PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	School-based intervention for the prevention of HPV among adolescents: a cluster randomised controlled study
AUTHORS	Grandahl, Maria; Rosenblad, Andreas; Stenhammar, Christina; Tydén, Tanja; Westerling, Ragnar; Larsson, Margareta; Oscarsson, Marie; Andrae, Bengt; Dalianis, Tina; Nevéus, Tryggve

VERSION 1 - REVIEW

REVIEWER	Elsebeth Lynge University of Copenhagen
REVIEW RETURNED	11-Sep-2015

GENERAL COMMENTS	This is an intervention study to improve primary prevention of human papillomavirus infection. Although this is a valid intention, I am not too positive about the paper. It is very long, and very difficult to read. It is not so clear what is the intervention and what is the control situation. I assume that for the control situation it is 1 hour health interview which includes sexual health, and for the intervention group 30 minutes especially about HPV is added. But this is not clear from the text. To understand the outcome of the intervention, the reader would like to know what was the endpoint value for the intervention group and what was the endpoint value for the control group, what was the RR or difference, and what was the p-value. This cannot be found in the tables. The authors mention an effect on [3]intention to use condoms[2] but no effect on [3]use of condoms[2]. They play this down in the Discussion. But actually, I find this to be some of the most interesting info in the paper. Either this manuscript needs considerable shortening and revision, or it is not acceptable for publication.

REVIEWER	Alice Forster UCL, UK
REVIEW RETURNED	24-Sep-2015

GENERAL COMMENTS	The authors report the findings of a randomised controlled trial of an intervention to prevent HPV infection among adolescents. The study benefits from being a large, multi-centred study, evaluating an intervention that was well described and incorporated relevant psychological theory. It is great to see interventions being trialled in this area. The manuscript could be improved by providing greater

	detail of the method (specific detail described below) and clarification of the study design. The authors describe a cluster randomised controlled trial, but do not report it as such. I would like a statistician to comment on this aspect of the manuscript, including the appropriateness of the sample size calculation for a cluster RCT and analysis. Missing detail It is very useful that the authors have supplied a completed CONSORT checklist, however, I query some of the items that have been coded as N/A. In particular, I want to know if there was/was not blinding. I understand that the nurses may not have been blind to allocation, but was the individual who collected follow-up responses and who analysed the data? Lack of blinding would be a limitation of this trial. Also, I want to know how randomisation was performed and any restrictions. Finally, were any harms of the intervention reported (I presume not as this was low risk, but it should be reported). The authors should report the abstract in line with the CONSORT checklist for abstracts. Additional missing detail includes: - Were the schools selected systematically or opportunistically? - What age are students going into the first year of upper secondary school? - The authors report that nurses completed logs to ensure fidelity to the intervention, however, fidelity was not reported. - It would be useful to report the questions used to assess the primary outcome. Related to this, how was the total HBM score computed? - How were missing data handled? Was the data analysed as intention to treat or per protocol? Minor comments I would encourage the authors to take care when describing the potential impact of their intervention (particularly the last line of the abstract). While cervical screening can prevent mortality from cervical cancer (attendance at screening/intentions to do so was not an outcome of the trial), we do not yet know that HPV vaccination reduces incidence of HPV-related cancer. Although this is likely, we cannot say this for certain yet. I would question whether the intervention really is feasible to integrate into the health interview. The intervention lasts 30 minutes and the usual interview lasts 1 hour (total 1.5 hours). Are schools happy to lose an additional 30 minutes of teaching? Do nurses have the capacity to increase their workload for their students by 50%? In the results, the findings relating to the questionnaire data describes numbers going up or down. It would be useful if the authors describe whether an increase in severity score, for example, means that students believed that HPV was more or less severe than at baseline. When discussing the health belief model in the methods section, the authors omit self-efficacy and cues to action. Was this intentional and if so, why were these aspects of the model not included in the intervention? In the results section, when discussing change in proportion of

	vaccinated girls, I would have liked to have seen the proportions (baseline and follow-up) for the control group as well as for the intervention group. The authors report different response rates in the abstract, results and discussion. Although I presume these are different calculations (i.e. total participating at baseline of those asked and total who participated at baseline and follow-up), it is not clear. Table 2 - I would encourage the authors to check their numbers. For example, the results for HPV vaccination in the intervention group, total n does not equal 390.

REVIEWER	Robin Bruyndonckx Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BIOSTAT), University of Hasselt, Belgium
REVIEW RETURNED	29-Sep-2015

GENERAL COMMENTS	The authors have studied the effect of a face-to-face intervention for prevention of HPV among adolescents. Whereas the study is well designed and conducted, I do have some concerns about the statistical analysis and its interpretation. I have formulated some comments and questions to improve the manuscript. Introduction --- The introduction is well written and provides the reader with the necessary information to understand the framework of the study. After reading the paper, the aim of the study is clear to me, but the paper would benefit from an elaboration on 'favourable beliefs' in the subsection 'aim and hypothesis'. How is this measured? Methods --- I have some questions related to the design and concerns related to the statistical analysis. Population and sample: - You mention that students that were unable to speak English and students with learning disabilities were excluded from the study. Why did you exclude these individuals? Would they not benefit from an intervention? Recruitment and randomisation: - You state that all nurses happened to be female. Do you expect the gender of the nurse providing the treatment to have an effect on the impact of the treatment? If no, this information becomes irrelevant. If yes, it might be good to mention this briefly in the discussion under the Section 'weaknesses'. - You write that the baseline questionnaire was completed at the nurses' office while the follow-up questionnaire was given in the classroom, with classmates present. Maybe you could briefly explain why the work method was changed? Intervention:

	- You mention that condoms were handed out to students in the treatment group at the time of the follow-up questionnaire. Could this trigger the students to only think about condom use at that point in time, not the week before or after, and hence bias the results on intention to use condom with a new partner? I understand that this cannot be changed now, but do think that this deserves some attention in the Discussion. - The manuscript would become more readable by explicitly mentioning the type of outcome variable (e.g. intention to use condom with a new partner: is this expressed as yes/no or as percentage or …) Statistical analysis: - In the first paragraph you mention categorical, ordinal, discrete and continuous variables. Is there a discrete variable present in your study? - In the first paragraph you mention the use of several different tests to assess the difference between the treatment and control group (Pearson's Chi-square, Mann-Whitney and student's t). You mention there that this is done both at baseline and at follow-up, but the results only mention the differences at baseline. - Why is a GEE model used for condom use, HBM total score, susceptibility, severity and barriers, but not for HPV vaccination? - In the second paragraph you mention that a two-sided p-value was used for all analyses, although the main hypothesis states that the intervention group would have more favourable beliefs towards HPV prevention than the control group, which is a one-sided hypothesis. Results ---------- I have some concerns about the results of the statistical analysis. - In the first paragraph you mention a response rate of 90.2%. Was this response rate similar for both groups? - In the second paragraph you mention that there were significant differences between the groups at baseline regarding sex and immigrant status. For this reason, only adjusted results should be presented. The unadjusted results give a distorted view of the reality. - In the second paragraph you also mention that the validity of the GEE for the outcome benefits was not certain. What does 'not certain' mean? - In paragraphs 4 to 8, you report the interpretation of an interaction term while the models (Tables 3 and 4) do not seem to include any interaction terms. You should either clarify that the terms in your tables are in fact interaction terms between group and sex and between group and immigrant status, rather than main effects for sex and immigrant status, or alter the interpretation of the results in paragraphs 4-8. - In paragraph 9 you mention results for the Mc Nemar test. As mentioned before, I would suggest to use a GEE model also for this outcome variable. Discussion ---------- Also in the discussion my main concern is on the interpretation of the results shown in Tables 3 and 4. You discuss your findings as if an interaction term was included in the model although I found no mentioning of an interaction term in the manuscript. Tables and Figures

	---------- I have a small concern about Figure 1 which presents the flowchart from recruitment of schools to final analysis. In the middle of the chart, you go from 60 classes including 1587 students to 394 students receiving intervention by excluding 252 students and not giving the intervention to 8 students. Why does this not add up?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1	**Author response**	
6) It is not so clear what is the intervention and what is the control situation. I assume that for the control situation it is 1 hour health interview which includes sexual health, and for the intervention group 30 minutes especially about HPV is added. But this is not clear from the text.	Thank you for this comment; we have now clarified this in the paper.	Page 7-8
7) To understand the outcome of the intervention, the reader would like to know what was the endpoint value for the intervention group and what was the endpoint value for the control group, what was the RR or difference, and what was the p-value. This cannot be found in the tables.	Since this is a cluster randomized RCT, we had to adjust for potential confounders using a statistical model, in this case a GEE model, in order to get an estimate of the real difference between the intervention and control groups. The reported slopes and accompanying p-values are thus the most informative measures to report. The actual endpoint values are thus of less interest, since these are biased and does not take the dependence within a cluster into account. Furthermore, to compute p-values for these differences, we still would have to resort to GEE, thus replicating the other analyses anyway. Thus, we feel that it is better to not report these values in a table.	
8) The authors mention an effect on ³intention to use condoms² but no effect on ³use of condoms². They play this down in the Discussion. But actually, I find this to be some of the most interesting info in the paper.	Thank you. We agree that this is interesting and believe this finding emphasizes that it is hard to change health behavior, especially with a short period of follow up. We have added a sentence about possible different results with a longer follow-up.	15
9) Either this manuscript needs considerable shortening and revision, or it is not acceptable for publication.	We have shortened the manuscript as much as we have deemed possible, while at the same time trying to provide all the clarifications and additional information requested by the referees and editors.	The full manuscript
Reviewer 2	**Author response**	
10) The authors describe a cluster randomised controlled trial, but do not	The second author, Dr Rosenblad, biostatistician and associate professor has	The full

report it as such. I would like a statistician to comment on this aspect of the manuscript, including the appropriateness of the sample size calculation for a cluster RCT and analysis.	been involved in this project since the design of the study. The paper was not reported as a cluster RCT since we had to recruit school nurses willing to participate before we could randomize schools and classes on cluster level. In addition, all analyses are performed on individual level. The power (sample size) is adequate for these analyses. However, as we discuss on page 16 it would have been even better with a larger sample size in order to be able to conduct more subgroups analyses. Nevertheless, this was not feasible due to logistic reasons. We have now revised the paper as a cluster randomised study. Considering the inherent uncertainties in and approximations of all power calculations, our statistician deemed these power calculations to be appropriate.	manuscript
11) Missing detail *It is very useful that the authors have supplied a completed CONSORT checklist, however, I query some of the items that have been coded as N/A. In particular, I want to know if there was/was not blinding. I understand that the nurses may not have been blind to allocation, but was the individual who collected follow-up responses and who analysed the data? Lack of blinding would be a limitation of this trial.*	Thank you for pointing this out, the CONSORT checklist has been updated. The study could not be double blinded; it was obvious to the school nurse if the student received the "treatment" or not. However, the research assistant who entered all data into SPSS did *not* know whether the questionnaires were intervention or control. This has now been clarified in the text. This project was a PhD project and in Sweden statistical analyses should be performed by the PhD student (with assistance by senior researchers, in this case the second author, a biostatistician). The analyses are based on measuring the effect on the intervention thus we had to mark the questionnaires as IG or CG in SPSS, otherwise we could not have compared the groups at baseline or at follow up. We do not believe that this has affected the results in this paper.	Page 7
12) The authors should report the abstract in line with the CONSORT checklist for abstracts.	We have revised the abstract according to CONSORT checklist for cluster randomised trials.	Abstract
13) Also, I want to know how randomisation was performed and any	The restrictions (inclusion and exclusion criteria) are reported in the paper on page 6. We have now also clarified how the	Page 5-6

restrictions.	randomisation was performed.	
14) Finally, were any harms of the intervention reported (I presume not as this was low risk, but it should be reported).	No harmful effects were reported. This has now been clarified.	Page 14
15) Additional missing detail includes: *- Were the schools selected systematically or opportunistically?*	The schools were selected opportunistically since we first had to recruit school nurses willing to participate. The included schools are the schools were the school nurses who agreed to participate worked. We have now clarified this.	Page 6
16) What age are students going into the first year of upper secondary school?	We have clarified the students' age (aged 16) in the paper	Page 5
17) The authors report that nurses completed logs to ensure fidelity to the intervention, however, fidelity was not reported.	Thank you for this insightful comment. We are currently performing a process evaluation of the study indicating high fidelity among school nurses. This is now briefly mentioned in this paper.	Page 16
18) It would be useful to report the questions used to assess the primary outcome.	Due to space restrictions the questions cannot be included *verbatim* in the main text. However, if asked to do so by the editor we can provide English translations and add those as an appendix.	
19) Related to this, how was the total HBM score computed?	For each individual item in the HBM index, the answers on the five-point Likert scale (Strongly agree to Strongly disagree) where gives scores 0-4 for negative questions, i.e., "strongly disagree" would imply higher health belief and 4-0 for reversed questions. "Do not know" was classified as answering the neutral option. Finally, scores from all included individuals were summarized to give a total score. This is now included in the paper.	Page 11
20) How were missing data handled? Was the data analysed as intention to treat or per protocol?	The missing data, i.e. the students (n=10) who did not complete the follow up questionnaire are not included in the final and reported analyses. All students that completed the follow-up questionnaire where analyzed according to intention to treat principles. However, none of the randomized	Page 11

	students switched from intervention to control group or vice versa.	
21) Minor comments *I would encourage the authors to take care when describing the potential impact of their intervention (particularly the last line of the abstract). While cervical screening can prevent mortality from cervical cancer (attendance at screening/intentions to do so was not an outcome of the trial), we do not yet know that HPV vaccination reduces incidence of HPV-related cancer. Although this is likely, we cannot say this for certain yet.*	OK. We have, with some reluctance, removed the last sentence of the abstract.	Abstract
22) I would question whether the intervention really is feasible to integrate into the health interview. The intervention lasts 30 minutes and the usual interview lasts 1 hour (total 1.5 hours). Are schools happy to lose an additional 30 minutes of teaching? Do nurses have the capacity to increase their workload for their students by 50%?	This question is certainly relevant. We do, however, only provide the evidence for the value and feasibility of the intervention. Whether, in the future, a similar block of information should be included as an add-on to the information given to every student or if something else (with less evidence of benefit) should be removed is a matter for school planners and politicians to decide. We have chosen not to devote space to this certainly interesting discussion in the paper, but we have clarified that the intervention was included in the general health interview scheduled about one hour (not 90 minutes). The control group only received the general information regarding sexual health, while the intervention group received specific HPV information. A discussion about school nurses beliefs regarding a national implementation of the intervention will be discussed in a future paper.	Page 7-8
23) In the results, the findings relating to the questionnaire data describes numbers going up or down. It would be useful if the authors describe whether an increase in severity score, for example, means that students believed that HPV was more or less severe than at baseline.	Thank you for this valuable comment; We have clarified this in the results section in the paper.	Page 13-14
24) When discussing the health belief model in the methods section, the authors omit self-efficacy and cues to action. Was this intentional and if so,	We omit *cues to action* and *self-efficacy* intentionally in this paper, to avoid confusion (since we do not present results about this). The questions about *cues to action* are not	

why were these aspects of the model not included in the intervention?	included since they are focused on the decision-making process about HPV vaccination and measured on other scale levels (dichotomy and open ended questions). *Self-efficacy* was not included in the original HBM model, and we did not measure self-efficacy specifically in the questionnaire. Nevertheless, we do believe that the students' self-efficacy was increased by the intervention, especially regarding HPV vaccine. The components *cues to action* and *self-efficacy* will be discussed in a future paper (the process-evaluation). Due to lack of space the above discussion is not included in the text.	
25) In the results section, when discussing change in proportion of vaccinated girls, I would have liked to have seen the proportions (baseline and follow-up) for the control group as well as for the intervention group.	The vaccination proportion for the control group was *unchanged* as stated in the paper on page 13. We have added the proportion *60.9%.*	Page 14
26) The authors report different response rates in the abstract, results and discussion. Although I presume these are different calculations (i.e. total participating at baseline of those asked and total who participated at baseline and follow-up), it is not clear.	Thank you for this comment, we describe this on page 12 and in the Abstract and have clarified the numbers. A total of 832 students were randomised. Of these, 751 agreed to participate and 741 students completed the follow up and are included in the final analyses and reported in the results.	Abstract Page 12
27) Table 2 - I would encourage the authors to check their numbers. For example, the results for HPV vaccination in the intervention group, total n does not equal 390.	Thank you, we have checked the numbers. Only girls are included in the national vaccination programme (we describe this in the introduction on page 4 and in Table 2, below the table; HPV vaccinated (only girls[a])), consequently the numbers do not add up to 390.	
Reviewer 3		
28) The introduction is well written and provides the reader with the necessary information to understand the framework of the study. After reading the paper, the aim of the study is clear to me, but the paper would benefit from an elaboration on 'favourable beliefs' in the subsection 'aim and hypothesis'. How is this measured?	Thank you. Favourable beliefs (i.e. positive attitude) are used since the theoretical model, HBM, uses beliefs and not attitude (although attitude is commonly used in studies based on HBM). Beliefs are measured based on the HBM concepts; perceived benefits, barriers, susceptibility and severity, as described in the paper. We have clarified this in the aim.	Page 4-5

29) Population and sample: *-You mention that students that were unable to speak English and students with learning disabilities were excluded from the study. Why did you exclude these individuals? Would they not benefit from an intervention?*	This is a very good question. Students with dyslexia and neuropsychiatric disorders (ADHD, Asperger syndrome etc.) who attend the regular classes are included. Students with severe learning disabilities and development disorders (such as Downs's syndrome etc.) attending the special classes or the schools (sometimes these schools are incorporated in the upper secondary schools and sometimes they are apart) are excluded for ethical and methodological reasons. Due to their medical condition and cognitive disability the intervention would have had to be strongly modified (the information would have had to be adjusted for this group and the questionnaire in oral instead of in writing etc.) in order to be feasible. Moreover, it would have required other preparations of the study (pilot studies among this group). According to the ethical guidelines in Sweden children should be aware of what it means to consent to participation in research studies, this would have been ethically challenging among this vulnerable group. To include adolescents not able to speak Swedish (not English) would also have required substantial methodological and logistic challenges. Moreover, some of these children cannot read or write when arriving in Sweden, and could not have completed the questionnaire (even if translated). Therefore, we excluded these adolescents. We have clarified this in the paper.	Page 5-6
30) Recruitment and randomisation: *-You state that all nurses happened to be female. Do you expect the gender of the nurse providing the treatment to have an effect on the impact of the treatment?* *If no, this information becomes irrelevant. If yes, it might be good to mention this briefly in the discussion under the Section 'weaknesses'.*	Thank you for this comment. The reason to include the sentence was as an explanation to us consistently referring to them with female pronouns in the text (i.e. not "he or she"). We have now removed the sentence.	Page 6

31) You write that the baseline questionnaire was completed at the nurses' office while the follow-up questionnaire was given in the classroom, with classmates present. Maybe you could briefly explain why the work method was changed?	After the process evaluation of the pilot study (feasibility study) it became clear that school nurses could not perform the follow-up on an individual level (it was considered a "logistical nightmare"). As suggested, we have briefly mentioned this in the section Pilot study.	Page 9
32) Intervention: *-You mention that condoms were handed out to students in the treatment group at the time of the follow-up questionnaire. Could this trigger the students to only think about condom use at that point in time, not the week before or after, and hence bias the results on intention to use condom with a new partner? I understand that this cannot be changed now, but do think that this deserves some attention in the Discussion.*	Thank you for this question. We realize that this is not clear in the paper. The condoms were handed out *after* the students had completed the follow up questionnaire, consequently we do not expect this to have affected the results on the intention to use condom if new partner. We have now clarified this.	Page 8
33) The manuscript would become more readable by explicitly mentioning the type of outcome variable (e.g. intention to use condom with a new partner: is this expressed as yes/no or as percentage or …)	Outcome variables: * HBM questions, five point Likert scale (Strongly agree to Strongly disagree) * Intention to use condom with a new partner, five-point Likert scale (Strongly agree to Strongly disagree) * HPV vaccinated yes/no/do not know * Condom use during their latest intercourse yes/no We have added this in the paper.	Page 8
34) Statistical analysis: *-In the first paragraph you mention categorical, ordinal, discrete and continuous variables. Is there a discrete variable present in your stud*	Thank you for this question, we have discrete variables in the questionnaire, although these questions are not presented in this paper. Thus we have removed references to discrete variable in the paper, as follows: "For descriptive statistics, categorical data are presented as frequencies and percentages, n (%), ordinal data as medians, means and standard deviations (SD), while continuous data are given as mean and SD.	Page 11

	Differences between the intervention and control groups at baseline are tested with Pearsons' Chi-square test for categorical data, Mann-Whitney test for ordinal data, and Student's independent samples *t*-test for continuous data."	
35) In the first paragraph you mention the use of several different tests to assess the difference between the treatment and control group (Pearson's Chi-square, Mann-Whitney and student's t). You mention there that this is done both at baseline and at follow-up, but the results only mention the differences at baseline.	Thank you for this insightful comment, we do not mention the follow-up results in this paper for all of these statistical tests and have revised this section.	Page 11
36) Why is a GEE model used for condom use, HBM total score, susceptibility, severity and barriers, but not for HPV vaccination?	This was only a subgroup analysis for a secondary outcome, since not all individuals would have any results for actual HPV vaccinations (some were already vaccinated) and actual condom use (some did not have sexual intercourse), which resulted in heavy biases, we thought it would not be worthwhile to perform a GEE analysis.	
37) In the second paragraph you mention that a two-sided p-value was used for all analyses, although the main hypothesis states that the intervention group would have more favourable beliefs towards HPV prevention than the control group, which is a one-sided hypothesis.	Thank you for this insightful comment, you are absolutely right and we have clarified this in the paper, with the following two-sided hypothesis: "The overall aim of the project was to improve primary prevention of HPV by promoting HPV vaccination and increased condom use among upper secondary school students. The hypothesis was that intervention was associated with favourable beliefs (positive attitude) towards HPV prevention at follow-up and that this influenced the actual behaviour."	Page 4
38) I have some concerns about the results of the statistical analysis. *-In the first paragraph you mention a response rate of 90.2%. Was this response rate similar for both groups?*	The response rate was similar, about 90%, for both groups, please see Flow chart (Fig 1) and below; Invited to participate (i.e. randomised): N=832 Intervention n=431	

	Control n=401 Declined participation: Intervention n=37 Control n=44 Response rate Baseline: 751/832=90.2% Intervention n=394/431 (91.4%) Control n=357/401 (89.0 %)	
39) In the second paragraph you mention that there were significant differences between the groups at baseline regarding sex and immigrant status. For this reason, only adjusted results should be presented. The unadjusted results give a distorted view of the reality.	Thank you, we appreciate this comment. We have revised this section accordingly; only adjusted results are now presented.	Table 3 and 4
40) In the second paragraph you also mention that the validity of the GEE for the outcome benefits was not certain. What does 'not certain' mean?	This sentence is now changed to the following wording: "Since the Generalized Estimating Equations model fit did not converge for the HBM construct *benefits*, results for this construct are not presented."	Page 12
41) In paragraphs 4 to 8, you report the interpretation of an interaction term while the models (Tables 3 and 4) do not seem to include any interaction terms. You should either clarify that the terms in your tables are in fact interaction terms between group and sex and between group and immigrant status, rather than main effects for sex and immigrant status, or alter the interpretation of the results in paragraphs 4-8.	We have clarified that it is not an interaction effect, but the main effect presented in the results.	Page 12-13
42) In paragraph 9 you mention results for the Mc Nemar test. As mentioned before, I would suggest to use a GEE model also for this outcome variable.	This was only a subgroup analysis for a secondary outcome, since not all individuals would have any results for actual HPV vaccinations (some were already vaccinated) and actual condom use (some did not have sexual intercourse), which resulted in heavy biases, we thought it would not be worthwhile	

	to perform a GEE analysis. (In addition, probably too small sample for using GEE).	
43) Also in the discussion my main concern is on the interpretation of the results shown in Tables 3 and 4. You discuss your findings as if an interaction term was included in the model although I found no mentioning of an interaction term in the manuscript.	Thank you for this comment, we have clarified this. Please also see our response for question number 41.	Page 12-13
44) I have a small concern about Figure 1 which presents the flowchart from recruitment of schools to final analysis. In the middle of the chart, you go from 60 classes including 1587 students to 394 students receiving intervention by excluding 252 students and not giving the intervention to 8 students. Why does this not add up?	We are thankful for this comment and have revised the figure.	Figure 1

VERSION 2 – REVIEW

REVIEWER	Alice Forster UCL, United Kingdom
REVIEW RETURNED	06-Nov-2015

GENERAL COMMENTS	Thank you for replying so comprehensively to the reviewer comments. I have only one minor comment which the authors may choose to address or not. Point 18 - It would be useful for the questions used to assess the primary outcome to be included as an appendix. This would be very beneficial for future research.

REVIEWER	Robin Bruyndonckx Hasselt University, Belgium
REVIEW RETURNED	12-Nov-2015

GENERAL COMMENTS	Although the authors have addressed all suggested clarifications, I still have some major concerns on the statistical analysis. **Introduction** You state the following: The hypothesis was that intervention was associated with **favourable** beliefs (**positive** attitude) towards HPV prevention at follow-up and that this influenced the actual behaviour. Since the outcome variables under assessment is beliefs towards primary prevention about HPV, and you state here that the intervention group has favourable beliefs, this hypothesis is still one-

	sided while it is tested as if it would be a two-sided hypothesis. The two-sided version of this hypothesis would be that intervention was associated with **different** beliefs (**different** attitude) about HPV prevention. **Methods** **Statistical analysis:** In the second paragraph you mention the use of a McNemar test for HPV vaccination and for condom use. These outcomes have been defined as yes/no/do not know and yes/no. To my knowledge, a McNemar test for a kxj table does not exist. This is a specific test for a kxk table. Therefore, you should either specify that you leave out the 'do not know' group for vaccination status or use a different test statistic here. **Results** Although you state in your response that both the interpretation and the tables are based on main effects rather than interaction terms, the results discussed in paragraphs 4 to 8 cannot be obtained without inclusion of an interaction term. For example, you state: Students with **immigrant** background in the **intervention** group, however, had a 3.291 higher score compared to **immigrant** students in the **control** group (p=0.003). This gives an interpretation for the interaction between immigrant status and treatment group. Interpretations for the main effects would state that students with immigrant background have a higher score compared to non-immigrant students and that students in the intervention group have higher scores compared to students in the control group. Using Table 4 you cannot make the conclusion you state in paragraph 5

VERSION 2 – AUTHOR RESPONSE

1. Introduction: The two-sided version of this hypothesis would be that intervention was associated with different beliefs (different attitude) about HPV prevention.

We agree and have changed accordingly.

2. Methods: Specify that you leave out the 'do not know' group for vaccination status or use a different test statistic here.

This is now clarified. We have specified that we have excluded the "do not know" group for vaccination in the analysis.

3. Results: Although you state in your response that both the interpretation and the tables are based

on main effects rather than interaction terms, the results discussed in paragraphs 4 to 8 cannot be obtained without inclusion of an interaction term. For example, you state: Students with immigrant background in the intervention group, however, had a 3.291 higher score compared to immigrant students in the control group (p=0.003). This gives an interpretation for the interaction between immigrant status and treatment group. Interpretations for the main effects would state that students with immigrant background have a higher score compared to non-immigrant students and that students in the intervention group have higher scores compared to students in the control group. Using Table 4 you cannot make the conclusion you state in paragraph 5.

We have now clarified this in the Results part of the text

VERSION 3 - REVIEW

REVIEWER	Robin Bruyndonckx Hasselt University - Belgium
REVIEW RETURNED	27-Nov-2015

GENERAL COMMENTS	Although I am happy with the changes made to the results section regarding the interpretation of Table 4, I would like to see the same changes being made to the discussion section. In the second paragraph you still mention "the intervention's effect among boys" and in the fourth paragraph "the intervention was especially effective among adolescents with an immigrant background". Please make sure that the Discussion and the Results section are in agreement.

VERSION 3 – AUTHOR RESPONSE

We have now made the minor changes in the discussion section as requested by Reviewer.