

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)	Visibility of smoking among schoolteachers in Spain and associations with student smoking: a cross-sectional study
AUTHORS	Escario, José-Julián; Wilkinson, Anna

VERSION 1 – REVIEW

REVIEWER	Dr Abhishek Shankar Assistant Professor Department of Preventive Oncology Dr BR Ambedkar Institute Rotary Cancer Hospital All India Institute of Medical Sciences New Delhi, Delhi - 110029, India
REVIEW RETURNED	04-Jun-2017

GENERAL COMMENTS	<p>This is an important study to correlate smoking pattern among school teacher I have gone through this manuscript and i am writing my suggestions after review.</p> <ol style="list-style-type: none">1. Smoking related facts in introduction needs to be corrected and recent facts and figures need to be mentioned. Kindly refer the recent Global report by WHO.2. Dont put details of your methods in introduction. there is no detail of rational of study in introduction. You need to discuss the background behind this research question which is missing. Introduction needs to be written in detail with proper and up to date evidence.3. Materials and Methods look ok.4. Results are very difficult to understand and very confusing. Kindly re frame the sentence to make it simple.5. Discussion is not on the line of results. So kindly take positive findings and try to discuss one by one to make it more simple to understand. Point wise discussion helps the reader to reach to a conclusion.
-------------------------	---

REVIEWER	Lotus Sofie Bast National Institute of Public Health, University of Southern Denmark
REVIEW RETURNED	20-Jun-2017

GENERAL COMMENTS	<p>Thank you for the opportunity to review this interesting manuscript. I have written my comments below. I think that the manuscript would benefit from a revision with a clear focus. Further, I have a major concern about the analyses, as they are not controlled for well known confounders of pupil smoking, i.e. friend and family smoking.</p>
-------------------------	---

	<p>ABSTRACT</p> <p>13#: “Not only is the decision to smoke or not considered, but also the number of cigarettes consumed”. I think that this sentence should be moved to the objectives. This is not methods.</p> <p>23#: unclear what the OR’s refer to, please provide an explanation.</p> <p>43#: Strengths and limitations. First bullet point is fine; the three others does not reflect the important issues in the manuscript.</p> <p>INTRODUCTION</p> <p>As I understand, the most interesting in this manuscript is on page 4 line 27-32: “Expanding the traditional measure of exposure.....”. Please provide more information about what is already known about different ways to report on smoking outcomes, and point to why it is important to expand the measure. This is the new thing in this manuscript. In my opinion, we already know that teachers smoking impacts student smoking. See i.e. publications from Poulsen et al. 2002, Hamidreza 2014 and Sabiston 2009.</p> <p>The very general introduction with numbers of annual deaths justifying the public health relevance is – in my eyes – unnecessary. This information is common knowledge within smoking prevention. You could place more focus on the school context, the teachers as role models and so on.</p> <p>The focus on differences between countries (page 4, line 20), I would leave out. The manuscript contains only data from Spain. Therefore, a comparison between countries is not possible. This also means that the objective in line 33-35 should be leaved out.</p> <p>METHODS</p> <p>The methods section is very well written. I have only minor comments.</p> <p>Page 5, #26: Please provide an explanation of HOW schools were selected.</p> <p>#29: Participation rate of WHAT?</p> <p>#31: Can you provide the overall participation rate of pupils in classes? And the total N of invited students.</p> <p>#56: May I propose that you use another title for this section? Controls may be confused with comparisons in an intervention study. You could use “exposure variables” or “covariates”.</p> <p>Page 6, #36: How answers this question? If pupils, are they able to answer this. Please provide a discussion about the reliability of this in the discussion section.</p> <p>#50: Income, as disposable money weekly in euros by the adolescents. I have never heard of this in a study – could you elaborate on the reliability of this measure, and WHY it is important here?</p> <p>Data analyses – nice section!</p> <p>RESULTS</p> <p>#46: In my opinion, you don’t have to provide the proportion of no-smokers. You have given the proportion of smokers.</p> <p>#51: is it 46% of all adolescents or of smokers?</p> <p>Page 8, #10: Comment to table 1. Below the table you write that weighting is used. However, there is no description of this in the text?</p> <p>DISCUSSION</p> <p>Generally I would lay more weight on discussing the results regarding the two different measures of smoking.</p> <p>Page 10, line #47: Following these theories we assume that an important part of this association could be interpreted as a causal effect. Are you really able to conclude on this? It might be that students who already smoke, are more prone to see teachers smoke???</p>
--	---

	<p>There is a major limitation in this study regarding discussion of methods. There is no methodological discussion. What about the reliability and validity of self-reportings from pupils? What about selection.</p> <p>Table 1: Please provide a column with p-values of difference. The table would be easier to read if you provided the numbers as percentages.</p> <p>Tables 2 and 3: Please indicate reference categories in the tables. Have I understood correctly, that none of the analyses are adjusted for confounders?</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dear Author

This is an important study to correlate smoking pattern among school teacher. I have gone through this manuscript and I am writing my suggestions after review.

We thank the referee for his/her comments which have truly helped to improve the paper. In the following pages we explain how we have dealt with them.

1. Smoking related facts in introduction needs to be corrected and recent facts and figures need to be mentioned. Kindly refer the recent Global report by WHO.

Response: Following the suggestion of the referee, we have updated these figures with the Who Reports (2013 and 2015).

2. Don't put details of your methods in introduction. There is no detail of rational of study in introduction. You need to discuss the background behind this research question which is missing. Introduction needs to be written in detail with proper and up to date evidence.

Response: Following his/her comment, the current version of the manuscript better describes in the introduction section the research background. We have also included some research questions. In addition, new references have been incorporated.

3. Materials and Methods look ok.

Response: We want to thank the referee for his/her positive opinion about this.

4. Results are very difficult to understand and very confusing. Kindly reframe the sentence to make it simple.

Response: We have explained what the OR and IRR refer to in the "Data analyses section" with the aim of making easier to understand the results. In addition, we have explained a little more the zero inflated model. Finally, results have been related to research hypotheses.

5. Discussion is not on the line of results. So kindly take positive findings and try to discuss one by one to make it more simple to understand. Point wise discussion helps the reader to reach to a conclusion.

Response: As we have pointed out in point 2, we have included some research questions. We believe that it facilitates the understanding of the Results and Discussion sections in a more point wise form. We have also related more important findings to research questions.

Reviewer: 2

Comment: Thank you for the opportunity to review this interesting manuscript. I have written my comments below. I think that the manuscript would benefit from a revision with a clear focus. Further, I have a major concern about the analyses, as they are not controlled for well known confounders of pupil smoking, i.e. friend and family smoking.

Response: We thank the referee for his/her comments which have truly helped to improve the paper. In the following pages we explain how we have dealt with them.

Starting with the major concern mentioned above and related with the last comment of the reviewer, we have pointed out, in the "Data Analysis" section, that the count models have been estimated including all covariates at once, consequently, the coefficients for all the variables are obtained controlling for the rest of variables or confounders. We apologize for had not clarified this point in the first version of the article.

ABSTRACT

13#: "Not only is the decision to smoke or not considered, but also the number of cigarettes consumed". I think that this sentence should be moved to the objectives. This is not methods.

Response: Following the suggestion of the referee, we have moved this sentence to the methods section.

23#: unclear what the OR's refer to, please provide an explanation.

Response: We have now explained what the OR and IRR refer to in the "Data analyses section". In addition, we have explained a little more the zero inflated model.

43#: Strengths and limitations. First bullet point is fine; the three others does not reflect the important issues in the manuscript.

The last tree bullet points have been modified. More concretely, the second bullet point is entirely changed and the term "heterogeneity" has been eliminated. We believe that now it is clear what "heterogeneity" means in point three.

INTRODUCTION

As I understand, the most interesting in this manuscript is on page 4 line 27-32: "Expanding the traditional measure of exposure.....". Please provide more information about what is already known about different ways to report on smoking outcomes, and point to why it is important to expand the measure. This is the new thing in this manuscript. In my opinion, we already know that teachers smoking impacts student smoking. See i.e. publications from Poulsen et al. 2002, Hamidreza 2014 and Sabiston 2009.

Response: Following the referee, we have included more details about what is already known and about the different measures. For example, I have pointed out that two studies used tree level measures (Poulsen et al., 2002; Roohafza et al., 2014) and that Roohafza et al. (2014) found that higher perceived teacher smoking was associated with student smoking among girls. In addition, the suggested references have been included, although we believe that the suggested Hamidreza 2014 refers to Roohafza et al. (2014)

Comment: The very general introduction with numbers of annual deaths justifying the public health relevance is – in my eyes – unnecessary. This information is common knowledge within smoking prevention. You could place more focus on the school context, the teachers as role models and so on. In this point we have used, as the other referee suggested us, the last Global Report by WHO. We have also put more focus in school context.

The focus on differences between countries (page 4, line 20), I would leave out. The manuscript contains only data from Spain. Therefore, a comparison between countries is not possible. This also means that the objective in line 33-35 should be leaved out.

Response: Following the referee advice, we have eliminated this sentence. We just meant that it was necessary results for more countries.

METHODS

The methods section is very well written. I have only minor comments.

Response: We want to thank the referee for his/her positive opinion about this.

Page 5, #26: Please provide an explanation of HOW schools were selected.

Response: We have added a more detailed explanation of the sampling method. To be more specific, we have pointed out that it follows a bi-etapic sample design in which, schools (primary units) and classes (secondary units) have been selected in a random way. We have also pointed that the sampling error is 0.6% at the 95% confidence level. The methodology employed to collect data is explained (in Spanish) in more detail in the following link (page 64):

http://www.pnsd.msssi.gob.es/profesionales/sistemasInformacion/sistemaInformacion/pdf/8_ESTUDE S_2012_Informe.pdf

#29: Participation rate of WHAT?

Response: This has been now clarified in the text. More concretely, the methodology notes pointed out that 86.5% of the randomly selected schools participated in the survey.

#31: Can you provide the overall participation rate of pupils in classes? And the total N of invited students.

Response: This information is not available in the methodology published in the link mentioned above. We only know the participation rate of selected schools and the total N of students that fill in the survey, but there is not information about the proportion of students that skipped classes and, consequently, didn't fill in the survey.

#56: May I propose that you use another title for this section? Controls may be confused with comparisons in an intervention study. You could use “exposure variables” or “covariates”.

Response: As the referee suggest, we have used “Covariates” as the new title.

Page 6, #36: How answers this question? If pupils, are they able to answer this. Please provide a discussion about the reliability of this in the discussion section.

Response: There are several questions that are filled in by the staff of the survey, these questions include the school type.

#50: Income, as disposable money weekly in euros by the adolescents. I have never heard of this in a study – could you elaborate on the reliability of this measure, and WHY it is important here?

Response: The referee is right in pointing out that this variable is not common in part of the literature. However, it could be due to unavailability of the information in most surveys about tobacco, alcohol and drugs. This variable is fundamental in consumption or demand theory and, consequently, it could be a basic confounder to control for. Other studies that use this survey have controlled for this variable, see for example Escario and Wilkinson (20015). This variable is also included in several studies in the field of health economics. Moreover, this variable appears as statistically significant in both decisions: i) being a smoker or not; and ii), conditioned by being a smoker, the quantity smoked.

Data analyses – nice section!

Response: We want to thank the referee for his/her positive opinion about this.

RESULTS

#46: In my opinion, you don't have to provide the proportion of no-smokers. You have given the proportion of smokers.

Response: We have eliminated this proportion.

#51: is it 46% of all adolescents or of smokers?

Response: The information in Table 1 is referred to the entire sample except the consumption conditioned in positive consumption (Smoking Consumption >0).

Page 8, #10: Comment to table 1. Below the table you write that weighting is used. However, there is no description of this in the text?

Response: The weight variable is provided in the survey data. According to the methodology published in the previous link, this weighting is done taking into account regions (there are 19 regions) and school ownerships in order to correct for the proportion differences between the sample and the total population.

DISCUSSION

Generally I would lay more weight on discussing the results regarding the two different measures of smoking.

Response: We agree with the referee in this point, we would like to lay more weight in this topic, but we had focused in the smoking social context (teachers, friends, and parents), and only parent smoking appears as a significant predictor in the number of cigarettes consumed. Obviously, this has been pointed out in the text. Moreover, and following the referee suggestion, we have discussed the role of other variables like Income in the level of consumption.

Page 10, line #47: Following these theories we assume that an important part of this association could be interpreted as a causal effect. Are you really able to conclude on this? It might be that students who already smoke, are more prone to see teachers smoke???

Response: The referee is right in this point and we have toned down the assumption eliminating "important" of the sentence and we have pointed out the reverse association.

There is a major limitation in this study regarding discussion of methods. There is no methodological discussion. What about the reliability and validity of self-reportings from pupils? What about selection. These limitations have been commented in the limitations in the paper. To mitigate one of these concerns we have provided a reference (Dolcini et al., 1996) that concludes that self-reports yield reliable estimates of substance use when data are gathered confidentially.

Table 1: Please provide a column with p-values of difference. The table would be easier to read if you provided the numbers as percentages.

Response: Following the suggestion of the referee, we have provided p-values for the tests and we have provided all the numbers as percentages, except for Income (in euros) and Smoking Consumption (number of cigarettes).

Tables 2 and 3: Please indicate reference categories in the tables.

Response: We have indicated the reference categories in Tables 2 and 3.

Have I understood correctly, that none of the analyses are adjusted for confounders? As we have mentioned at the beginning of this response, we have pointed out, in the "Data Analysis" section, that the count models have been estimated including all covariates at once, consequently, we report confounder-adjusted estimates. We apologize for had not made clear this point in the first version of the article.

VERSION 2 – REVIEW

REVIEWER	Paul Zarogoulidis Pulmonary Department, ``Theageneio`` Anticancer Hospital Thessaloniki, Greece
REVIEW RETURNED	09-Sep-2017

GENERAL COMMENTS	An interesting paper in its field
-------------------------	-----------------------------------

REVIEWER	Alessandro Coppo Eastern Piedmont University, Italy
REVIEW RETURNED	19-Sep-2017

GENERAL COMMENTS	<p>The study is of interest even if the question is not completely new. My suggestion is to compare the work with Poulsen 2002 (cited in the paper) that provided some methodological indications here used that could be acknowledged, for example the frequency of teachers' smoking observations (every day, sometimes, never). Furthermore the same study described the association measure adjusting for potential confounders such as parental smoking, best friends smoking, proportion of smokers in the class, and sex. I think this paper should discuss why this analysis was not performed.</p> <p>Minor revision: 1) Reference 8 appears to be outdated 2) Could you define the term "school ground" (such as: inside school building, inner courtyards, close the entrances)?.</p>
-------------------------	--

	This aspect is of importance as many countries have adopted smoking restrictions also in the school premises. This topic could be analysed in the discussion taking into account what the current law provides for Spanish schools.
--	---

REVIEWER	Ankur Singh The University of Adelaide, Australia
REVIEW RETURNED	01-Oct-2017

GENERAL COMMENTS	<p>The authors have tested associations between visibility of smoking among teachers and presence of smoking among Spanish adolescents. Additionally, they examine associations between presence of smoking among peers/parents and presence of smoking among the adolescents. They further examine the association between these exposures and numbers of cigarettes smoked. Although, the research question on the association between visibility of smoking among teachers and smoking practices of adolescents is interesting, the manuscript very hard to follow. The manuscript requires extensive editing for English language to improve its readability. The reviewer has highlighted concerns across the sections that needs to be addressed:</p> <p>Abstract:</p> <ol style="list-style-type: none"> 1. The authors start the objective with “Limited research has examined the correlation between smoking among schoolteachers and student smoking”. This is a key problem as it is not clear across the text whether the authors formulate their hypotheses on a prediction model or a causal model. Given that they start with correlation and test associations, there is a lack of inconsistency. The authors should clarify if they are interested in prediction/correlation or association. 2. Clearly the pathway that teachers smoking may impact student smoking through a range of factors and not limited to only whether students consider teachers as their ‘role models’ or not. It will be helpful if the authors be specific about their objectives. 3. The following sentence does not seem appropriate: Not only is the decision to smoke or not considered, but also the number of cigarettes consumed. Should be rephrased to improve understanding. 4. In the methods section the authors report that count data regressions were applied. I am not sure if count data regressions are a type of regression model. There are different kinds of regression models that are applied for outcomes that are count variables, such as poisson/negative binomial regression. Within the results section the authors mention odds ratios that are certainly not derived from negative binomial regression models but from the logistic regression models. <p>Background:</p> <ol style="list-style-type: none"> 1. In the second paragraph the authors mention a range of risk factors that influence uptake of smoking. Labelling all of them as risk factors may not be appropriate as social support can be a protective factor for smoking. 2. “few studies have investigated the effects of teacher smoking behaviour on adolescent smoking (e.g. *10).” – Referencing in this manner seems inappropriate. 3. The authors have listed five hypotheses (H1a, H1b, H2, H3 and H4). The consistency between the five different hypotheses has not been established within the background under the social ecological framework. Therefore, the hypotheses appear as disjointed.
-------------------------	--

	<p>4. The authors mention that their hypotheses in based in the social ecological model. However, the theoretical basis for their four different hypotheses is very poorly supported within the background section.</p> <p>5. Having five hypotheses makes it difficult to ascertain what the authors aim to achieve.</p> <p>6. The last paragraph is hard to follow.</p> <p>Methods</p> <p>1. What does bi-etapic mean?</p> <p>2. Dependent variable is hard to understand. What do the authors mean by “For the dependent variable, Smoking Consumption, students who selected the two first options were coded as zero mean number of cigarettes per day, and for students who reported daily smoking, the self-reported mean value was used.” Please specify clearly how this variable was categorized. What does it mean to have zero mean number of cigarettes per day?</p> <p>3. Can the authors justify their choice of measures through references? Are these measures validated?</p> <p>4. The authors have clubbed all the other variables under covariates. They should follow a structure where they should specify the exposures based on theory and their research question rather than putting all of them as covariates. This is a key challenge as the applied approach seems like investigating predictors of smoking among adolescents rather than the association between specific exposures and outcomes. This limitation is also observed in the statistical modelling approach.</p> <p>5. The details to explain the process of estimation of IRR from negative binomial regression does not seem relevant to the study here. The logic behind the application of zero inflated negative binomial regression model due to excessive zeroes seem sufficient.</p> <p>Results</p> <p>1. How can you present “brief descriptive analysis”? You can only present the findings from the descriptive analysis.</p> <p>2. The authors should refrain from using p-values to demonstrate statistical significance of findings. Confidence intervals are more informative and appropriate. Given the large sample size, the p-values are likely to be small in most cases.</p> <p>3. The odds ratio of 20.8 is extremely high for those having a smoker as a close friend. This result needs to be discussed if the authors are looking at predictors of tobacco use.</p> <p>Discussion</p> <p>1. It will be helpful if the authors could specify in the first paragraph about what they found for each hypothesis rather than just saying that the results are consistent with the four hypotheses.</p> <p>2. The authors state: “following these theories we assume that part of this association could be interpreted as a causal effect. Other part of the association could be due to the fact that smoker students are more prone to see teachers smoke.” Due to the cross-sectional nature of the data, the authors cannot establish anything on the temporal sequence between the exposure and the outcome. This is a far-stretched conclusion and the authors should be cautious on making such conclusions.</p> <p>3. Authors should refrain from presenting odds ratio within the discussion section.</p> <p>4. The discussion should mainly focus on discussing the results on the main objective of the study.</p>
--	--

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Comment: An interesting paper in its field after review.

Response: We thank the referee for his/her previous comments, which have truly helped to improve the paper.

Reviewer: 2

Comment: The study is of interest even if the question is not completely new. My suggestion is to compare the work with Poulsen 2002 (cited in the paper) that provided some methodological indications here used that could be acknowledged, for example the frequency of teachers' smoking observations (every day, sometimes, never). Furthermore the same study described the association measure adjusting for potential confounders such as parental smoking, best friends smoking, proportion of smokers in the class, and sex. I think this paper should discuss why this analysis was not performed.

Response: We thank the referee for his/her comments, which have truly helped to improve the paper. We compared our work with Poulsen's work, and acknowledge that we use frequency of teachers' smoking visibility instead of binary indicators. Poulsen et al. (2002) used the three level measure in the descriptive analysis (Table 1), but then, they dichotomised this frequency (never versus everyday or sometime) when visibility of teacher smoking was introduced, as a control variable in the logit regression (Table 2).

Consistent with the approach of Poulsen et al. (2002), we adjusted for similar confounders, such as parental smoking, best friends smoking, and sex. Unfortunately, we cannot use the proportion of smokers in the class, as the 2012 survey does not provide an identification code for school and/or class, which is necessary to compute such a proportion.

Minor revision:

1) Reference 8 appears to be outdated.

Response: Taking into account the referee's suggestion, we have updated the reference, and now use the recent 2014 survey targeted to teachers and carried out by the Spanish Government's Delegation for the National Plan on Drugs.

2) Could you define the term "school ground" (such as: inside school building, inner courtyards, close the entrances)? This aspect is of importance as many countries have adopted smoking restrictions also in the school premises. This topic could be analysed in the discussion taking into account what the current law provides for Spanish schools.

Response: Following the suggestion of the referee, we have pointed out that the term "school ground" includes school buildings and outdoors on the school premises. Moreover, the Spanish smoking restrictions have been commented in the Discussion section.

Reviewer: 3

Comment: The authors have tested associations between visibility of smoking among teachers and presence of smoking among Spanish adolescents. Additionally, they examine associations between presence of smoking among peers/parents and presence of smoking among the adolescents. They further examine the association between these exposures and numbers of cigarettes smoked. Although, the research question on the association between visibility of smoking among teachers and smoking practices of adolescents is interesting, the manuscript very hard to follow. The manuscript requires extensive editing for English language to improve its readability. The reviewer has highlighted concerns across the sections that needs to be addressed:

Response: We thank the reviewer for completing such a close read of the manuscript. We apologize for the lack of clarity throughout, and have edited the entire document to improve readability.

Abstract:

1. The authors start the objective with "Limited research has examined the correlation between smoking among schoolteachers and student smoking". This is a key problem as it is not clear across the text whether the authors formulate their hypotheses on a prediction model or a causal model. Given that they start with correlation and test associations, there is a lack of inconsistency. The authors should clarify if they are interested in prediction/correlation or association.

Response: Thanks for noting this confusing sentences. Given the cross-sectional nature of the data, this is not a causal study. The first, second and third sentences of the abstract have been modified in order to avoid the previous confusion. Our study primarily examines the extent to which student smoking is associated with perceived smoking of teachers after controlling for socioeconomic characteristics.

2. Clearly the pathway that teachers smoking may impact student smoking through a range of factors and not limited to only whether students consider teachers as their 'role models' or not. It will be helpful if the authors be specific about their objectives.

Response: We agree with the reviewer. As noted above, we have modified the second and third sentences in the abstract to eliminate the term "role model".

3. The following sentence does not seem appropriate: Not only is the decision to smoke or not considered, but also the number of cigarettes consumed. Should be rephrased to improve understanding.

Response: We have revised the sentence to say, "We examine both smoking behavior as well as the number of cigarettes consumed."

4. In the methods section the authors report that count data regressions were applied. I am not sure if count data regressions are a type of regression model. There are different kinds of regression models that are applied for outcomes that are count variables, such as poisson/negative binomial regression. Within the results section the authors mention odds ratios that are certainly not derived from negative binomial regression models but from the logistic regression models.

Response: As the reviewer notes, ORs are not derived from negative binomial regression models, but from logistic regression models. Although in the manuscript we only comment on the two zero inflated models, we have estimated four count regression models: Poisson, Negative Binomial, Zero Inflated Poisson, and Zero Inflated Negative Binomial. We have used LR and Vuong tests that have selected the Zero Inflated Negative Binomial model as the preferred model (Cameron and Trivedi, 2013).

The Zero Inflated Negative Binomial model combines a logistic regression and a Negative Binomial that are jointly estimated by a maximum likelihood function. Consequently, it is a regression model, and the exponentiated coefficients are OR in the logistic part and IRR in the Negative Binomial part. We choose to present ORs and IRRs to facilitate understanding for our intended readers, the public health audience.

Background:

1. In the second paragraph the authors mention a range of risk factors that influence uptake of smoking. Labelling all of them as risk factors may not be appropriate as social support can be a protective factor for smoking.

Response: Thank you for noting this. We now say “lack of social support” to clarify.

2. “few studies have investigated the effects of teacher smoking behaviour on adolescent smoking (e.g. *10).” – Referencing in this manner seems inappropriate.

Response: We have removed the asterisk.

3. The authors have listed five hypotheses (H1a, H1b, H2, H3 and H4). The consistency between the five different hypotheses has not been established within the background under the social ecological framework. Therefore, the hypotheses appear as disjointed.

Response: Following the referee’s suggestion, we have explained how the hypotheses H3 and H4 are based on the Social Ecological Model.

4. The authors mention that their hypotheses in based in the social ecological model. However, the theoretical basis for their four different hypotheses is very poorly supported within the background section.

Response: We have pointed out that the last two hypotheses also are based on the Social Ecological Model as they refer to essential parts of the social environment.

5. Having five hypotheses makes it difficult to ascertain what the authors aim to achieve.

Response: The referee is right in saying that five hypotheses appear too many. However, in this paper, the first three hypotheses examine the association between student smoking and teacher smoking visibility, while the last two examine the association between student smoking and smoking of other people in the social environment (friends and parents). The objective is to examine associations between student smoking and exposure to key individuals in their social environment who smoke – in particular teachers, but also friends and family members.

6. The last paragraph is hard to follow.

Response: Following the referee’s suggestion, we have rewritten this paragraph.

Methods

1. What does bi-etapic mean?

Response: It is a sampling design technique in which the final units (classes in this data set) are randomly selected between bigger units (schools) that have been previously randomly selected.

2. Dependent variable is hard to understand. What do the authors mean by

“For the dependent variable, Smoking Consumption, students who selected the two first options were coded as zero mean number of cigarettes per day, and for students who reported daily smoking, the self-reported mean value was used.”

Please specify clearly how this variable was categorized. What does it mean to have zero mean number of cigarettes per day?

Response: The paragraph containing the definition of the dependent variable has been rewritten. We have also eliminated the word “mean” in order to simplify the reading.

3. Can the authors justify their choice of measures through references? Are these measures validated?

Response: We apologize for this oversight. The data presented in the paper are drawn from a nationally representative survey of substance use behavior in Spain. The data are used by Spain and the European Union to monitor substance use among school age children. As such all measures are valid, moreover research is conducted to ensure the validity and reliability of the measures. We have added two references to the technical documents in support of the validity of the measures.

4. The authors have clubbed all the other variables under covariates. They should follow a structure where they should specify the exposures based on theory and their research question rather than putting all of them as covariates. This is a key challenge as the applied approach seems like investigating predictors of smoking among adolescents rather than the association between specific exposures and outcomes. This limitation is also observed in the statistical modelling approach.

Response: To improve clarity for the reader, we have added a heading: “Main exposures of interest” which refers to teachers, friends and parents smoking behavior. The heading “Covariates” now covers the demographic and socioeconomic variables, which are standard covariates in many analyses.

5. The details to explain the process of estimation of IRR from negative binomial regression does not seem relevant to the study here. The logic behind the application of zero inflated negative binomial regression model due to excessive zeroes seem sufficient.

Response: The explanation about the interpretation of the IRR’s was introduced to the text because one of the other reviewers suggested that we provide an explanation of “what the OR’s refer to”. To facilitate interpretation of the results for the readers, we extended our explanation to cover both ORs and IRRs.

Results

1. How can you present “brief descriptive analysis”? You can only present the findings from the descriptive analysis.

Response: This has been corrected in the manuscript.

2. The authors should refrain from using p-values to demonstrate statistical significance of findings. Confidence intervals are more informative and appropriate. Given the large sample size, the p-values are likely to be small in most cases.

Response: Following the referee’s suggestion, we have eliminated p-values and replaced them with confidence intervals.

3. The odds ratio of 20.8 is extremely high for those having a smoker as a close friend. This result needs to be discussed if the authors are looking at predictors of tobacco use.

Response: We agree that this is a high OR. However, our result is consistent with previous work, as other researchers have found that this association is high. For example, Poulsen et al., (2002) found ORs of 12.9 and 12.2, which they justify by two hypotheses: smokers tend to spend time with other smokers (selection hypothesis), and smokers are more aware of other smokers (projection hypothesis). We have added this to the discussion. Also the confidence intervals are quite tight, which lends support to the validity of the result, as they suggest limited measurement error.

Discussion

1. It will be helpful if the authors could specify in the first paragraph about what they found for each hypothesis rather than just saying that the results are consistent with the four hypotheses.

Response: Thank you for this suggestion. We have rewritten this section.

2. The authors state: “following these theories we assume that part of this association could be interpreted as a causal effect. Other part of the association could be due to the fact that smoker students are more prone to see teachers smoke.”

Due to the cross-sectional nature of the data, the authors cannot establish anything on the temporal sequence between the exposure and the outcome. This is a far-stretched conclusion and the authors should be cautious on making such conclusions.

Response: The referee is correct. We do not mean to imply that our results can be interpreted causally; rather the theories that we used to guide our research have causal implications. We have rewritten this section to clarify our intention.

3. Authors should refrain from presenting odds ratio within the discussion section.

Response: We no longer present odds ratios within the discussion section.

4. The discussion should mainly focus on discussing the results on the main objective of the study.

Please note we modified the discussion to accommodate other reviewer’s suggestions, but overall we believe that the discussion focusses on the main objective of this study. This is to say, we have mainly focused on the association between adolescent smoking and perceived smoking of teachers, best friends, and of parents.

VERSION 3 – REVIEW

REVIEWER	Ankur Singh University of Adelaide, Australia
REVIEW RETURNED	16-Nov-2017

GENERAL COMMENTS	The authors have sufficiently and appropriately addressed all the concerns raised in the the previous version of the manuscript. They must be commended for the revisions that have further improved this paper and the conclusions drawn are limited to their findings. Minor suggestion: searching on bi-etapic on google and PubMed shows that only very few studies have used this terminology. For the international audience, it may be helpful if the authors could include the meaning of bi-etapic or use two-stage cluster sampling.
-------------------------	--

VERSION 3 – AUTHOR RESPONSE

Reviewer: 3

Comment: The authors have sufficiently and appropriately addressed all the concerns raised in the previous version of the manuscript. They must be commended for the revisions that have further improved this paper and the conclusions drawn are limited to their findings. Minor suggestion: searching on bi-etapic on google and PubMed shows that only very few studies have used this terminology. For the international audience, it may be helpful if the authors could include the meaning of bi-etapic or use two-stage cluster sampling.

Response: Thank you for your comments and guidance throughout the review process, which have truly helped to improve the paper. In this new revision we have changed the terminology following the referee's suggestion.