS. Dr. Eddie Mukooyo, a co-author of the manuscript, is head of the NHIRC and subscribes to the paper findings.

In addition, the Uganda HMIS collect information from the private sector as well, and is probably the only African country where the private sector is part of the HMIS. Many private sector facilities were part of the sample as well. Thus, the study does cover the whole HMIS system. In fact, in 2004, the PRISM framework and its tools were used to assess the education information system in Uganda as well. However, this information is provided for the reviewer to dispel the impression that PRISM is only suited to health or public sector information system and not relevant to the paper at hand. For more detailed information on the PRISM framework, please see Aqil et al. (2009) through the link below – the paper is available free of cost.

http://heapol.oxfordjournals.org/cgi/content/abstract/czp010

1. The RHIS performance variable for 2004 is whether charts, graphs etc are displayed at the facility. And there is no assessment of whether the graphs, charts displayed were accurate. I have been in many health facilities in a range of low income countries, most of them had charts, graphs, tables displayed, many of them were out of date and/or inaccurate. I’m not sure that this is a very useful indicator of RHIS performance – at least it needs some justification in this paper.

We understand the reviewer’s concern about the indicator of RHIS performance for 2004. By design, this indicator is a measure of the use of information, and not of the quality of information, which is assessed through the record review of RHIS reports and registries available at the time of the facility visit. There are a number of reasons why we chose not to assess the accuracy of RHIS information displayed. First, designing instruments to assess the accuracy of displayed data would be difficult, as the government has not issued guidelines on what type of data should be displayed in facilities. In fact, the survey enumerators noted a wide variety of types of information displayed, involving various types of services, time periods (monthly, quarterly, yearly), and methods of display (pie charts, bar charts, tables and other forms of display). Second, displayed data could be inaccurate due to the staff having limited capacity to create graphs, charts and maps, as opposed to the capacity to produce accurate data for RHIS reports (which is assessed in the study). Third, the source of displayed data could be the RHIS, surveys, or both, which complicates efforts to assess accuracy. Addressing these issues was beyond the scope of the current
study, but we believe they are interesting and worth considering in the design of future RHIS assessments.

Moreover, the index of the use of RHIS was substantially revised in the version of the facility instrument administered in 2007 by both including additional questions on how RHIS data are used and by collecting information on whether the data displays had been updated within the three month period prior to the facility visit. For the pooled 2004-2007 data analysis, however, we were forced to use the more limited index. Given these limitations, we believe that the results of the pooled data analysis on the use of RHIS information should be interpreted with caution, and that the 2007 cross-sectional analysis provides more insights on the use of information.

In order to respond to the reviewer’s comments above, we have revised the methods section of the manuscript as follows. First, we more clearly state that the 2004 indicator that we used in the pooled data analysis measures use of RHIS information, and not overall RHIS performance. Second, we more clearly describe the limitations of the 2004 indicator (i.e. it measures only the display of data and not the quality of data). And third, we describe that the accuracy of information is assessed through the record review of RHIS reports and registries and not of the display of information.

2. in 2007 there is no assessment of the extent to which questionnaire responses seem to relate to what people actually do (display graphs etc).

We are not sure we completely understand this comment, but assume the reviewer is referring to whether the study assesses the associations between the indices of the potential determinants of RHIS performance and what health workers actually do with RHIS data (i.e. display data, use data for decisions, etc.). As mentioned in the revised manuscript, the overall purpose of the PRISM approach is do exactly this – assess the RHIS by measuring whether and how technical, organizational and behavioral determinants influence data quality and data use. The study instruments include many questions on what health workers do, including the types of information collected, how they are trained, how they are supervised, and how RHIS data are used (i.e. in displays, in meetings and to help make decisions). Moreover, the study methods include both bivariate and multivariate techniques that relate the PRISM determinants to the use of RHIS information in order to provide empirical evidence of the hypothesized relationships. All of this helps address the objectives of the study, which are a) to assess the reliability and validity of the PRISM instruments that capture the various components of PRISM framework, and b) to assess the validity of the PRISM framework. We hope this response addresses the reviewer’s comment.

3. the paper seems ambivalent on whether there were changes between 2004 and 2007 that might have affected quality. Thus, on pages 13 and 20 it is stated that no significant RHIS strengthening interventions were carried out between the surveys; on pages 15 and 21 it is stated that the training on revised data collection tools might have contributed to changes in quality; then on page 16 it is said to be possible that revision of data collection forms might have given staff the impression that management wanted to improve the culture of information; but the results are said to show a deterioration in data quality between the surveys (p14).

We thank the reviewer for providing us with an opportunity to clarify this issue. In response to this comment, we conducted additional analysis of our data on RHIS data quality to determine whether there were meaningful changes from 2004 to 2007. New Chi-square test results suggest that data quality, as measured by the accuracy of RHIS data available in the sample facilities, did not significantly change over the period. Regarding the comments on RHIS interventions, our understanding is that the Ugandan government did not introduce any major RHIS interventions during the 2004 to 2007 interval, but did
incorporate into the RHIS additional information on HIV/AIDS services in order to respond to emerging health needs. In the revised manuscript, we now present the new test results described above (in the results section) and have revised the discussion on the RHIS interventions carried out during the period between the surveys (in the discussion and conclusion sections).

4. The Methods do not actually describe how the staff interviewed were selected – e.g. at random from a staff list or opportunistically.

In the revised manuscript, we have clarified the process used to select the facilities and the facility staff. Facility selection was random and then the facility in-charge was selected along with any other staff involved in data collection. If the facility in-charge was not present, his or her substitute was selected.

5. there was a major shift in the composition of the staff interviewed between 2004 and 2007 – in 2004 ‘other’ staff were 45% of respondents whereas in 2007 this category constituted 73%. This is a major change and needs to be taken into account in comparing the two surveys. Is this change in staff category correlated with age and/or sex of the respondent?

This confusion arises because the code responses for staff categories were inadvertently different in the 2004 and 2007 surveys. As result, we are not able to compare the composition of the staff between 2004 and 2007. In the revised manuscript, we now clarify this issue.

6. Table 8 – the footnote to this table refers the reader to Appendix Table 1 which, in turn, has a footnote that says all the indicators of use of information were verified through record review and facility observation. But on page 21 it is stated that the study was not able to assess the completeness of monthly report data due to problems in administration of the survey. The extent of the missing data by variable and level needs to be presented in the paper.

Thank you for these comments. In the revised manuscript, we have now added the N’s used to generate a) the Cronbach’s alpha analysis, b) the test-retest analysis and c) the multivariate results. This information is presented in the tables of results.

In response to the comment on the assessment of data completeness, we have now clarified why we were not able to carry out this analysis. We mention in the paper that the denominator is the total number of data elements that are supposed to be reported in the monthly reports and the numerator is how many elements were actually filled and reported. Thus, the completeness of monthly report was supposed to be assessed by counting the number of items filled. During the survey, counting the number was not done properly due to a misunderstanding on whether to count all data items or only 10 percent of the total data items. Thus, the data on the completeness of the monthly reports could not be analyzed and reported.

7. in Table 9 the only consistently significant independent variable is having an RHIS assistant on staff. The dependent variable is display of graphs etc. This is not particularly surprising as this presumably would be part of the job description of that person. But this may have little relationship to the RHISI performance and use (see comment 1 above).

We appreciate the reviewer’s observation. As described above, in response to an earlier comment, the dependent variable for the model consists of a much more limited index of the use of RHIS data than that
used in the 2007 cross-sectional analysis. In the revised manuscript, we now offer the comment above as a possible explanation for our findings, and include a sentence regarding the limitations of the model.

Discretionary Revisions.

An alternative approach to this paper that the authors might wish to consider is to make it a review of the Uganda RHIS in which, after summarizing the health system and describing the RHIS in some detail, they concentrate on the 2007 data to provide an overall assessment of performance.

While we agree that this type of review paper would be very useful, we believe it is outside the objectives of this particular research study, which is to assess the reliability and validity of the PRISM tools and conceptual framework.

Reviewer 3

1. While the paper is well written, it is a bit difficult to follow because the authors are attempting to address two separate goals in the manuscript – assessment of the psychometric properties of the instrument and evaluation of the implementation of the instrument. The paper needs to be either simplified with the goals more clearly stated or divided into two separate manuscripts.

We agree with the reviewer that the goals of the paper need to be more clearly stated. As described above, we have revised the paper by re-stating the objectives of the study. As now mentioned in the abstract and background section, the specific objectives of the study are: a) to assess the reliability and validity of the research instrument, and b) to assess the validity of the overall PRISM conceptual framework.

In revising the manuscript, we tried to create a balance in describing psychometric properties of the tools and the general validation of the PRISM framework. The paper is for a general audience of policy makers, health services providers, HIS practitioners and health services researchers and thus detailed psychometric properties analyses were avoided, for which we need a separate paper. The main premise is that without reliable and valid tools, it is difficult to validate the framework. Thus, the major reliability measures for the tools are presented while construct validation is subsumed under PRISM framework validation.

2. Related to the issue in #1, some of the tables are unnecessary (e.g., Tables 1-3) and the information could be placed in the text.

As suggested, we have dropped from the manuscript Tables 1-3 and have placed most but not all of the information in the text.

3. Please define types of facilities (IV, III, II).

In the revised manuscript, we added to the background section definitions of the types of health centers included in the analysis (IV, III and II). This is part of the description of the Uganda context for our analysis, as suggested by Reviewer 2.
4. Why did the authors use principal components analysis instead of confirmatory factor analysis given that the items were expected to fall within particular constructs?

Principle component analysis was applied in order to construct an index measuring the use of routine health management information. It is supposed to be a one dimensional construct. Thus, all items are supposed to fall with in one factor, which is what the analysis showed. The rationale for our decision to use PCA in the analysis is explained above in our responses to Reviewer 1.

The exploratory and confirmatory analyses for other scales, such as the promotion of a culture of information, is beyond this paper and as suggested by the reviewer, will be part of a separate paper that is planned.

5. Were statistical analyses conducted for the test-retest estimates? It is unclear if test-retest was assessed at the group level or the individual level (the correct way, I believe) using paired t-tests or correlations. Please state the method used and argue for its appropriateness.

As described above, we have substantially revised the description of the test-retest analysis, which is now included in the methods section. It is to be noted that although 98 percent of the facilities were revisited but we did not record how many respondents were part of the previous survey. Thus, test and re test participants were not 100 % same. Therefore, calculation a correlation coefficient of reliability using the same respondents was not possible. However, the respondents were not different in their socio-demographic characteristics allowing a comparison in their responses using group mean scores. Pedhauzer et al. (1991) argue that having more than two week interval for test retest reliability is problematic, as it is difficult to separate changes in respondents from reliability (stability) of the scale to real changes in respondents. Cooke and Szumal (1993) have used test retest mean score for assessing the reliability of organizational cultural inventory with a two year difference. Thus, we have used test-retetest mean score of scale between 2004 and 2007 as reflection of reliability as well as sensitivity of the scale. Cronbach’s alpha along with lack of changes in mean score showed that scales were reliable. While internal consistency among various variables was used to validate the scales were sensitive to pick up real change.

The individual responses were combined for creating a mean score on motivation, perceived promotion of culture of information, perceived RHIS task self-efficacy and RHIS task competence. A comparison is made of the mean scores responses between 2004 and 2007 (Table 5) by using two samples t test. We have also used chi square test to show whether diagnostic tool was consistent in measuring in data accuracy over time. This was done at the facility-level.


Cooke RA, Szmul JL (1993), Measuring Normative Beliefs and Shared Behavioral Expectations in Organizations: the Reliability and Validity of the Organizational Culture Inventory, Psychological Reports, 72, 1299-1330.

6. Elements of discussion points are presented in the results – these should be moved to the discussion (e.g., top of page 16).

We have moved the discussions points from the results section to the discussion section. Thank you for this comment.
7. Please discuss the use of the dependent variable and why it is appropriate to use.

The goal of the RHIS is to use routine information for evidence-based decision-making, creating transparency and accountability. Thus, under the PRISM framework, the key “output” of interest is RHIS performance which is composed of two dimensions – data quality and the use of information. In the study, we have chosen to emphasize the determinants of the use of RHIS information. As mentioned in response to earlier comments, we believe our index of RHIS data use from 2007 is substantially improved, as compared to the measure of 2004, as it includes not only indicators of the display of data, but also how RHIS data was used for other purposes. In the attached manuscript, we have revised the description of the dependent variable by mentioning these points.

8. Please provide more specifics regarding the types of interventions suggested by the findings (in the Conclusions section).

We think this is an important suggestion. We now present more specifics regarding the types of interventions suggested by the findings in the conclusions section.

Minor Essential Revisions

1. Please use principal not principle components.

This change was made throughout the paper. Thank you.

2. Please use Likert not likert.

This change was made throughout the paper.

3. Please discuss the use of Uganda as a test country – how typical is it? How can the information be used in that specific country.

As described above in response to reviewer 2, we have now added to the background section a brief discussion on the Ugandan context and how the information can be used by government stakeholders.

Discretionary Revisions

Please provide more descriptive information of the health center in terms of size, location, types of providers, and patients seen.

In the description of the Ugandan context, we now include details on the government-run referral system, which includes information on the types of health centers that are in place in a typical district.