Author's response to reviews

Title: Human papillomavirus infection and cervical dysplasia in female sex workers in Northeast China: an observational study

Authors:
Haiqing Jia (haiqing725@163.com)
Xiaobin Wang (lnwdif703@163.com)
Zaiqiu Long (intv@vip.sohu.ne)
Liankun li (Li_liankun@sina.cn)

Version: 3
Date: 13 June 2015

Author's response to reviews: see over
Dear Editor-in-Chief,

Thank you very much for offering us an opportunity to revise our manuscript. We are also greatly grateful to the reviewers for their careful review of the manuscript with constructive comments. We have revised the manuscript based on the reviewers’ comments. While the changes made in the manuscript are highlighted using the track changes feature, our point-by-point responses to the reviewers’ comments are detailed below.

Reviewer #1

1) What is the period of recruitment of FSWs sample and controls? Also, the authors report that 1000 women who underwent regular gynecologic examinations were considered as a representative sample of the general population, however I suggest to provide more details, such as, in addition to the period of recruitment, if the recruitment was consecutive, and if any criteria of exclusion or inclusion were applied. Please, although the recruitment has been carried out in Northwest China, specify if sample and controls were of Northwest China ethnicity.

Response: FSWs were sampled at the same time as they were detained at the police station within a period of three months. The controls were all employees from one single factory in Shenyang. They were recruited within a month, a time period when their free annual gynecologic examinations were performed. As a random sample of the general population, no specific inclusion and exclusion criteria were applied to the control subjects. There was no difference in the cultural and racial/ethnic backgrounds between FSWs and control subjects; they were all Han Chinese.

2) I was surprised that age at first sexual intercourse was a significant risk factor, while the time as sex worker was not. How was the time as sex worker measured? There was any measure about the rate of working activity? Some details about the questionnaire on this might be helpful.

Response: We were also puzzled by this finding. Nevertheless, the association between the age at first sexual intercourse and the risk of HPV infection has been also reported in previous studies (another relevant reference has been added to the revised manuscript). Currently, the exact mechanisms underlying this documented association are not fully understood, but in our manuscript we have provided possible explanations in the Discussion section (lines 217-225, page 9). Moreover, given that HPV is transmitted only through sexual contact, the younger the age at first sexual intercourse, the longer the total years of chance to expose to HPV and the higher the risk of HPV infection. We have added this to the Discussion section as an additional explanation of the association between the age at first sexual intercourse and increased risk for HPV infection.

The time as a sex worker was measured as the period from first commercial sex to the detention at the police station. Regarding the rate of commercial sex working activities, there were no accurate approaches to measure.

3) What do the author mean with menopause: age at menopause or if the women are already in menopause?

Response: We actually mean post-menopause. Corresponding changes have been made throughout the revised manuscript.
4) Since the HPV prevalence in FSWs in this study is especially higher among women aged less than 20 or more than 51, why age was not added in the multivariate analysis? Other major comments are about the necessity of more details in the statistical analysis section and the revision of the tables. In detail:

**Response:** The reviewer is absolutely right that age can be included in the multivariate regression analysis. Nevertheless, since we have clearly shown in Table 2 that the risk of HPV infection is significantly increased in two age groups: ≤20 years and >51 years, inclusion of unstratified age in the multivariate regression analysis is less meaningful than separation of ≤20 years and >51 years into two independent variables.

More detailed information on the statistical approaches used has been added to where appropriate as advised.

5) “Putative risk factors… were analyzed by univariate and multivariate analyses” (line 145), please specify which test was used (logistic regression?)

**Response:** Yes, logistic regression analyses were performed. This has been specified in the revised manuscript.

6) Table 1: which test was used to calculate the statistical significance? Pearson’s chisquare? Comparing what? I assume that for HPV infection the four cells were 210, 191, 790 and 118, but I wonder how it was calculated for ASCUS and CIN. Please, explain better.

**Response:** As the title implies, the present paper involves two major aspects: HPV infection and cervical intraepithelial lesions. Presented in Table 1 are data on differences in HPV infection and various cervical intraepithelial lesions between healthy control subjects and FSWs. Yes, the data were analyzed by Pearson’s chi-square test.

7) Table 2: please, add N and HPV prevalence for controls, and discuss these data with respect to the literature about Chinese population (if any). Add also 95% confidence intervals in both tables 1 and 2.

**Response:** The data presented in Table 1 are about differences in HPV infection and cervical lesions between normal control subjects and FSWs and those in Table 2 about differences in HPV infection between different age groups, analyzed by chi-square test. In our case, 95% confidence intervals may provide no additional meaningful information with respect to what are to be analyzed.

8) Table 3: the title reports “linear regression”. I guess it’s a typo and the correct analysis is logistic regression, otherwise explain how the analysis was performed. In addition, the p-value reported on the first line for the risk factors with more than two categories (i.e. 0.36 in the line that refers to the marital status single) does not have any sense, since that line refer to the reference category by definition (OR=1). I guess the authors meant the significance in general; in this case please, move the p-value to the heading line (i.e. marital status). Besides, when the overall test is not significance, it is of small use to report the significances for the distinct categories (this last comment is discretionary).

**Response:** Yes, “linear regression” is a typo and has been corrected. The p-value on the first line was a mistake and has been deleted.
9) Why the result for the univariate analysis on Menopause is not reported in Table 3 and Use of vaginal medications and Regular Ob/Gyn exams (and only those) are not reported in the multivariate analysis (Table 4)?

Response: The variable ‘menopause’ was mistakenly omitted in Table 3 and has been added in the revised version.

Minor comment:

10) In the statistical analysis section, “data on the prevalence rates of HPV and CIN were analyzed by Pearson’s chi-square test” (line 144): I guess the authors mean “data on the prevalence rates of HPV and CIN in FSWs in comparison to controls and also considering age stratification were analyzed by Pearson’s chi-square test”. Please, specify.

Response: The reviewer is absolutely right; the analysis was performed to compare differences between normal control subjects and FSWs. Changes have been made as advised.

11) Refs 9 and 10 do not seem to be about FSW (lines 97-98)


12) “HPA infection” in Table 1: is it a typo?

Response: Yes, it is a typo. Correction has been made in the revised manuscript.

13) is it available any data about age at first sex intercourse and menopause among controls? It would be interesting to know if the results obtained on the FSWs sample is valid also for the general population.

Response: We agree that inclusion of the data on the age at first sexual intercourse and menopause in control subjects would make the paper more informative. Unfortunately, however, these data were not collected since the control subjects included in this study were employees from a single local factor in Shenyang and they were recruited in during their free annual gynecologic examinations (as a part of their employment benefits).

Reviewer #2

1) LINES 30-31: Please, revise the sentence (Major Compulsory Revision).

Response: Changes have been made.
2) LINE 37: It should be explained the reason why the Authors decided to enroll in study design a sample size 309 FSWs compared to a control group of 1000 healthy subjects (Major Compulsory Revision).

Response: Based on the results from a meta-analysis, there is a huge variation (12.8%-84.8%) in the prevalence rate of HPV infection in FSWs across Asian countries. Therefore, it is hardly possible to use the reported prevalence data as a reference to accurately calculate sample size needed to achieve the expected power of analysis. And the sample size in this study was chosen primarily based on the available number of subjects at the time of recruitment. The above statements have been added to the revised manuscript.

3) LINE 43: Please, revise the sentence (Major Compulsory Revision)

Response: Done as required.

4) LINES 43-44: Please, revise the English grammar in the sentence (Major Compulsory Revision).

Response: Done as required.

5) LINES 47-48: P<0.1 is not statistically significant and also the sentence is not well proposed. Please, correct and rephrase properly (Major Compulsory Revision)

Response: This was a typo and correction has been made.

6) LINES 48-51: these are discussion of data and it should be moved in the conclusion section (Minor Essential Revisions).

Response: Rewording has been done.

7) LINE 52: it should be better explain what did you mean for age at first intercourse, whether it is referred to young age is not a risk factor because in the multivariate analysis OR resulted <1. It is clear with an OR>1 that only the variable menopause is a risk factor for HPV acquiring in FSWs Please provide correction and or a better explanation (Major Compulsory Revision).

Response: In the multivariate regression analysis, both age at first sexual intercourse (P=0.045) and menopause (P=0.032) were significantly associated with HPV infection.

PAPER

8) LINES 138-139: Was the interview administered only to FWSs ? What about the control group? Please, clarify (Major Compulsory Revision).

Response: Yes, the interview was administered only to FWSs but not to control subjects. This is because the primary focus of this study was on the difference in HPV infection rates between FWSs and healthy controls and the major risk factors for HPV infection in FWSs in Northeastern China.

9) LINES 170-171: The variable menopause was not investigated in the univariate analysis; could you explain why? Please, clarify (Major Compulsory Revision)
Response: Menopause was included in the univariate analysis indeed. The result was not presented in Table 3 due to a mistake and has been added in the revised version.

10) Discussion: the discussion section should be shortened, with a proper discussion of the main findings relevant collected from data analysis; moreover, there are some points to better clarify.

Response: Changes, where appropriate, have been made.

We hope that you and the reviewers will agree that we have adequately addressed all concerns raised in the first round of review. We will be more than happy to cooperate with you if further revision is needed.

Sincerely yours,

Haiqing Jia