Author’s response to reviews

Title: Difference in overall and age-specific prevalence of human papillomavirus infection in Italy: evidence from NTCC trial

Authors:

Iacopo Baussano (baussano@iarc.fr)
Silvia Franceschi (franceschis@iarc.fr)
Anna Gillio-Tos (giliotos.demarco@cpo.it)
Francesca Carozzi (f.carozzi@ispo.toscana.it)
Massimo Confortini (m.confortini@ispo.toscana.it)
Paolo Dalla Palma (paolo.dallapalma@apss.tn.it)
Margherita De Lillo (anapat@ausl.imola.bo.it)
Annarosa Del Mistro (annarosa.delmistro@ioveneto.it)
Laura De Marco (giliotos.demarco@cpo.it)
Carlo Naldoni (cnaldoni@regione.emilia-romagna.it)
Paola Pierotti (paola.pierotti@ausl.bologna.it)
Patrizia Schincaglia (p.schincaglia@ausl.ra.it)
Nereo Segnan (nereo.segnan@cpo.it)
Manuel Zorzi (manuel.zorzi@ioveneto.it)
Paolo Giorgi-Rossi (giorgirossi@asplazio.it)
Guglielmo Ronco (guglielmo.ronco@cpo.it)

Version: 3 Date: 22 December 2012

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

We thank you for your e-mail of 16 November 2012 regarding our manuscript entitled “Difference in overall and age-specific prevalence of human papillomavirus” by Baussano et al. We appreciate the opportunity you give us to revise our manuscript, and you will find the revised version attached.

We have addressed all major and minor points raised by the reviewers. Modifications introduced in the main text have been highlighted in yellow to facilitate the reviewers' work.

Please find below in red fonts a brief statement answering each comment made by the reviewers. Rebuttal to major comments is briefly summarized after each comment. For minor revisions, we did not report every single change in the letter below unless it was deemed necessary for clarity sake. We preferred to only amend the manuscript.

We very much hope BMC Infectious Diseases will consider our manuscript suitable for publication, and look forward to your response.

Yours sincerely,

Dr Iacopo Baussano
Scientist, Infections and Cancer Epidemiology Group
baussano@iarc.fr

ENCL: (1)
Reviewer's report  Title:  Difference in overall and age-specific prevalence of human papillomavirus infection in Italy: evidence from NTCC trial

Version: 2
Date: 14 August 2012
Reviewer: BEIBEI LU

Reviewer's report:
The authors assessed differences in age trends of oncogenic HPV prevalence across nine regions of Italy using a well chosen statistical method. The study contributes important regional age-specific HPV prevalence information.

Major revision
1. The statement in the background, “...the impossibility of simply inferring the natural history of HPV from age distribution and...”, is somewhat out of place as providing inference on HPV natural history is not the focus of the study. Please consider removing the statement.
Authors: The statement has been modified to account for this observation, and to focus the reader’s attention on the geographic heterogeneity in age-specific HR-HPV prevalence in different populations (page 4, lines 13-16).

2. A briefly statement with respect to the following in page 5 first paragraph will be necessary to describe the study population: (1) study design of the NTCC study (randomized controlled trial?) and its endpoints/main objectives; (2) the source of study population – the type of institutions from which the participants were recruited.
Authors: The first paragraph of the Methods section has been modified to account for these observations (page 5, lines 2-5).

3. Hybrid Capture 2 (HC2) assays was used to detect 13 oncogenic HPV types in the current study. Please provide the definition of a valid HC2 test (page 6 first paragraph).
Authors: A HC2 positivity definition has been provided in the Methods section “... with a cut point of 1 relative light unit (RLU) for HPV positivity ...” (page 5, line 26 & page 6, line 1).

I would also like for the authors to consider the potential impact of undetected non-oncogenic HPV types on the overall age-specific HPV prevalence in the current study and discuss how it may affect the comparability of findings between the current study and previous studies.
Authors: The NTCC trial was not designed to provide an accurate description of age-specific curves of LR-HPV types; therefore, we cannot satisfactorily address this comment. Anyhow, low-risk HPV types are less frequently detected but strongly correlated with high-risk types. Hence, the inclusion would not alter our geographical comparisons.

4. The authors raised an interesting discussion to explain the higher HPV prevalence observed among young-to-middle-aged women in Northern Italy in page 8 second paragraph. The authors appeared to suggest (1) the increasing delay in age at first marriage could result in more sexual promiscuity before marriage and thus increased HPV exposures; (2) the birth cohort effect was mainly due to sexual behavioral changes in younger birth cohorts, for instance, earlier initiation of sexual activity. If so, the relationships need to be clarified in the text.
Authors: The text has been amended to account for the reviewer's comments. In particular, we explicitly mentioned that delay in marriage and migrations may affect the increase in sexual promiscuity and affect HPV circulation within the population (page 9, lines 15-21).

/./.


5. The authors suggested in page 8 last paragraph that the steeper decline in HPV prevalence reported by the study conducted in Central and Southern Italy compared to the findings of the current study was likely due to more recent change in sexual behaviors among youngest group of women in Southern Italy. Please explain the type of sexual behavioral changes that recently took place in Southern Italy and discuss how the changes might have affected the age-specific HPV exposures.

Authors: We slightly modified this section, focusing our attention on age-specific difference of sexual patterns across Italy. We also modified the bibliography and removed the somehow confusing (see comment n. 15 of the third reviewer) reference to smoking activity changes (page 10, lines 1-7).

6. The authors stated in the discussion “The largest migration to Turin from Southern and North-eastern Italy, however, had occurred in the 1960s. Almost half the women migrated to Turin from Southern Italy arrived during such period and would have not fallen in the age range eligible for the NTCC study [39].” The relevance of internal migration to possible population mixing in Turin and resulting heterogeneous HPV prevalence observed in the current study is unclear in the statements. Please clarify.

Authors: The text has been amended to account for the reviewer’s comments. In particular we distinguished more clearly the potential role of present and past immigration on HPV infection circulation (page 10, lines 16-23).

Minor essential corrections.
Authors: The manuscript has been amended according to all minor essential corrections suggested by the reviewer.

1. Page 3, Background: Please specify the observation was made among women.

Authors: The text has been amended to take into account the reviewer’s comment (page 4, lines 18-21).

2. Page 4 second paragraph line 4: it should be “…cohort studies suggesting that…”; Line 8: “by” should be removed.

Authors: The text has been amended to take into account the reviewer’s comment (page 4, second paragraph).

3. Page 6 second paragraph: I believe the R package used for this analysis is “ljr” not “lrj”.

Authors: The text has been amended to take into account the reviewer’s comment (page 6, line 18).

4. Page 6 second paragraph first sentence: “Age trends for the prevalence of 13 oncogenic HPV types were assessed for each recruitment centre using a logistic joinpoint regression analysis.”

Authors: The text has been amended to take into account the reviewer’s comment.

Suggested changes
1. Page 5 second paragraph line 6: “Seventy-four percent of eligible women consented to the study and were randomized.”

Authors: The text has been amended to take into account the reviewer’s comment (page 5, lines 19-20).

2. Page 5 third paragraph second sentence: “The Hybrid Capture 2 (HC2) hybridization assay (Qiagen, Hilden, Germany) targeting HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68, was used to determine HPV presence with a cut point of 1 relative light unit (RLU) for HPV positivity, as recommended by the manufacturer.”

Authors: The text has been amended to take into account the reviewer’s comment (page 5, lines 24-26).

./..
3. Page 6 first paragraph second sentence: “HPV prevalence was standardized to the 2004 Italian female population aged 25-60 years.”
Authors: The text has been amended to take into account the reviewer’s comment (page 6, lines 3-5).

4. Page 6 last paragraph last sentence: “...and studied differences between centres as suggested by the intercepts and the slopes.”
Authors: The text has been amended to take into account the reviewer’s comment (page 6, lines 12-13).

5. Page 7 third paragraph first sentence: “No significant join-points were observed in six of the nine centres...”; last sentence: “While HPV prevalence showed a steeper decline with age in Bologna, Florence and Imola than in Trento, Verona and Viterbo..., the test for heterogeneity of age slopes did not reach statistical significance (p=0.08).”
Authors: The text has been amended to take into account the reviewer’s comment (page 7, lines 14-19).

6. Page 7 fourth paragraph line 2: “In Ravenna HPV prevalence decreased linearly until 45 years of age and reached a plateau thereafter, with a slightly higher prevalence observed in the oldest women. In Turin two significant join-points were observed: HPV prevalence remained high up to age 30 years, decreased steeply between age 30 and 38 and maintained a very slow decline after age 38. In Padua, HPV prevalence showed a very slow decline between age 25 and 39, with a higher prevalence for women aged 35-39 than in any other centre (14.1% vs. 9.5% on average), and subsequently had a steep downturn with the lowest prevalence observed in the oldest age groups.”
Authors: The text has been amended to take into account the reviewer’s comment (page 7, lines 23-26 & page 8, lines 1-2).

7. Page 8 first paragraph: “...with the highest age-standardized prevalence in Ravenna (close to a touristic sea resort area) nearly double the lowest prevalence in Trento...”
8. Page 8 second paragraph line 2: “..., notable age patterns were observed in three regions (Ravenna, Turin and Padua).”
Authors: The text has been amended to take into account the reviewer’s comment.

9. Page 8 second paragraph last sentence: “The center in Ravenna had the highest overall HPV prevalence with a second peak in prevalence detected above age 45, ...”
Authors: The text has been amended to take into account the reviewer’s comment.

Level of interest: An article of importance in its field
Quality of written English: Needs some language corrections before being published
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: I declare that I have no financial or non-financial competing interests.

Reviewer's report
Title: Difference in overall and age-specific prevalence of human papillomavirus infection in Italy: evidence from NTCC trial
Version: 2
Date: 9 August 2012
Reviewer: Hormuzd Katki
Reviewer's report:
The NTCC is one of the most important clinical trials for incorporating HPV testing into cervical screening programs. Analyses of data from this study will provide important evidence, not only for screening, but also on the natural history of HPV and cervical precancerous lesions. This paper presents one basic analysis of data from NTCC.

Major Compulsory Revisions
1. Use 5-year age intervals (don't switch to 10 at ages 40+) for table 1. This is important because perimenopause is between 45-55, and so using 40-49 and 50-60 will obscure the effect of perimenopause, which is a key hypotheses being tested. I recognize the models were fit to single-year data, but I find the raw evidence more compelling than the model fits because you have so much raw data.
Authors: We are grateful to the reviewer for this comment. Table 1 has been modified as suggested (page 21).

2. Can you stratify the analysis by the concurrent cytology result? This is useful because HPV prevalence may differ by age for each cytology. In particular the HPV prevalence in ASCUS may drop in perimenopause due to hormonal changes in cells that create "look-alike" ASCUS unrelated to HPV infection. If so, that would be interesting to note.
Authors: The article focuses on age trends of HPV prevalence in the general population to have insights on how much can be interpreted as a result of the natural history of infection or of human behaviour, related to the cohort of birth and migration. The analysis proposed is no longer related to the interpretation of cytology. Indeed, we already reported that the proportion of HPV-positives among women with ASC-US changed by age in a paper on triage (Ronco et al. Europ J Cancer 2007; 43:476-80). Although in Ronco's paper the analysis was by quite wide age groups, and a more detailed analysis in post-menopausal women could have been done, we do not think it is pertinent to the present work.

3. I suggest plotting table 1, as this represents the evidence in the copious raw data. I did not find the model estimates of age-specific prevalence useful because you have so much raw data that you can just present that. Of course, you have to have the model fits to compare across cities, but I didn't find the graph enlightening.
Authors: We plotted the findings reported in Table 1. However, since we prefer to provide readers with row aggregated data the plot is presented as part of the supplementary information section (Figure S1).

4. The city-specific discussion is interesting. What other covariates do you have -- do you have sex behavior in particular [it seems like you do]? If you do, please adjust the prevalences for sex behavior to test your hypothesis that differences between cities are due to sex behavior differences.
Authors: Unfortunately, we cannot adjust out analyses by sexual behaviour, because we did not collect this kind of information in NTCC trial.

Minor Essential Revisions

Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:
I declare that I have no competing interests

-----------------------------
Reviewer's report

Title: Difference in overall and age-specific prevalence of human papillomavirus infection in Italy: evidence from NTCC trial
Version: 2
Date: 13 August 2012
Reviewer: Christine Pierce Campbell

Reviewer's report:
The authors use data from a randomized clinical trial of cervical cancer screening in Italy to calculate the overall (age-standardized) and age-specific prevalence of high-risk HPV across 9 geographical areas of northern and central Italy. In addition to estimating HPV prevalence, the authors use joinpoint analyses to describe changing patterns of HPV infection across the lifespan. Overall, the manuscript does not seem to provide a useful contribution to scientific knowledge regarding the HPV epidemiology. Perhaps this study fills a gap in HPV knowledge in Italy, however, the authors do not make this statement. Furthermore, I am concerned with the use of age-specific prevalence measures in joinpoint analyses; joinpoint analyses are typically used to assess time trends in age-standardized incidence or mortality rates, or occasionally, time trends in prevalence. I believe the authors are attempting to demonstrate a non-linear relationship between age and HPV prevalence for each geographical area; however, I do not know if joinpoint analyses are the most appropriate method for this aim. The authors may want to consult a statistician if they are unsure of the applicability of their method.

Major Compulsory Revisions
1. Methods, fourth paragraph. To assess heterogeneity across study centers, you used meta-analysis methods. Why did you decide to use this particular method instead of a grouped randomized control trial method, since all centers were within the same study (NTCC)?
   Authors: The method we have adopted allows us to provide an estimate of the heterogeneity based on a random effects model, which provides a weighted average of the effect sizes of a group of studies. In fact, we analyzed the NTCC trial data as if each center conducted an independent, but methodologically identical, trial. The random effects approach explicitly assumed that the observed differences result from both between-center variability as well as within-center variability. In general, this approach provides an opportunity to assess variations in outcomes across centers (see Schmid, CH. Exploring heterogeneity in randomized trials via meta-analysis. Drug Information Journal, Vol. 33, pp. 211–224, 1999).

2. Methods, fifth paragraph. Confirm that the use of joinpoint analysis is appropriate for the study aims.
   Authors: As confirmed by the statistical reviewer the joint-point method is appropriate for our analysis. Although it has been so far mainly used in the evaluation of trends in cancer incidence and mortality, joint-point analysis is a method, which allows us to identify significant changes in slopes, with which epidemiologists are relatively familiar.

Minor Essential Revisions
Authors: The manuscript has been amended according to all minor essential revisions suggested by the reviewer; please see below for more detailed comments.

In general: For consistency, use one term to describe the HPV types examined (high-risk, oncogenic, or carcinogenic). For each instance of HPV in the results and discussion, preface the word with high-risk/oncogenic/carcinogenic (e.g., high-risk HPV prevalence) since the authors did not examine low-risk HPV types.
Authors: The text has been amended to take into account the reviewer’s comment.
1. Title. Include the term “high-risk” HPV in the title, since the analyses included only high-risk HPV types.

Authors: The text has been amended to take into account the reviewer’s comment.

2. Abstract, methods. Briefly describe the intervention arm of the NTCC study. The words “changes in prevalence trends by age” is confusing and implies a time component. Perhaps use “differences in the age-distribution of HPV infection.”

Mention there are 9 centers.

Authors: The text has been amended to take into account the reviewer’s comment (page 3, line 12 & page 5, lines 2-5).

3. Abstract, results. “No significant joinpoints were observed” provides little to no information to a reader unfamiliar with joinpoints. “Age slopes did not differ significantly” from what...across centers?

Authors: The text has been amended to take into account the reviewer’s comment (page 3, lines 12-13 & 15-16).

4. Abstract, conclusions. A summary of the findings would be helpful.

Authors: The text has been amended to take into account the reviewer’s comment (page 3, lines 21-22).

5. Background, second paragraph. The first two sentences review literature on HPV incidence and persistence. Perhaps a review of HPV prevalence literature would be more useful. The third sentence is unclear; why can (can’t) we infer the natural history of HPV from age distribution? If you are referring to geographical variation, perhaps this should be explicitly stated. Also, mention why it is important to consider geographical variation in age-specific prevalence of HPV.

Authors: The text has been amended to take into account the reviewer’s comment (page 4, lines 13-16).

6. Background, first paragraph. The first sentence states that previous studies were unable to study cohort effects; this implies that the present study examines cohort effects, which it does not. The aim/purpose of this study is unclear. The last paragraph states what was done, however, it should be written as an aim instead. The use of the words “age shape of HPV prevalence” is confusing. Perhaps say “age-distribution of HPV prevalence.”

Authors: The text has been amended to take into account the reviewer’s comment (page 4, lines 18-21).

7. Methods, fourth paragraph. First sentence should add “...was considered in determining HPV prevalence.” Second sentence could read as follows, “The overall age-standardized HPV prevalence for each center was estimated by direct standardization using the standard female Italian population for the year 2004, truncated from 25-60 years.” Third sentence could read, “The heterogeneity in age-standardized HPV prevalence across centers was tested using the [NAME] method proposed by ...” Name and describe this method, along with why you chose to use meta-analysis methods to analyze heterogeneity across centers within the same study (NTCC). Also, I could not find this data in the results section.

Authors: The text has been amended to take into account the reviewer’s comment (page 6, lines 5-7).

...
8. Methods, fifth paragraph. Which methods were used to calculate the age-distribution of HPV prevalence? Again, “the shape of age-specific prevalence” is confusing; use “The age-distribution of HPV prevalence.” A brief overview of joinpoint analysis should be provided, as not all readers will be familiar with this method. Briefly mention the purpose of joinpoint analysis and what can be achieved through its use, along with a definition/interpretation of a statistically significant joinpoint. The following source provides a clear description of joinpoint methodology: http://www.portal.state.pa.us/portal/server.pt?open=18&objID=901769&mode=2
The R package is “ljr” not “lrj”. Also, the statistical software (R) itself should be cited (see below). Furthermore, you should state the name of the statistical test used within the “ljr” package, if available.


With which methodology did you categorize age into groups? State that the level of statistical significance was set at p<0.05.

Authors: The text has been amended to take into account the reviewer’s comment (page 6, lines 18-23).

9. Results, first paragraph. Mention that the 47,369 women were in the NTCC study. Clarify that the “HPV group” is the HPV testing arm. Perhaps say “...46,900 remained and were included...” Are there any other participant characteristics that you can report (e.g., median age)?

Authors: The “HPV Group” has been clarified (page 7, line 1).

General characteristics of the population investigated in NTCC can be found in references 24-27, and for clarity sake are not reported again in the present manuscript.

10. Results, second paragraph. Present data from Figure 2 prior to Table 1, as Figure 2 provides the most general data (HPV prevalence by center) compared with Table 1 (HPV prevalence by center and age). A statistical test for heterogeneity (?) is mentioned, but this was not described in the methods. Also, this test is not provided as part of Figure 2. Perhaps modify the second sentence to say “Age-standardized high-risk HPV prevalence varied significantly (p<0.001) across centers with ...” All remaining centers had values between 8.2% and 9.4%.” I could not find the numbers 8.0% and 9.5% listed in the manuscript. Table 1 provides some interesting information that should be mentioned in the results text. For example, there was an overall high-risk HPV prevalence of 8.4%, prevalence varied geographically (range: 5.7%-9.7%), and for all centers the prevalence of high-risk HPV was highest among women 25-29 years of age and lowest among women 50-60 years of age.

Authors: Since Table 1 shows the aggregated raw data, we prefer to keep it as the first result presented to the readers. We amended the Table to take into account comment n. 1 made by the second reviewer. We also plotted the findings reported in Table 1, as suggested by the second reviewer (comment n.3). We also present in the main text the overall prevalence results. Center specific data are better represented by Figures 2-4.

11. Results, third paragraph. Suggestion: “Joinpoint regression analyses suggest that no significant joinpoints occurred in age-specific high-risk HPV prevalence in 6 centers...” Which statistical test was used to test for heterogeneity in intercept? Slope? It appears as though the intercept is for women age 25, but you mention this is for overall prevalence; please clarify.

Authors: The text has been amended to take into account the reviewer’s comment (page 3, lines 12-13 & 15-16).

12. Results, fourth paragraph. Explain the “fitted regression”; is this a line of best fit for the data? Why are you fitting a line to the data? What does it represent? This was not mentioned in the methods section. It may be helpful to present the data for Turin, then Ravenna, and finally Padua. Modify the order of centers in Figure 4 accordingly. For Revenna, you mention “low values” and “high values”, but do not define what is considered low or high.

Authors: The text has been amended to take into account the reviewer’s comment (page 7, line 20).
13. Discussion, first paragraph. Instead of saying that you observed different prevalences, perhaps say “We observed geographical variation in age-standardized and age-specific prevalence of high-risk HPV infection...almost twice that of Trento (a mainly rural mountain area), the lowest prevalence.” Are there any data to support the claimed variation in sexual behavior across Italy? If so, please cite these studies.
Authors: The text has been amended to take into account the reviewer’s comment (page 4, lines 2-7).

14. Discussion, second paragraph. In the first sentence, “...non-negligible differences were found...in the shape of age specific prevalence.” Again, perhaps it would be better to say “...differences were found in the age-distribution of high-risk HPV prevalence across areas.” What does non-negligible mean? Are you referring to a statistical test? Instead of age-pattern, I would say age-distribution. I would be careful of using “Persistence of high prevalence” as the term “persistence” has a special meaning in the HPV literature. How are you defining “less young women”? In the fifth sentence, how did you come to these conclusions? The connection is not immediately apparent. Perhaps replace the word husbands with partners, if you did not specifically ask for their marital status or sexual orientation.
Authors: The text has been amended to take into account the reviewer’s comment (page 9, lines 9-25).

15. Discussion, third paragraph. In the second sentence, you mention changes in sexual behavior, but do not describe these “changes”. For what changes are you referring? Also, please elaborate on the smoking habits.
Authors: The text has been amended to account for the reviewer’s comments. Please also refer to comment n.4 from the first reviewer (page 9, lines 15-21).

16. Discussion, fourth paragraph. Elaborate on these thoughts. They sound incredibly interesting but are not well-developed, coherent sentences.
Authors: The text has been amended to account for the reviewer’s comments. Please also refer to comment n.5 from the first reviewer.

17. Discussion, fifth paragraph. Shouldn’t population mixing decrease geographical variation? Perhaps you are referring to a rural-to-urban migration (or vise versa) that may explain an increase in the variation of HPV prevalence. If so, please clarify.
Authors: The text has been amended to account for the reviewer’s comments. Please also refer to comment n.6 from the first reviewer.

18. Discussion, sixth paragraph. Suggestion: “The NTCC study population is large and is considered to be representative of Italian women between the ages of 25 and 60 years who participate in cervical cancer screening ...” In the third sentence you mention no significant effects of education, marriage, or place of birth; were any significant effects found in other variables?
Authors: Age was also considered, but was not significantly associated with differences in participation in screening.

Fourth sentence: I suggest removing the word “especially”, and explaining why you believe selection bias to be unlikely. In the weaknesses, it is important to state that you have no data from southern Italy. The final weakness listed (i.e., power) should be mentioned at the beginning of your weaknesses.
Authors: The text has been amended to take into account the reviewer’s comment (page 11, lines 2-3).
19. Conclusions. I did not see data on differences in age-specific HPV prevalence across centers. Also, I did not see data for birth cohort effects. The final two sentences are interesting and important, and should have been mentioned in both the background and in the discussion sections. These statements require some explanation, and would be best addressed in the background/discussion.

Authors: Figure 3 (along with the significant differences in intercepts of logistic regression) shows the differences in age-specific HPV prevalence across centres. Cohort effects were postulated as possible explanation of changes in the slopes of age-specific prevalence in Turin, Padua, and Ravenna. We have amended the text so to clarify these aspects.

20. Table 1. Title: Prevalence of high-risk cervical HPV by center and age. Column heading, “HPV+/Tested women (%).” Whole numbers should have a decimal place (e.g., 1.0%).

Authors: The text has been amended to take into account the reviewer’s comment (page 21).

21. Figure 2. Title. The age-standardized prevalence of high-risk cervical HPV, by recruitment center. Provide a p-value for test of heterogeneity.

Authors: The text has been amended to take into account the reviewer’s comment (page 7, lines 9-10).

22. Figure 3. Title. Model-based age-specific prevalence of high-risk cervical HPV for centers with linear age trends. Y-axis should be, “High-risk HPV prevalence (%).”

Authors: The text has been amended to take into account the reviewer’s comment.

23. Figure 4. Title. [Model-based?] age-specific prevalence and fitted regression line of high-risk cervical HPV for centers with non-linear age trends. Y-axes should be, “High-risk HPV prevalence (%).” Re-order panels based on appearance in text.

Authors: The text has been amended to take into account the reviewer’s comment.

**Discretionary Revisions**

1. Abstract, background. Clarify that the decreasing prevalence of HPV with increasing age is among women. This is not true among men.

Authors: The text has been amended to take into account the reviewer’s comment (page 3, line 2).

2. Background, first paragraph. “HPV infection varies between populations.” I would use the term “across” rather than between.

Authors: The text has been amended to take into account the reviewer’s comment (page 4, line 2).

3. Methods, first paragraph. The first sentence could add (Figure 1) to the very end, and delete the words “shows the location of...” List the 9 centers in the same order that their population size is described. Adriatic Sea is without an “o”.

Authors: The text has been amended to take into account the reviewer’s comment.

4. Methods, second paragraph. In the first sentence, specify that it is the NTCC study.

Authors: The text has been amended to take into account the reviewer’s comment (page 5, lines 2-5).

5. Methods, third paragraph. Instead of “In the HPV group”, I would recommend saying “Among women in the HPV testing arm”, which is a bit more specific. What is phase 1 and 2, and is it important to mention here? Instead of “targeting any of the oncogenic HPV types” I would say “was used to identify any of the high-risk HPV types ... and was considered positive at 1 relative light unit.” There is no need to add (RLU) since the term is not used again in the manuscript. “High reproducibility of HPV testing between laboratories...” should say “across laboratories...”

Authors: The text has been amended to take into account the reviewer’s comment (please see replies above).
Level of interest: An article whose findings are important to those with closely related research interests

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

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

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

Reviewer's report
Title: Difference in overall and age-specific prevalence of human papillomavirus infection in Italy: evidence from NTCC trial
Version: 2
Date: 15 October 2012
Reviewer: H. Hyune-Ju Kim

Reviewer's report:
See the attached.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

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

Review of “Difference in overall and age-specific prevalence of human papillomavirus infection in Italy: evidence from NTCC trial”
This article studies prevalence patterns of HPV infection in 9 regions of Italy using a logistic regression model. To investigate if there are changes in the pattern, the authors used the joinpoint approach developed for a logistic regression model, and the authors concluded that “there are substantial differences in overall and age-specific HPV prevalence across Italian areas.”

Some specific comments and questions are written below.

(a) Page 5, second paragraph: The authors described the randomization used in the study to assign each subject to one of the two groups: conventional cytology group and HPV testing group. However, I do not see any comparison between the two groups, and it is not clear what is the motivation/goal of the randomized two group study.

Authors: The comparisons of the two arms enrolled in NTCC trial can be found in references 24-27, and for clarity sake are not reported again in the present manuscript.

(b) Page 6, line 4: It was indicated that the heterogeneity between centres was tested using the method by DerSimonian and Laird. Please provide a short summary on their method such as parametric analysis, multiple comparison, etc.

Authors: The text has been amended to take into account the reviewer's comment (page 6, lines 5-7).

(c) Page 6, second paragraph: Provide more details on logistic joinpoint regression analysis, especially on how the model is selected. For example, include the overall significance level and the p-value(s) if the testing approach was used and also what type of testing method was used.

Authors: The text has been amended to take into account the reviewer's comment (page 6, lines 18-23).

(d) Page 7, third paragraph: Since the logistic regression model is used, expressions like “Age slopes appeared slightly steeper...” may mislead readers to understand it as a slope between age and prevalence (%). Make it clear that the slope is the slope parameter of the logit function.

Authors: The text has been amended to take into account the reviewer's comment (page 7, lines 14-19).
(e) Page 8, second paragraph, “In Turin and Padua, however,...”: It was not clear whether “high HPV infection prevalence up to about age 30 and 40” was relative to other centres. If so, it would be helpful to provide the mean prevalence rates during these periods and compare them with corresponding values in other centres to check if the rates in Turin and Padua are significantly higher than those of other centres.

Authors: Here the comparison was drawn within each center and not between centers. The text has been amended to take into account the reviewer’s comment (page 9, lines 12-15).

(f) Figure 2: Is Torino same as Turin?
Authors: The text has been amended to take into account the reviewer’s comment.