Author's response to reviews

Title: Effectiveness of a peer-led HIV prevention intervention in secondary schools in Rwanda: results from a non-randomized controlled trial

Authors:

Kristien Michielsen (Kristien.Michielsen@ugent.be)
Roxanne Beauclair (Roxanne.beauclair@gmail.com)
Wim Delva (Wim.Delva@ugent.be)
Ronan Van Rossem (Ronan.VanRossem@ugent.be)
Marleen Temmerman (Marleen.Temmerman@ugent.be)

Version: 2 Date: 10 July 2012

Author's response to reviews: see over
From: Kristien Michielsen  
   International Centre for Reproductive Health, Ghent University  
   E-mail: kristien.michielsen@ugent.be

To: BioMed Central Editorial Office

Subject: Resubmission of manuscript “Limited effectiveness of a peer-led HIV prevention intervention in secondary schools in Rwanda: results from a non-randomized controlled trial” (Manuscript reference number: 1435799946697974)

10 July 2012

Dear Editor,

We thank you for your extensive review and constructive comments on our manuscript entitled “Limited effectiveness of a peer-led HIV prevention intervention in secondary schools in Rwanda: results from a non-randomized controlled trial” (Manuscript reference number: 1435799946697974). We have addressed each of the reviewers’ comments point by point and have amended our manuscript accordingly. Please see below a detailed response.

We hope that you will reconsider this submission for publication in BMC Public Health.

Yours faithfully,

Kristien Michielsen
Reviewer #1

This is a thought-provoking paper on the evaluation of a peer education programme for school-going young people in Rwanda. The implementation of the programme (which is often flawed) was done well. The evaluation was well planned and many confounding variables were taken into account. Still, the peer-led intervention had limited effectiveness – when effectiveness was measured using the conventional indicators: knowledge, attitudes and behaviour.

The paper was written in a critical style. The procedures used were described as it happened, and critically reviewed. The evaluation was well-planned. The sample was determined through careful calculations. The evaluation included three measurements and various variables. The inclusion of a participation score is a benefit – since previous research often blamed non-effectiveness to low levels of participation. This was also true in this research, but level of participation did not determine the outcomes.

The results were described well and the discussion raised many issues that need to be taken into account when planning, implementing and evaluating peer-led interventions, as well as other HIV prevention interventions. I think this paper provides a good evaluation and a summary of thoughts about peer-led HIV interventions that could influence future HIV prevention planning. I would really like to see this published.

We thank the reviewer for this positive feedback.

Minor essential revisions. Some attention can be given to the following:

1. Is it standard practice or ethical to give the names of the participating schools?

Even though the results are aggregated and cannot be linked to the individual schools, we do share the reviewers’ concern for anonymity and replaced the schools’ names by numbers.

Table 1 now reads:

Table 1. Characteristics of participating schools

<table>
<thead>
<tr>
<th>Name of the school</th>
<th>Location</th>
<th>Lower/higher secondary education</th>
<th>Number of students (2009)</th>
<th>Public or private</th>
<th>Religious background</th>
</tr>
</thead>
<tbody>
<tr>
<td>School 1</td>
<td>Bugesera (I)</td>
<td>Urban lower + higher</td>
<td>&gt;750</td>
<td>private</td>
<td>mixed</td>
</tr>
<tr>
<td>School 2</td>
<td>Bugesera (I)</td>
<td>Rural lower + higher</td>
<td>&gt;750</td>
<td>private</td>
<td>Catholic</td>
</tr>
<tr>
<td>School 3</td>
<td>Bugesera (I)</td>
<td>Rural lower</td>
<td>251-500</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 4</td>
<td>Bugesera (I)</td>
<td>Rural lower</td>
<td>251-500</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 5</td>
<td>Bugesera (I)</td>
<td>Urban lower</td>
<td>501-750</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 6</td>
<td>Bugesera (I)</td>
<td>Rural lower + higher</td>
<td>&lt;250</td>
<td>private</td>
<td>Islamic</td>
</tr>
<tr>
<td>School 7</td>
<td>Bugesera (I)</td>
<td>Urban higher</td>
<td>501-750</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 8</td>
<td>Bugesera (I)</td>
<td>Rural lower + higher</td>
<td>&gt;750</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 9</td>
<td>Rwamagana (C)</td>
<td>Rural lower + higher</td>
<td>501-750</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 10</td>
<td>Rwamagana (C)</td>
<td>Rural lower + higher</td>
<td>501-750</td>
<td>public</td>
<td>mixed</td>
</tr>
<tr>
<td>School 11</td>
<td>Rwamagana (C)</td>
<td>Rural lower + higher</td>
<td>251-500</td>
<td>public</td>
<td>mixed</td>
</tr>
</tbody>
</table>
2. The alpha of severity is described as relative low – it is very low and actually not acceptable to use in research (0.28!!). There are two other scales with scores round about 0.5 which is also relatively low.

We are aware that a Cronbach’s alpha of 0.28 is very low, even for a four-item scale. We deleted this scale from the analysis and replaced it with one measure of perceived severity. The manuscript was changed into:

“Attitudes towards HIV/AIDS were divided into three attitudinal constructs. Perceived susceptibility to HIV (range 0-16, Cronbach’s alpha 0.67) was measured using the validated 4-items scales of Lux & Petosa [1]. Perceived severity of HIV was measured in one single item ‘At present, the danger of AIDS has almost passed’. Enacted stigma was measured by two items asking if the students would refuse to be taught by a HIV-positive teacher and if they thought HIV-positive students should be expelled from school (scale 0-1, Cronbach’s alpha 0.76).”

The sexual self-concept scale was developed by Winter (1988) and was found to be reliable in different populations of adolescents (Pearson, 2012; Deardorf, 2008). Nevertheless the Cronbach’s alpha was relatively low (0.52). After further analysis, we found that removing the last item of the scale would increase the Cronbach’s alpha to 0.6. This is still not very high, but acceptable since the scale on which it was based was already validated. The manuscript was changed into:

“3) sexual self-concept using an adapted scale made up of 13 5-point Likert items, with scores ranging from 13 to 52 (Cronbach’s alpha of 0.60) [2].”

Knowledge of HIV protection modes was measured by presenting the respondents with 11 true or false ways of protecting oneself against HIV infection, and asking them to indicate the correct ways. We mistakenly reported this measure as a scale (ranging from 0-11 with a Cronbach’s alpha of 0.58), while it is actually a measure of knowledge and should not be considered a scale. We changed the manuscript to:

“Knowledge of HIV protection modes was measured with 11 items proposing true and false statements about ways of protection, which were consequently summed to get a score from 0 (all answers incorrect) to 11 (all answers correct).”

3. A more extensive explanation of propensity score can help the reader to understand how that was used to deal with differences at baseline.

We added the following explanation to the manuscript:

“Propensity scores are used to reduce selection bias when assignment to study arms is not randomized. A participant’s propensity score is its probability of being assigned to a specific study arm given a set of known covariates.” [3]
4. Comment on the complexity of data – such a differences at baseline, change in both groups and how levels of change can be compared.

We added the following paragraph to the manuscript:

“Marginal linear and logistic regression analyses, using Generalized Estimating Equations (GEE), were conducted to determine the likelihood of experiencing different outcomes based on which group the participant belonged to, while accommodating for repeated, correlated measures. Specifically and most importantly for this analysis, these models allow for investigation of group effects, time effects, and group by time interactions. Our analysis specified an unstructured correlation matrix and a binomial or Gaussian distribution depending on which dependent variable was analyzed.”
Reviewer #2

Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

The study is a non-randomised controlled evaluation of a peer-led HIV prevention programme in secondary schools in Rwanda. It is good to see peer-education evaluations from Sub-Saharan Africa. I feel that the study has merit and have made suggestions about how the manuscript could be improved. Although I know it looks daunting I do think that the manuscript would benefit from these changes/additions.

We thank the reviewer for sharing her experience in evaluating peer education interventions and publishing on this topic. We are convinced that the integration of these comments has significantly improved our manuscript.

1. Rwanda has instituted anti-AIDS clubs in secondary schools and 98% have installed such clubs. The authors say however that such clubs often remain inactive. How did this affect the intervention? For example, how active were these clubs in the intervention schools?

While the Rwandan Government obliged all schools to install an anti-AIDS-club, no government means were provided for their functioning. In practice, this means that a small number of students (mostly between ten and twenty) form an anti-AIDS-club, which gathers regularly (e.g. every fortnight or every month). However, since they are not trained in matters related to HIV/AIDS or adolescent sexuality, and do not have financial or material means, they only internally discuss these topics and do not organize activities for other students.

In Rwanda, the CDLS (Comité de District de Lutte contre le SIDA – District AIDS Control Commission) coordinates HIV-related activities of governmental and non-governmental organizations, in order to avoid duplicity. In the two districts included in our study we were informed that no other organizations were working with the anti-AIDS-clubs in the schools. Before the start of our study, the anti-AIDS-clubs in the schools involved in our study (both intervention and control) therefore remained inactive because of lack of training and support.

In the manuscript, we specifically state that the anti-AIDS-clubs are activated:

“The general objective of the peer education program was to reduce sexual risk behavior and to promote sexual and reproductive health in the secondary school communities by activating the anti-AIDS-clubs in the schools.”

2. How were the schools selected to receive extra support for the clubs. The authors say that ‘in most cases’ this support consisted of a peer education programme. What happened in other schools?

The Peer Education intervention in secondary schools in Bugesera was part of the larger “Integrated Health Program” of the Rwandan Red Cross. The district of Bugesera was selected to implement this program because of its low scores on general health indicators (as malaria, access to drinkable water, access to primary health care). Within this district, all secondary schools received support for the anti-AIDS-clubs using peer education, and no selection of schools was made.
When stating ‘in most cases’ we referred to other non-governmental organizations who may support anti-AIDS-clubs in schools in other districts, through peer education or other support (e.g. distribution of booklets).

3. Is the question posed by the authors well defined? The authors evaluate the HIV prevention peer education programme implemented by the Rwandan Red Cross society in Bugesera. Later on under ‘study design and sample size’ the authors mention that the primary objectives were to assess whether the intervention influenced the time trend in condom use and recent history of sexual intercourse. These primary objectives need to be stated at the end of the background section also.

The main objective of the Peer Education intervention was to reduce sexual risk behavior and to promote sexual and reproductive health in the secondary school. While the intervention equally focused on reducing sexual risk behavior (condom use and recent sexual activity), increasing knowledge and reducing stigma, sample size was expected to be largest for measuring changes in sexual behavior and condom use, since their prevalence was expected to be the lowest. See also response to Comment 10.

At the end of the background section we completed the sentence “This paper evaluates the effectiveness of the HIV prevention peer education intervention implemented by the Rwandan Red Cross Society in the district of Bugesera” with “in increasing HIV related knowledge, reducing sexual risk behaviors and changing attitudes”.

In the section ‘study design and sample size’ we adapted the manuscript as follows: “We based the sample size calculation on the study’s objective of assessing whether or not the intervention influenced the time trend in condom use and recent history of sexual intercourse.”

4. Are the methods appropriate and well described? In the first paragraph there is no mention of the theoretical background employed in the development of the programme. Can the authors say a little more about this?

The expert centre that gave input in intervention development and in the training of trainers used an integrated theoretical model based on a number of behavioral theories: Theory of Reasoned Action, Social Learning Theory, Diffusion of Innovations Theory and Health Belief Model. For its delivery, the intervention was based on the Theory of Participatory Education.

In the Methods section, we added: “The intervention was based on an integrated theoretical framework that included aspects of the Theory of Reasoned Action, the Social Learning Theory, the Diffusion of Innovations Theory, and the Health Belief Model. The intervention delivery was based on participatory learning techniques”.

The questionnaire included constructs of the Theory of Reasoned Action, the Health Belief Model and the Social Cognitive Theory. Separate analyses are being done on the extent to which these constructs explain variety in sexual behavior.

5. What did the 6-day training consist of?

The training consisted of four main parts:
- background on the Red Cross and its main principles
- proving information on HIV/AIDS, sexually transmitted diseases, family planning and pregnancies
- the role of the peer educator: what is expected of a peer educator and what is the deontology of a peer educator?
- teaching/communication methods: how to best approach students and how to transmit messages and counsel?

We added the following sentence to the manuscript:

“The training consisted of information on the Red Cross and its main principles, HIV/AIDS, sexually transmitted diseases, family planning and pregnancies, the role of the peer educator (what is expected of a peer educator and what is the deontology of a peer educator?) and teaching and communication methods (how to best approach students and how to transmit messages and counsel?).”

6. How were the peer educators chosen? Were they volunteers, did teachers or some other adult choose them? Were they credible sources of information? This is important (Mason-Jones, Flisher and Mathews 2011)

Selection criteria for peer educators were developed:

- personal characteristics: peer educators should be open, approachable, confidential, take initiative, have integrity, have leadership capacity, and have exemplary behavior
- half of the peer educators have to be girls
- peer educators have to be in the 2nd, 4th or 5th year of secondary education, to continue activities over several years (drop-out is the largest after the 3rd year, being the end of lower secondary education)

These criteria were given to the ‘prefet de discipline’ (disciplinary teacher) who lives with the students and knows them well, and who is asked to make a selection of peer educators.

We added the following sentence to the manuscript:

“The peer educators were selected by the disciplinary teacher, who lives in the school and knows the students well, based on a number of predefined criteria (personal characteristics, sex, study year).”

7. The authors say that the peer educators were tasked with teaching their fellow students how to adopt positive and responsible behaviours towards HIV/AIDS through group and individual counselling, drama songs and other interactive methods. Did the training include counselling skills? What were the ‘positive and responsible behaviours?’ You need to describe this in more detail.

The training did include workshops on how to approach students, how to address sensitive topics, and included role plays on counseling (see response to comment 5). Behaviors and attitudes that were emphasized were mutual respect within relationships, personal responsibility for protective behavior both for boys and girls and non-stigmatizing attitudes towards sex before marriage and hiv-positive persons.
In the manuscript, we added:

“During the course of the intervention, the peer educators were tasked with teaching their fellow students how to adopt positive and responsible behaviors towards HIV/AIDS, such as respect within relationships and personal responsibility for protective behavior.”

8. The intervention ran for one year (August 2009 to November 2010). What activities took place during this time in the intervention schools? (Fidelity monitoring)

The peer educators were expected to report each trimester on the activities they organized. These reports detailed the types of activity (e.g. theatre, song, discussion, teaching), the topic discussed (e.g. HIV prevention, anatomy, pregnancy prevention) and the number of students who attended the activity. These reports indicate that a part of the activities were organized only for the members of the anti-AIDS-club (sometimes with extension to the Red Cross club). Other students were reached by activities as songs, dances, theatres and condom demonstrations. Furthermore, peer educators were available for counseling at all times. They were introduced to the other students at the beginning of the intervention, and could be identified through a special t-shirt.

The process analysis [4] indicated that the activities organized by the peer educators can be roughly divided into four types:

- meetings with the club members
- events for larger groups of students or for the whole school, such as theatre plays or competitions
- spontaneous personal and group discussions or conferences
- field visits.

The most frequently used technique were conferences and group discussions. In this case, the peer educator prepared a certain topic and presents this to the students. These conferences could take place in the anti-AIDS-club, but the peer educators also visited classes to present a topic. After the peer educator has given the explanation, students were given the opportunity to ask questions.

The meetings with the club members took place once or twice a week during leisure times. The clubs generally had a fixed day in the week to meet. During these meetings they discussed themes such as reproductive health and HIV/AIDS, or the working of the Red Cross.

The bigger events were organized for a larger public of students. They took place on average once or twice a semester. These activities differed greatly in techniques used and interactive participation of the students. The different techniques used were theatre plays, songs and dance, poems, games and competitions. Theatres and sketches were popular means to organize prevention activities amongst both peer educators and students. The peer educators tried to present their activities in the form of leisure activities to attract students. Theatres were popular amongst the students if they contained enough humor and drama, and they were good means to involve other students in the preparations.

We observed an evolution in the frequency of activities during the course of the program. There seemed to be substantial decline in the frequency of activities organized in the last part of 2010. This was observed both in the activity reports as in interviews and focus group discussions. Moreover not only the amount of activities within the schools declined. In the beginning of the program there were
interschool activities, regular meetings between peer educators, and visits to the schools by the Red Cross. Although they were planned, such interventions were not organized in the second part of the period. The content of the activities, however, was stable.

In the section ‘Factors associated with the implementation of the intervention’ we added the following sentence:

“Consequently, in the second part of the intervention, we observed a reduction in the number of activities organized by the peer educators in all schools.”

9. Under study design and sample size, there is no mention of how the control schools were selected. Were they matched with the intervention schools? Did they also complete the survey at baseline, T1 and T2?

The control schools were selected the same way as the intervention school. We started with the inclusion of school who offered higher secondary education (5 in the intervention district, 4 in the control district), since they include older youth for which the intervention might be more important. Following, we tried to add variety in the sample by including schools from (in this order) different setting, different religious background, different size and different funding. All students, both in the intervention and control group, completed the survey at baseline, T1 and T2.

In the manuscript, we changed the paragraph on school selection into:

“Schools were selected on a purposive basis. We aimed to include the greatest variety of schools in the study and applied several selection criteria: education offered (lower/higher secondary education), location (urban/rural), religious background, number of students (small/large) and funding (public/private). All selected schools were willing to participate in the study. Since no roads directly connect intervention and control sites, cross-site contamination was unlikely.”

10. Was the sample size adequate? I suggest that this needs to be looked at by a statistical expert. Are you saying that your assumptions were 30% baseline condom use? Did you account for the intra-cluster correlation coefficient within the schools? You say under ‘statistical analysis’ that the correlation between students within schools was ignored as it was weak and non-significant. I think this needs a more robust explanation (and possibly a post hoc power calculation)

Our sample size calculation for the study protocol and the post hoc power calculation were done by a statistician. Adjusting for intra-cluster correlation was considered for the sample size calculation included in our protocol, but was omitted in post hoc power calculations because of weak, non-significant correlations between students in each of the school. However, we did find that there was a stronger within participant correlation for the repeated measures, so we factored in a design effect of 2 to take this into account.

As stated in the manuscript, for the logistic regression models, to detect an adjusted odds ratio of 2 or more with 80% power at the 0.05 significance level, under conservative assumptions of 30% baseline prevalence of the outcome variable and no changes over time in the control group, we needed to obtain 1241 observations. For linear regression models, a minimum of 348 observations were required to detect a small standardized effect size (Cohen’s d) of 0.3 with 80% power at the
0.05 significance level. You are indeed correct in inferring that we thought a 30% baseline prevalence of our binary dependent variables, including condom use, would be a conservative estimate.

We calculated that information for all time points (T0/T1/T2) from at least 1241 participants were necessary. Assuming 25% loss-to-follow-up over the course of the study, the target sample size at T0 was 1655. At baseline we had 1950 participants, and the total retention rate after the third survey (T2) at 18 months was 71.8% (1400/1950), thus meeting the sample size estimated prior to the commencement of the study.

11. How were the 8 schools selected out of the 15? Did you randomly select the schools in a stratified sample of urban/rural, public/private, small/large number of students etc?

See comment 9.

12. Why did you select student in 2nd and 5th year of school? What about 3rd year?

The study ran over more than one academic year (2009 and 2010). Since we preferred to collect longitudinal data, we wanted to make sure that most students who took part in the baseline survey would still be in school at T1 and T2. The lower secondary education stops after the 3rd year, with a consequence that many students either stop school or change to another school (especially when the school only offers lower secondary education). Therefore we only included students from the 2nd and the 5th year.

In the manuscript this is stated as follows:

“The study targeted all students who were in their second and fifth year of school, since they had a higher chance of still being in school at the end of the survey. Drop-out rates are highest after the third year (the end of lower secondary education) and of course the sixth and final year.”

13. ‘Coding system guaranteeing confidentiality’ needs a little more explanation.

The students needed to complete an informed consent form before the start of each survey, which stated their names and a number. The same number figured on the questionnaire they completed. The informed consent form information was stored in a separate database, containing name, date of birth and questionnaire number. Another database contained the questionnaire number and answers, but no name or date of birth. The numbers allowed us to merge answers from the same respondent, while the identifying information and the questionnaire answered did not figure in the same dataset at any point in time.

We changed the manuscripts to:

“A coding system guaranteeing confidentiality, separately storing identifying information and questionnaire answers, was used to match students over the three waves.”

14. You say that all schools agreed to be involved. How did you approach the schools?

Schools were first selected, based on the criteria described above. Then two representatives of the study team (the main researcher and the district coordinator of the Rwandan Red Cross) visited the schools to explain the intervention and the study and to explain what was expected from the schools.
15. Did parents give informed consent for their child to be involved in the research study? Did the students give written informed consent?

The respondents provided written informed consent at baseline, T1 and T2, before starting the completion of the questionnaire.

Even though the legal age of majority is 21 years in Rwanda, and our study included participants from the age of 14 years onwards, we did not ask parental consent. The first submission of our research project to the Rwandan National Ethical Committee was not approved for that reason. Argumentation in the second submission convinced the Committee members that parental consent could be waived. Two main arguments were given. First, practical considerations: many students live in boarding schools and return home only two or three times a year. Parents rarely come to the school. The school does not possess all addresses or phone numbers of the parents, making it impossible to visit them or contact them by letter or phone. Low literacy rates and a low number of phone ownership was an additional problem. Second, we argued for a more developmental approach to adolescence and adulthood and stressed the importance of collecting data directly from adolescents. Based on guidelines from the World Health Organization [5], the Society for Adolescent Medicine [6], the American Sociological Association [7] and scientific literature [8, 9] we argued that adolescents are capable of understanding studies and have the cognitive capacity to take decision concerning participation and that the generally accepted age limit for this capacity is 14 years.

In the manuscript we added the word ‘written’ to the following sentence:

“Before the start of the survey, the students gave written informed consent after concepts of voluntary participation and confidentiality were explained to them.”

16. More detail is needed on the instruments used (surveys) for collecting the data. Why were the questions chosen? Were these related to the theoretical basis of the programme?

The questionnaire was developed mainly using existing, validated questions and scales [1, 2, 10, 11]. The items were chosen to measure the specific objectives of the intervention (for further information on the items used, see response to comment 19). The questionnaire was pretested for clarity on a group of 30 students from the first year of secondary education.

While the intervention was based on an integrated model based on a large number of behavioral theories, we measured the theoretical constructs of only three theories in our questionnaire (Theory of Reasoned Action, the Health Belief Model and the Social Cognitive Theory - an analysis is being done on the extent to which these constructs actually explain variation in sexual behavior). Including more constructs would have made the questionnaire too long and complex.

17. The statistical analysis needs to be explained in more detail. Particularly relating to the use of a propensity scoring model. If the propensity model misses an important reason why subjects are selected to treatment or control there may be a problem. I suggest a statistical expert. Eg. Dr Thomas Love, Case Western Reserve University (thomaslove@case.edu). Also the model includes variables that are primary outcomes (ever had sex, had sex in last 6 months, used a condom at last intercourse). This needs to be explored further as to whether this is appropriate.
We added the following explanation of propensity scores to the methods section, and cited an additional key paper on this statistical method:

“Propensity scores are used to reduce selection bias when assignment to study arms is not randomized. A participant’s propensity score is its probability of being assigned to a specific study arm given a set of known covariates [3].”

You are correct in your suggestion that failing to miss an important variable in the propensity score may be problematic. However, this is no more so problematic than failing to adjust for any potential confounder variable in analyses of any observational study. Unfortunately, almost every study ever conducted (including RCTs that break randomization after baseline) fails to do this. The authors of this paper, in consultation with a statistician, are confident they chose a range of variables for the propensity score which might explain the selection bias. They were chosen because they had significant differences between participants in intervention and control groups at baseline. Essentially, creating the propensity score is the same as adjusting for all eight of those variables, but reducing it to one parameter in the model. It allows us to achieve greater statistical efficiency, i.e. greater power to detect the effects of primary interest.

Additionally, we are aware that our propensity score contains variables that are also dependent variables in our models. Importantly, however, the propensity score only contains baseline information, whereas the main focus of the analysis lies on the time-trend and group-time interactions, i.e. how outcomes change from T0 to T1 and T2. In a way, our approach of adjusting for propensity scores which include baseline information is similar in spirit to what has been dubbed a difference-in-differences analysis: In this type of analysis, treatment effect sizes are estimated, adjusting explicitly for differences between treatment arms at baseline.

Finally, as per your suggestion, we further explained how the models were constructed to take into account the complex longitudinal data. We added a description of the Generalized Estimation Equations method of modelling:

“Marginal linear and logistic regression analyses, using Generalized Estimating Equations (GEE), were conducted to determine the likelihood of experiencing different outcomes based on which group the participant belonged to, while accommodating for repeated, correlated measures. Specifically and most importantly for this analysis, these models allow for investigation of group effects, time effects, and group by time interactions. Our analysis specified an unstructured correlation matrix and a binomial or Gaussian distribution depending on which dependent variable was analysed.”

18. Good that you used a scale for programme involvement. This is very useful. 19. There is not enough detail about the instrument used for measuring variables and whether they were validated for this population. How many items were there in total and what scales were used? You have mentioned some in this first section and then later on. It would be useful to describe the scales in one place in the methods. More detail on the scales. Often Cronbachs alpha was quite low (0.58 for Knowledge of HIV protection modes health locus of control, 0.59, sexual self concept 0.52), or very low, (0.28 for severity of HIV). Usually a Cronbachs alpha of over 0.70 is considered a ‘good’ scale (Bland and Altman 1997).

As described in the response to comment 16, the questionnaire was largely developed using existing, validated questions and scales [1, 2, 10, 11]. The items were chosen to measure the specific objectives of the intervention. Not all items were validated for the specific population of school-going youth in Rwanda, which may be an explanation for the relatively low Cronbach’s alpha’s.
The questionnaire consisted of items measuring socio-demographic characteristics (8 items), knowledge on HIV transmission, protection modes and sexual and reproductive health (18 items), sexual behavior (30 items, not all had to be completed by all respondents) and items measuring theoretical determinants of sexual behavior (79). The questionnaire took approximately 1 hour to complete.

For a response to the relatively Cronbach’s alpha we refer you to our response to comment 2 of reviewer 1.

Results

20. Why was attrition more in the intervention group? Did you do a sensitivity analysis to find out if this was related to any of the outcomes of interest?

We compared baseline characteristics of students who did not participate in the study at T1 or T2 (see table below), to see if there were large difference that may have meaningfully influenced our analysis. For most variables, we found only small differences between trop-outs and those retained in the study.

One variable showed a large difference: sex. In the control area, more girls dropped-out than boys. This can be traced back to one school, a private girls-only school. In this school, we were confronted with a 53% drop-out of participants, all having left school (on average 17.7% of all students had left their school between baseline and T1). This school had to reduce the number of classes drastically: in the second year of secondary education (where we did the baseline) there were three classes. This was reduced to one single class the next year at T1. Young people prefer to go to public schools, since these are for free. The school concerned was a rather expensive school, which might explain why so many girls left.

Since sex was included in the propensity score, our main analysis controlled for this difference.
Baseline characteristics for those retained in the study and those lost to follow according to study site and time of measurement

<table>
<thead>
<tr>
<th></th>
<th>Intervention group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td></td>
<td>Retained</td>
<td>Lost-to-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Follow-up</td>
</tr>
<tr>
<td>Sex n (%)</td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td></td>
<td>414 (51.43)</td>
<td>129 (48.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Mean (sd)</td>
<td>18.35 (2.14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.01 (2.39)</td>
</tr>
<tr>
<td>Ever tested for HIV n</td>
<td>Yes</td>
<td>418 (55.73)</td>
</tr>
<tr>
<td></td>
<td>14 (1.87)</td>
<td>5 (1.98)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 (0.83)</td>
</tr>
<tr>
<td></td>
<td>Socio-economic status n (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>263 (32.79)</td>
</tr>
<tr>
<td></td>
<td>22 (2.74)</td>
<td>16 (6.11)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Live during the year n</td>
<td>Boarding school,</td>
<td>150 (19.58)</td>
</tr>
<tr>
<td></td>
<td>off campus</td>
<td>21 (2.77)</td>
</tr>
<tr>
<td></td>
<td>Parents/ Family</td>
<td>313 (40.86)</td>
</tr>
<tr>
<td></td>
<td>336 (45.65)</td>
<td>14 (10.14)</td>
</tr>
<tr>
<td>Baseline HIV knowledge</td>
<td>Mean (sd)</td>
<td>5.92 (1.87)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retained</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.65 (2.02)</td>
</tr>
<tr>
<td>Ever had sex n (%)</td>
<td>Yes</td>
<td>216 (27.38)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retained</td>
</tr>
<tr>
<td></td>
<td></td>
<td>165 (22.48)</td>
</tr>
</tbody>
</table>
21. There was significant baseline imbalance suggesting that the intervention and control school students were different. Was this related to how they were selected?

Since the intervention was planned to take place in all secondary schools in the intervention district, we could not randomize the students to the intervention or control group. Given the difference existing between the intervention and control district (e.g. the control district had a higher socio-economic status) it is possible that the baseline imbalance was related to the selection procedure. To counter this imbalance, we controlled our analyses for these baseline differences in the propensity score.

22. Why was the knowledge significantly different in the control group after the intervention?

We did not find a sound explanation for this observation. We added this in the discussion session.

23. You also mention that the intervention students were significantly more likely to think of HIV as a severe disease but as mentioned previously, the scale may not very good measure of perceived severity in this population.

In the revised version, we have adapted the measure of perceived severity, and reduced it to one single measure (see response to comment 2 of reviewer 1). Consequently, we did no longer find a significant change in this variable. Reports on the significant increase in perceived severity score of intervention students were left out of the manuscript.

24. You mention in the discussion that schools in both groups were asked if there were other activities taking place but do not mention this anywhere in the text earlier on. This is important to include.

In the section ‘School and participant selection’, we added the following sentence to our manuscript:

“In the schools included in our study, no other organizations were implementing activities in the field of HIV/AIDS or sexual and reproductive health.”

25. Are the data sound? As only 12-14% of students actually participated in the intervention and some students were excluded from the analysis do you think your analysis is sound?

Even though only a small number of students said to have actively participated in the intervention (13.7% and 11.6% of students had a participation score above six at T1 and T2 respectively), about half of the respondents (56.6% and 43.5% at T1 and T2 respectively) participated in one or more activities. Furthermore, through diffusion of information, the messages spread by the peer educators, may well have reached more young people. The first analyses compared intervention with control schools, while an additional analysis compared participants with different levels of participation.

We were aware of the potential selection bias, hence adapted our analysis accordingly. The exclusion of students who fell out of the area of common support in the propensity score analysis was done to make the data more sound for the main analysis.

26. You imply that participation in activities was not necessary or meaningful. The discussion then does suggest that the peer education activities and the way they were implemented were not
necessarily acceptable to the students. Can you give more detail about this. You reference the dissertation published from this study but can you say something here which gives some insight into why this may be?

We apologize that it seemed we applied that participation in activities was not necessary or meaningful. We meant to say that more active participation in the intervention did not result in larger changes in sexual behavior, as was demonstrated by the dose-effect analysis.

Neither did we want to imply that the activities or the way they were implemented were not acceptable to the student. In the section ‘Factors associated with the implementation of the intervention’ we do say that, due to internal organizational problems, the intervention was not implemented as designed in the second part of the intervention. See also our response to comment 8.

We added the following section to the manuscript:

“Consequently, in the second part of the intervention, we observed a reduction in the number of activities organized by the peer educators in all schools.”

27. How can you suggest that the peer education programme could have succeeded in creating a communication climate? You don’t give any other evidence for this.

We agree with the reviewer that ‘communication climate’ might not be the best wording. Given the reduction in stigmatizing attitude, we wanted to say that the peer education intervention might have created a more positive attitude, a more positive climate around HIV. In our manuscript we changed the words “communication climate” into “positive, less-stigmatizing climate”.

Discussion

28. Before going on to talk about reasons for limited effectiveness I think it is important for you to provide details of the limitations of the study itself. How the groups were selected is one issue that needs to be addressed, especially as they are clearly different at baseline. Factors associated with the implementation of the intervention

We have placed the limitations section before the discussion of possible reasons for limited effectiveness.

29. This section needs to be stronger and a little more detailed. Was the implementation problematic in all schools?

Yes, as discussed in our responses to comments 8 and 26, due to internal organizational problems, the implementation of the intervention did not take place as planned. This was the case for all intervention schools.

In the manuscript, we added the following sentence:

“Consequently, in the second part of the intervention, we observed a reduction in the number of activities organized by the peer educators in all schools.”
30. Were some peer educators followed up and not others? You mention ‘limited monitoring and follow up’ but don’t give very much detail.

The monitoring plan of the intervention planned for trimestral visits of the coordinator to all schools. This was followed in the first half of the intervention. However, due to internal organizational issues, in the second part of the interventions, the coordinator could not visit the schools. All peer educators had to hand in activity reports to the coordinator. We found that, in all schools, the number of activities reduced in the second part of the intervention. Furthermore, they all asked for additional support of the Red Cross coordinator.

In the section describing the intervention, we added the following sentence:

“Peer educators had to hand in an activity report each trimester.”

In the section ‘Factors associated with the implementation of the intervention’ we added the following sentence:

“In the activity reports of the second part of the intervention, more peer educators requested additional support of the intervention coordinator.”

31. You also mention a ‘second round’ of training here but not in the methods section. Just make sure there is consistency throughout in reporting the study.

The manuscripts was revised for internal consistency. We added the following sentence to the section describing the intervention:

“Additional trainings were planned in the second part of the intervention, but did not take place (see Discussion section).”

32. You say that the intervention was ‘thoroughly developed’ but don’t give details about this in the methods section. What are ‘validated manuals’ for example? What does this mean?

In the methods section, we added the sentence:

“The intervention development was the result of a collaboration between the Rwandan Red Cross and two expert organizations in the field of HIV and sexual and reproductive health, and was based on a number of peer education manuals [12-14].”

In the discussion section, we replaced ‘validated manuals’ by ‘peer education manuals’.

33. Again you don’t mention the theoretical basis of the programme (if indeed there was one at all). What was the theory behind the design of the intervention?

See response to comment 4.

34. You mention that the methodologies were limited. However you don’t give enough detail as to why information giving, counselling, drama, theatre or song are not appropriate. I suggest adding references here.
The methodologies used in the intervention in our study mainly focused on transferring knowledge to the target population. This is done through different methods, as songs, drama, etc. However, the main focus is on transferring knowledge on HIV and sexual and reproductive health to the students.

In the manuscripts we completed the sentence “It has been amply demonstrated and discussed that increasing knowledge alone will not change sexual behavior [15-17]” with “, since sexual behavior is also determined by a number of other factors.”

35. You also start mentioning use of health services here which was presumably not part of the evaluation at all. Maybe just stick to aspects that were actually part of the intervention.

We mentioned health services as an example of why the intervention focus was, in our view, too narrow, and hence as a possible explanation of the observed limited effectiveness of the intervention. It would have been beneficial if participants with a sexually transmitted infection, or an unintended pregnancy, could have been referred to a youth-friendly health service, which was not the case in the intervention.

In the manuscript, we have deleted the sentence: “We cannot expect young people to seek health care when having sexual health problems if services are stigmatizing.”

36. You also mention that peer education programmes should be seen in context such as contributing to a change in communication and breaking taboos. This is not supported by evidence or data from the study.

See response to comment 27. In the manuscript we changed the words “communication climate” into “positive, less-stigmatizing climate”.


Rwandan school-going youth often live in boarding schools. Except for their teacher, they rarely have contact with other adults. Furthermore, due to the taboo on sexuality of adolescents, it is difficult for them to talk about these issue with adults. Consequently, they often rely on each other for information on these topics. Hence, we find that, intuitively, educating young people to spread correct information on HIV and sexual and reproductive health, seems like a good strategy. We changed the manuscript into:

“Since many Rwandan school-going youth stay in a boarding school and only return to their families two or three times a year, they have no other option than to rely on peers or teachers for HIV/SRH information. However, this does not mean they want to rely on them.”

38. In the section above you also start to introduce data and a table about preferred sources of information that was not presented in the results. It needs to go in the results if you want to discuss it further here.

We are aware that we are introducing new information in the discussion section. However, we prefer having a discussion in which a limited amount of new information is introduced, but which tries to give a complete and honest explanation for the observed limited effectiveness of the evaluated intervention. We remain confident that this is the best place to present these data. However, we
diminished the importance given to these new data by deleting the table and summarizing them in one sentence:

“In our study students were asked to indicate the two main channels through which they would prefer to receive information on HIV: friends ranked sixth as a preferred source of information, after radio, parents, television, teachers and medical experts (doctors/nurses).”

39. You need to explain more about the ‘peer education approach’ and why it is used, where it comes from, what the theoretical basis for it is. You do include some of these issues but this section needs more work.

We included extra information in the manuscript to give more insight in the peer education approach:

“Peer education implies that certain members of a group (peer educators) can be influential in convincing their peers to change their behavior. The strategy has proven successful in other fields of health promotion as drug or alcohol use (e.g. [18, 19]. However, when it comes to HIV prevention among young people, not disregarding the capacities they have, it is a very tall order to expect a young person – possibly discovering his/her sexuality him/herself - to act as an expert and guide, counsel, teach and advise peers on a personal, sensitive and complex issue as sexuality.”

40. Factors associated with the evaluation of the intervention. This raises issues related the sexual health research on adolescents more generally and ethical and adult oriented judgments about ‘what is the correct behaviour’ and therefore how to measure appropriate outcomes. Whilst a valid and very interesting discussion, how it relates to the data is not totally clear.

In this section, we discuss the possibility that the observed limited effectiveness of the evaluated intervention can be linked to the indicators used to evaluate it. Since the objective of the intervention was to reduce sexual risk behavior, we discuss that the indicators we used to measure this, might not be adequate to measure the actual sexual risks taken.

We added the following sentences:

“An indicator that would appropriately measure sexual risk behavior should include aspects of exposure (relationship and partner characteristics), transmission (type of sex and protective measures) and infectiousness (HIV infection and stage of infection of the partners). The development such contextualized, composite sexual behavioral measures is essential to measure the real risks young people are taking, hence to determine the real effectiveness of HIV prevention interventions.”

41. I don’t think that you can say that the study has illustrated limited effectiveness of peer education. Whilst I’m not convinced about the merits of peer education as it is currently conceived, your study has many limitations which mean that you need to be much more circumspect in your conclusions.

In the revised version of the manuscript, we speak of ‘observed limited effectiveness’, indicating that however we have observed limited effectiveness, it might still be possible that the intervention was effective, but that our study was subject to limitations making it impossible to observe this effectiveness. In the discussion section, we then further elaborate on these limitations and other
factors (such as the indicators used) that might influence the fact that we observe limited effectiveness.

42. Were peer educators in this study aware of key specialists and services? You mention it here but nowhere else in the manuscript.

The Peer Educators were informed about the health centres operational in the vicinity of their schools. However, collaboration between the peer education intervention and health facilities was not formal.

43. You say “Peer education should not be a stand-alone intervention and should be embedded in a larger strategy” but don’t give enough detail as to why this might be.

As was discussed in our response to comment 35, we argue that only spreading information will not result in the desired changes in sexual behavior, given the numerous other factors influencing sexual decisions.

44. The next section on the involvement of young people needs re-writing.

The section was rewritten:

“Third, peer education interventions seemingly actively involve young people in the prevention efforts. However, true participation goes further than only involving them in the implementation phase. Young people can participate and provide their input in the needs assessment, intervention development, the follow-up and monitoring and evaluation. Their input in these phases might be more valuable than in the implementation of the intervention.”

45. You say “This distance between prevention researchers and prevention implementers hinders the development of novel interventions” but provide neither explanation nor references for this assertion.

We deleted this sentence from the manuscript.

46. You mention sexual competence only in the second to last sentence. Again this needs to be mentioned somewhere in the introduction and discussed more thoroughly.

In the revised version of the manuscript, we discuss that HIV prevention interventions for young people should focus more on the process of decision making, than on the behavioral outcomes as such. As an example, we explain the concept of sexual competence, thereby reducing the weight that was given to this concept in the first version of the manuscript:

“Finally, we must be clear about what we really want to accomplish with HIV prevention interventions. Focusing narrowly on ‘sexual activity’ and ‘condom use’ is not optimal, since these behaviors are only the result of a decision process influenced by many factors that are inadequately addressed by current peer education interventions. In our view, it would be more useful to focus on the decision-making process itself and empower young people to make their own conscious, responsible decisions. A useful concept in this perspective is ‘sexual competence’ [20], meaning that, sexual intercourse should be protected, consensual, undertaken without regret and as a result of an autonomous decision.”
47. Does the manuscript adhere to the relevant standards for reporting and data deposition? More information is needed about the study.

As a response to the previous comments, the intervention and study were described in more detail.

48. Are the discussion and conclusions well balanced and adequately supported by the data? The conclusions are based not only on this study but include other studies. The conclusions are rather overstated for this study and the authors should ensure that conclusions are based on the data at hand.

We have adapted the discussion section based on this and the above comments.

49. Are limitations of the work clearly stated? Yes limitations section is good.

50. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished? Yes they do.

51. I’m not sure that the title is appropriate. Was the peer education programme truly limited in its effectiveness? Lack of effectiveness may well be related to the study being underpowered or poorly implemented or the outcome measures being inadequate. More detail is needed by answering the questions raised in this review.

We changed the title into: Effectiveness of a peer-led HIV prevention intervention in secondary schools in Rwanda: results from a non-randomized controlled trial.

52. Generally the paper is well-written. I just have a few comments to make:

53. Use past tense throughout.

The manuscripts was revised for the consistent use of the past tense.

54. Keep to US or UK spellings and be consistent. A mixture of US and UK English is used for example in the section on ‘Variables’ second paragraph you use programme and program in the one sentence. This is the same for behaviour and behavior which you use interchangeably.

The manuscripts was revised for the consistent use of US English.

55. The manuscript needs some editing and proof reading.

For example, changes to sentences such as ‘pulling it out of the taboo space’

Typos: susceptibly should be susceptible (2nd para in Results “ On the other hand they were less likely to feel....) - OK

Large fraction should read large proportion.(3rd para in Results) - OK

..Sensitising methodologies as... should read ...such as...(4th para in Discussion) - OK

..while peer educators are now the centre of the intervention informing...needs a comma between intervention and informing (Conclusion 1st para) - OK

If you keep this sentence.....recognize these interventions in their true value- i.e contributing to a changing communication climate etc...it should read...recognise the true value of these interventions in changing the communication climate and breaking taboos. - OK
Editors’ comments

1. Figure 1 is after Figure 2, please make sure the Figures are in chronological order.

The manuscript only includes one figure.

2. Competing interests - Please include a 'Competing interests' section between the Conclusions and Authors' contributions. If there are none to declare, please write 'The authors declare that they have no competing interests'.

We included a section in which we state that “The authors declare that they have no competing interests”.

References


