PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Is the risk of cervical atypia associated with the interval between menarche and the start of sexual activity? A population based cohort study
AUTHORS	Adhikari, Indira; Eriksson, Tiina; Luostarinen, Tapio; Apter, Dan; Lehtinen, Matti

VERSION 1 - REVIEW

REVIEWER	Kirsten Frederiksen Danish Cancer Society, Denmark
REVIEW RETURNED	08-Apr-2019

GENERAL COMMENTS	This study aimed to assess, whether a short interval between age at first intercourse (FSI) and age at start of oral contraception (OC) respectively is associated with risk of cervical atypia or worse. The question is addressed using a subpopulation of 914 women among the 2,399 women enrolled in the PATRICIA trial, who received HAV-vaccine. Cases were defined as those with cervical atypia or worse at at least one of the semi-annual samples taken during the four-year follow-up period. Regarding the analyses however, several problems with the analytical approach exist, and I recommend reanalyzing data using a different analytical approach. First of all the timing is problematic. Some of the covariates like for instance smoking and life-time number of sexual partners are measured by the end of follow-up; whereas the outcome measure used is a combination of samples taken during the total four years of follow-up. A woman observed with cervical atypi at one of the first semiannual visits might for instance start smoking, become HPV/Chlamydia positive and so on later; meaning that the outcome value is measured before the exposure assessment. To have the correct timing of exposure and outcome, exposure should be evaluated before outcome. For the two main exposure variables, which are whether OC and FSI respectively is initiated within the first three years after menarce (and perhaps the combination to account for interaction), a traditional follow-up analysis would start following the women up for cervical abnormalities from three years after menarche, and exclude those who were positive before this point in time. In the analyses in this paper all women are followed from inclusion (age 16-17 years) excluding those being positive at baseline, regardless of age at

	menarche. Among other things this means that for those with late age at menarche a positive outcome measurement before menarche age+3 years is counted as an event connected with exposure information based on future knowledge. Another issue is the mixing up of exposure and outcome, where the effect of each factor on the risk of cervical atypia or worse is studied in subgroups of women with short and long time from menarche to FSI and OC respectively (with non-users categorized together with users in “Interval ge 3 yrs”). The purpose of these analyses is not clear to me. And again problems with timing exists, as the categorization into groups before and after 3 years after menarche might be based on information occurring after the outcome. Furthermore the role of the covariates other than the two main exposure variables is not described, neither in the paper itself nor in the STOBE Statement Checklist. Do they serve as confounders, mediators or why are they included in the models?

REVIEWER	Deborah KONOPNICKI Saint-Pierre University Hospital, Brussels, Université Libre de Bruxelles, Belgium
REVIEW RETURNED	13-Apr-2019

GENERAL COMMENTS	Thank you for asking me to review this article. The study is based on a large and well documented cohort. I have red several times the article and the author’s answers to other reviewers but I am still wandering about its findings. As an example, the conclusion of the abstract and of the article are not the same; what is the taken home message? The authors have looked at the association between menarche and the start of sexual activity and cervical atypia. “Cervical atypia” includes both low grade cytological abnormalities (that reflects HPV infection most of the time) and high grade cytological lesions that are precursor of cancer but that may also spontaneously regress before becoming cancer. “Cervical atypia” includes also atypical abnormalities of unknown significance that might be benign (ASCUS) or that might be at high risk of high-grade dysplasia (ASCH). There are already several studies looking at the association between the risk HPV infection and the interval between menarche the start of sexual activity. However, there are few studies looking at the association between real cancer precursors (high grade cervical lesions) and the interval between menarche and the start of sexual activity. The authors have used “cervical atypia” as endpoint to retain statistical power however what means their results and what is its application in a clinical point of view? The discussion of their results is too short and disappointing: what are the biological plausibility’s to explain why their results are different from other authors (for example: metaplasia, hormonal influence, age of partner, genotypes of HPV, etc.). Could they clarify better the impact on HPV infection and the impact on cyto/histological abnormalities? Regarding cervical abnormalities, two well known risk factors have not been taken into account in the present study and should be

	discussed, too: HIV coinfection and high-risk HPV other than 16 and 18. 1/HIV is a well-known cofactor that increases both risk of HPV infection and HPV-associated dysplasia and is linked to number of life time partner and use of condom. 2/ High risk HPV other than 16/18: In Ruiz's article ("Proximity of first sexual intercourse to menarche and risk of high-grade cervical disease". JID 2012:206 (15 Dec)) which has a similar design as the current study, 1012 Finnish young women were included. Looking at HPV 16 and 18, the proportion of young women infected with these HPV at the end of the study was similar (11-12%) whether the interval was < or > 3 years between menarche and sexual activity; however, the infection with other high-risk genotype was significantly higher in the <3 years group. This data should also be taken into account in the present study. At last, some other data should be described such as ethnicity, the proportion of different cyto/histological abnormalities (n=198), the p-value for the odd ratios and the reason to use both age at first sexual intercourse and start of oral contraception as covariate should be detailed.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Kirsten Frederiksen

Institution and Country: Danish Cancer Society, Denmark

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This study aimed to assess, whether a short interval between age at first intercourse (FSI) and age at start of oral contraception (OC) respectively is associated with risk of cervical atypia or worse. The question is addressed using a subpopulation of 914 women among the 2,399 women enrolled in the PATRICIA trial, who received HAV-vaccine. Cases were defined as those with cervical atypia or worse at least one of the semi-annual samples taken during the four-year follow-up period.

Regarding the analyses however, several problems with the analytical approach exist, and I recommend reanalyzing data using a different analytical approach.

Question- First of all the timing is problematic. Some of the covariates like for instance smoking and life-time number of sexual partners are measured by the end of follow-up; whereas the outcome measure used is a combination of samples taken during the total four years of follow-up. A woman observed with cervical atypi at one of the first semiannual visits might for instance start smoking, become HPV/Chlamydia positive and so on later; meaning that the outcome value is measured before

the exposure assessment. To have the correct timing of exposure and outcome, exposure should be evaluated before outcome. For the two main exposure variables, which are whether OC and FSI respectively is initiated within the first three years after menarce (and perhaps the combination to account for interaction), a traditional follow-up analysis would start following the women up for cervical abnormalities from three years after menarche, and exclude those who were positive before this point in time. In the analyses in this paper all women are followed from inclusion (age 16-17 years) excluding those being positive at baseline, regardless of age at menarche. Among other things this means that for those with late age at menarche a positive outcome measurement before menarche age+3 years is counted as an event connected with exposure information based on future knowledge.

Answer- We should like to thank the reviewer for this comment. First of all, there were no women who had positive atypia findings before menarche. Moreover, in our longitudinal study only outcomes occurring after the exposure(s) were eligible. We re-checked our data as per the reviewer's request. As for one atypia case within the menarche+3 years category, the concomitant start of OC use (exposure) and withdrawal of cervical sample where the atypia diagnosis was made (outcome) could not be excluded. This case was removed, and the data was reanalysed without material effect to the point estimates, however. The study design and setting have now been clarified in the Materials and Methods section, page number 2, second paragraph and page 3, third paragraph.

Question- Another issue is the mixing up of exposure and outcome, where the effect of each factor on the risk of cervical atypia or worse is studied in subgroups of women with short and long time from menarche to FSI and OC respectively (with non-users categorized together with users in "Interval ge 3 yrs"). The purpose of these analyses is not clear to me. And again problems with timing exist, as the categorization into groups before and after 3 years after menarche might be based on information occurring after the outcome.

Answer- The purpose of the analyses stratified by short and long time from menarche was to be able to study the independent role of the cervical atypia risk factors. As mentioned above, no suggested mixing has occurred.

Question- Furthermore the role of the covariates other than the two main exposure variables is not described, neither in the paper itself nor in the STOBE Statement Checklist. Do they serve as confounders, mediators or why are they included in the models?

Answer- The covariates were included in the multivariable analysis to check if the interdependency of the main exposure variable differs but they did not act as confounders or mediators. This is mentioned in the text, in the result section page no. 10, last paragraph.

Reviewer: 2

Reviewer Name: Deborah KONOPNICKI

Institution and Country: Saint-Pierre University Hospital, Brussels,

Université Libre de Bruxelles,

Belgium

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

Thank you for asking me to review this article.

The study is based on a large and well documented cohort.

Question- I have red several times the article and the author's answers to other reviewers but I am still wandering about its findings. As an example, the conclusion of the abstract and of the article are not the same; what is the taken home message?

Answer- The requested changes have been made in the conclusion of abstract.

The authors have looked at the association between menarche and the start of sexual activity and cervical atypia. "Cervical atypia" includes both low grade cytological abnormalities (that reflects HPV infection most of the time) and high grade cytological lesions that are precursor of cancer but that may also spontaneously regress before becoming cancer. "Cervical atypia" includes also atypical abnormalities of unknown significance that might be benign (ASCUS) or that might be at high risk of high-grade dysplasia (ASCH).

There are already several studies looking at the association between the risk HPV infection and the interval between menarche the start of sexual activity. However, there are few studies looking at the association between real cancer precursors (high grade cervical lesions) and the interval between menarche and the start of sexual activity.

Question- The authors have used "cervical atypia" as endpoint to retain statistical power however what means their results and what is its application in a clinical point of view?

Answer- Even if ours is a sizeable study, the numbers of high-grade dysplasia cases is small. In addition to retaining statistical power, we think that also the clinical validity of the obtained results (the Discussion section is now expanded, please, see below) is better with the general end-point.

Question- The discussion of their results is too short and disappointing: what are the biological plausibility's to explain why their results are different from other authors (for example: metaplasia, hormonal influence, age of partner, genotypes of HPV, etc.). Could they clarify better the impact on HPV infection and the impact on cyto/histological abnormalities?

Answer- The discussion has been expanded, and now raises the two issues of C.trachomatis associated and HPV-associated cervical atypias separately concomitantly raising the impact of these overlapping two entities. Please see in the discussion section, 3rd paragraph page number 12.

Question- Regarding cervical abnormalities, two well known risk factors have not been taken into account in the present study and should be discussed, too: HIV coinfection and high-risk HPV other than 16 and 18.

1/HIV is a well-known cofactor that increases both risk of HPV infection and HPV-associated dysplasia and is linked to number of life time partner and use of condom.

Answer- The annual number of new HIV infections in Finnish females aged 15 to 24 years during the last decades has been negligible (below 10 in the entire country for the female population of 300.000). HIV co-infection is not an issue in our study as now noted in the discussion page 11, second paragraph.

2/ High risk HPV other than 16/18: In Ruiz's article ("Proximity of first sexual intercourse to menarche and risk of high-grade cervical disease". JID 2012:206 (15 Dec)) which has a similar design as the current study, 1012 Finnish young women were included. Looking at HPV 16 and 18, the proportion of young women infected with these HPV at the end of the study was similar (11-12%) whether the interval was < or > 3 years between menarche and sexual activity; however, the infection with other high-risk genotype was significantly higher in the <3 years group. This data should also be taken into account in the present study.

Answer- Participants to the Ruiz et al. study were solely from the Helsinki Metropolitan area or from the four next biggest cities in Finland. While HPV16 infection has been and is highly prevalent in adolescent and young adult females in the entire country the relative proportions of other high-risk HPV types have remained low in the vast majority (13 out) of our 18 study site communities.

Question- At last, some other data should be described such as ethnicity, the proportion of different cyto/histological abnormalities (n=198), the p-value for the odd ratios and the reason to use both age at first sexual intercourse and start of oral contraception as covariate should be detailed..

Answer- At the time when the study participants were enrolled in the study, they were a homogenous population of Caucasian Finnish women as indicated in the text (page 11, second paragraph). Our study, therefore, did not look for ethnicity. The proportion of different cyto/histological abnormalities has been added in the result section page number 6, second paragraph. About the p-value, this is a post-hoc analysis of a clinical trial, where the p-values cannot be used.

VERSION 2 – REVIEW

REVIEWER	Kirsten Frederiksen Danish Cancer Society
REVIEW RETURNED	04-Jun-2019

GENERAL COMMENTS	As stated on top of page 3 end-of study questionnaire information was used to define exposures and co-variables; whereas outcome was defined during the four year study period. This means that smoking, life-time number of sexual partners, condom use and sexual transmitted infections are measured after the outcome. Could that affect your results? Also the information about women with cervical atypia on page 6 relates to information at the end of follow-up (average 23 years) which is after atypia. I suggest adding information about when atypia occurs either in text or in table. Some discrepancies between the objective stated in the abstract, in the introduction and the results presented exist. In the abstract it is stated that the objective is to investigate whether the risik of atypia is associated with short interval between menarche and FSI or start of OC use respectively. The estimates in the top part of

	Table 3 however evaluates totaly different questions (not mentioned before). That is whether short time between menache and FSI/start OC modifies the association between other risk factors (HPV, Chlamydia,..) and risk of atypia; or whether HPV, Chlamydia,... are associated with risk of atypi within subgroups according to time between menarche and FSI/start OC use. The rationale for and hypothises underlying these analyses should be stated beforehand in the manuscript. Regarding the main exposure variables, Table 3 is rather confusing and overly complicated. Some of the many NA's must be due to the fact that there are no women with the combination "more than 3 years from menarche to FSI" and "less than 3 years from menarche to OC start". What are the rationale for and the hypothesis behind the two presented contrasts? According to the aim in the abstract I would expect the relevant contrasts to be the marginal comparisons (FSI: 54/248 versus 132/434, OC 31/161 versus 155/421); whereas the presented contrasts are more in line with the aim presented in the introduction. A simple table only presenting the four (in fact three) possible combinations of the two variables would be much more easy to understand. The linear estimates from the model including the two exposure variables as continuous variables in the model, mentioned on page 10 in the text, could be included in this table as well.

REVIEWER	Deborah KONOPNICKI Saint-Pierre University Hospital
REVIEW RETURNED	30-May-2019

GENERAL COMMENTS	The discussion remains poor; there is no real discussion on the reasons why the authors find results that differs from another very large study.

VERSION 2 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 2

Reviewer Name: Deborah KONOPNICKI

Institution and Country: Saint-Pierre University Hospital

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

The discussion remains poor; there is no real discussion on the reasons why the authors find results that differs from another very large study.

Answer: The Discussion has been improved (please, see paragraphs 3 and 5, page number 11 and 12) especially by explaining out comprehensive over-time evaluation of risk-taking behaviour by the questionnaire data yet noting that the exposure was taking place before the outcome (paragraph 3). These differences compared to earlier studies are also discussed (paragraph 5).

Reviewer: 1

Reviewer Name: Kirsten Frederiksen

Institution and Country: Danish Cancer Society

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

As stated on top of page 3 end-of study questionnaire information was used to define exposures and co-variables; whereas outcome was defined during the four year study period. This means that smoking, life-time number of sexual partners, condom use and sexual transmitted infections are measured after the outcome. Could that affect your results?

Answer: Characteristics of life habits such as smoking and risk-taking sexual behaviour were measured in a questionnaire that gathered mostly cumulative information like "life-time number of partners". As now discussed this is the most comprehensive way of evaluating possible impact of this kind of behaviour characteristics.

Also the information about women with cervical atypia on page 6 relates to information at the end of follow-up (average 23 years) which is after atypia. I suggest adding information about when atypia occurs either in text or in table.

Answer: The data is about atypia accumulating by the end of the follow up – changes made in the text, result section page no. 6, paragraph 1.

Some discrepancies between the objective stated in the abstract, in the introduction and the results presented exist. In the abstract it is stated that the objective is to investigate whether the risik of atypia is associated with short interval between menarche and FSI or start of OC use respectively. The estimates in the top part of Table 3 however evaluates totaly different questions (not mentioned before). That is whether short time between menache and FSI/start OC modifies the association between other risk factors (HPV, Chlamydia,..) and risk of atypia; or whether HPV, Chlamydia,... are associated with risk of atypi within subgroups according to time between menarche and FSI/start OC use. The rationale for and hypothises underlying these analyses should be stated beforehand in the manuscript.

Answer: New table 4 is made keeping the estimates from the lower part of table 3. Changes made in the text, material and methods section, page 3, paragraph 1 and results section page 8 and 9.

Regarding the main exposure variables, Table 3 is rather confusing and overly complicated. Some of the many NA's must be due to the fact that there are no women with the combination "more than 3 years from menarche to FSI" and "less than 3 years from menarche to OC start". What are the rationale for and the hypothesis behind the two presented contrasts? According to the aim in the abstract I would expect the relevant contrasts to be the marginal comparisons (FSI: 54/248 versus 132/434, OC 31/161 versus 155/421); whereas the presented contrasts are more in line with the aim presented in the introduction. A simple table only presenting the four (in fact three) possible combinations of the two variables would be much more easy to understand. The linear estimates from the model including the two exposure variables as continuous variables in the model, mentioned on page 10 in the text, could be included in this table as well.

Answer: Changes made in the tables. Table 3 has now only the co-variables. A new table 4 has been made with only two categories “less than 3 years from menarche to FSI” and “more than 3 years from menarche to OC start”, page 10. Interval was also used as a continuous variable as noted in the text.