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)	Sport and scholastic factors in relation to smoking and smoking initiation in older adolescents: a prospective cohort study in Bosnia and Herzegovina
AUTHORS	Sekulic, Damir; Sisic, Nedim; Terzic, Admir; Jasarevic, Indira; Ostojic, Ljerka; Pojskic, Haris; Zenic, Natasa

VERSION 1 - REVIEW

REVIEWER	Fawaz Mzayek University of Memphis, USA
REVIEW RETURNED	18-Oct-2016

GENERAL COMMENTS	Re: Sport and Scholastic Factors in Relation to Smoking and Smoking Initiation in Older Adolescents: a Prospective Cohort Study in Bosnia and Herzegovina The study examined longitudinal associations of factors related to academic and sport performance/involvement with smoking behavior in high school students in Bosnia Herzegovina. An important finding of the study is that quitting or low competitiveness in sports predicted smoking initiation in this sample. The study however has some problems: 1- The participants are rather old for studying factors associated with smoking initiation. At age 16 many students have already initiated smoking, which reduces the utility of this sample to study this outcome. 2- No information was gathered on important determinants of smoking (e.g., peer smoking, smoking environment in the school, nicotine dependence) 3- The independent variables are crude and may not actually reflect what intended to be measured (GPA on the last semester only, “no involvement in sport” a student maybe not be involved in formal sports but goes to the gym 5 times a week) 4- Smoking behavior measurements are not standard. 5- The study has limited generalizability: Bosnian students 16-18 years of age. Other points; 1- Page 9, line 19: this is a multi-stage cluster sample, not a simple random one. 2- Page 9, line 20: use” random” instead of “lottery” 3- Page 9, line 38: what is “passive IC means?” 4- Page 9, the part “this was one of the first studies…previously reported.” belongs to discussion. 5- Page 10, lines 36-53: the description of the study sample is long and confusing. A suggestion: “of the 1213 eligible students, 1059 (87%) had complete data in both waves. Of the latter, 872 who identified as Bosnians were included in the study (72% of the total eligible).

	6- Page 17: Calling estimates in Model 1 as crude is misleading. Crude estimates are already provided in the bivariate analysis. In fact, there is no need for Model 1—bivariate analyses and Model 2 are enough. 7- A related issue is Table 3; age and gender seem to be misplaced.

REVIEWER	Filippos Filippidis Imperial College London, United Kingdom
REVIEW RETURNED	12-Nov-2016

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. This is a potentially useful addition to the literature, especially with regards to smoking initiation. The longitudinal aspect of the study is its main strength, so I would suggest that the results on smoking initiation are particularly highlighted. I have some specific comments (see below). The introduction is rather long and could use with some editing in order to improve the flow. The authors do give appropriate background information in general, but the paragraphs are ot connected well and therefore the overall argument why this research is important does not come out strongly. Page 7, line 8. Why “theoreticians”? Is this opinion based more in theory compared to others? Is it supposed to mean that this is not based on empirical evidence? Not clear what the authors mean. Page 7, lines 10-15. IS this the only possible explanation? For example, not doing well in school may be an indication of other issues that might also affect smoking initiation. Or failing at school may reduce self-esteem and make adolescents more vulnerable to peer pressure etc. Page 9, line 24. What does “typical” mean here? Is it fair to assume that they are representative of the country’s population? If so, in what sense? It would be important to demonstrate that findings from these two regions could be generalizable to the entire country. For example, a supplementary table with demographics or socio-economic characteristics comparing the two regions with the entire country could serve this purpose. Page 9, line 41-54. This section seems out of place. It might fit better in the discussion. Page 10, line 32. I think there is something missing in this sentence. It needs some rephrasing in any case. Page 10. Even though I understand the rationale for including only ethnic Bosniaks in the analysis, it would be interesting to run a sensitivity analysis with all ethnicities. It would make results more generalizable both within B&H and in other countries. Table 1 could be a supplementary Page 12, lines 39-44. Not clear how smokers were defined. Those who smoke from time to time were considered smokers? Please clarify. Page 14, line 27. The MW value mentioned for sport achievement is

	not the same as the one shown in table 2. Page 16. I think it is good to show the results of Model 1 in the table, but I wouldn't mention them in the main text. Adjusted results are more relevant, so I'd suggest the authors focus on these in text. Results/Table 3. It is not clear what these ORs refer to. For example, are they shown for a 1-year change in age? Also, for categorical values, I don't think it is appropriate to present one OR. For example, school absence was assessed as almost never, rarely, etc. The change from almost never to rarely is not conceptually the same as the change from "from time to time" to "often". Please clarify what the OR refer to and use appropriate categories for categorical values. Discussion. While the authors adequately explore potential causal pathways (including reverse causality), they do not discuss confounding as a potential explanation for the detected associations. Quitting sports, poor academic performance or/an absence from school might be symptoms of adverse personal or family conditions. They might also show poor parental involvement. These factors, as well as others that were not controlled for (e.g. family income) could be the causes of both absence from school and smoking initiation. I believe that this should be an important part of the discussion. Discussion. Also, it is important for the readers to understand if results can be generalizable to adolescents in other countries. For that reason, it would be useful if some more contextual information is given. The authors do mention social acceptability and overall prevalence of smoking, but I think some more could be useful. For example, is smoking prohibited in schools? Are sales of tobacco products prohibited <18 years? How does the mean age of smoking initiation compare to other European countries? How does the prevalence of smoking among adolescents compare to other countries? Data from GYTS studies, ESPAD surveys and this study (https://www.ncbi.nlm.nih.gov/pubmed/25829499) may provide comparable information. Discussion. I found some results surprising. For example, 2-5 years' experience in sport seemed to be worse than no involvement at all. The authors don't discuss this at all. How would they explain this? If we believe this result, then discouraging adolescents from doing sports might be effective in smoking prevention –which is hard to believe. I would like to see some discussion on this. References. References 3 and 4 don't seem particularly relevant to the actual analysis. Also reference 6 (ESPAD) is not properly presented.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Fawaz Mzayek
Institution and Country: University of Memphis, USA
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

Re: Sport and Scholastic Factors in Relation to Smoking and Smoking Initiation in Older Adolescents: a Prospective Cohort Study in Bosnia and Herzegovina

The study examined longitudinal associations of factors related to academic and sport performance/involvement with smoking behavior in high school students in Bosnia Herzegovina. An important finding of the study is that quitting or low competitiveness in sports predicted smoking initiation in this sample.

The study however has some problems:

1- The participants are rather old for studying factors associated with smoking initiation. At age 16 many students have already initiated smoking, which reduces the utility of this sample to study this outcome.

RESPONSE: Indeed, most of the participants already started to smoke before baseline of this study (i.e. 16 years of age). However, main intention of this study was to prospectively investigate those adolescent groups already reported in cross-sectional studies done the region (it was originally stated at the beginning of the Methods section). Also, and as presented in the paper, a significant proportion of adolescents initiated smoking during the course of the study. However, since you suggested the age as one of the problems it is emphasized as one of the Limitation of this investigation (text now reads: "…the most important limitation comes from the fact that this study observed adolescents from 16 years of age, when many students had already started to smoke. Therefore the generalizability of the findings is limited to older adolescents"; please see page 19; Ln 8-12)

2- No information was gathered on important determinants of smoking (e.g., peer smoking, smoking environment in the school, nicotine dependence)

RESPONSE: Indeed, we did not present these determinants of smoking in the paper, but we have stated that overall permissibility and social-acceptance of smoking in the country is one of the most important reasons for high prevalence of smoking in adolescents, and that it is already established in previous cross-sectional studies (please see page 4, Ln 7-9). Also, the smoking environment and regulative is similar across all studied schools, therefore we believe that it didn't differentially influenced smoking across studies adolescents. In this version of the papers we have added discussion on smoking environment in schools, smoking regulative etc. (text now reads: "Bosnia and Herzegovina is a country which is traditionally oriented toward tobacco consumption since the country was part of the former Ottoman Empire, and tobacco farming has been an important part of the economy in some cantons for more than 300 years (e.g. particularly in the Herzegovina-Neretva Canton, which is the Mediterranean part of the country) [6]. As a result, smoking is socially accepted in public, and cigarettes are relatively cheap. Next, although smoking is prohibited in schools, such regulations are really only imposed for closed high-school buildings. It is probably even more important that smoking is not prohibited in places of social gatherings (e.g. pubs, cafe bars, and disco clubs). Finally, although is formally allowed only for those over 18 years, there is no ID control for purchasing cigarettes, while cigarette vending machines are also common"; please see page 19; Ln 12-22). The factors such as peer smoking, parental smoking had not been observed mainly because we have tried to stay focused on sport and scholastic factors as possible determinants of smoking (smoking initiation primarily), while factors such as peer smoking and smoking environment are frequently studied and reported to be strongly associated to smoking in adolescence. However, we agree that these problems should be highlighted. Therefore, the reasons for not including these variables in this investigation is briefly explained at the end of the Variables subsection (text reads: "As a methodological remark we must note that variables such as peer-smoking and parental-smoking were not included in the study due to their known association with smoking in adolescents [16 27]. Also, there were no differences in anti-smoking regulations among schools, while anti-smoking regulations (i.e. smoking allowance in public, selling tobacco to minors) are equal across the two studied communities"; please see page 10, Ln 4-8)

3- The independent variables are crude and may not actually reflect what intended to be measured (GPA on the last semester only, “no involvement in sport” a student maybe not be involved in formal sports but goes to the gym 5 times a week)
RESPONSE: By all means, the “recreational” (i.e. fitness/gym training) involvement in sport is a specific and highly important problem especially as it related to substance use and misuse. However, in this paper we were mainly focused on “competitive sport”, to extend previous reports on specific associations evidenced between competitive sports and substance misuse. Therefore independent variables were actually focused on these issues (i.e. achievement, experience in years). GPA on the last semester is used as one of scholastic-achievement measures, and we certainly understand your doubts about it (i.e. why didn’t we include previous semesters also?). However, the scholastic system in Bosnia and Herzegovina puts special emphasis on “end of the school year”, and in our case “end of the 2nd high school years is actually “last semester” (i.e. baseline testing was done at the beginning of the 3rd year). At the same time, the half of the school year (i.e. 1st, 3rd, 5th and 7th semester in Bosnia and Herzegovina) is not important with regard to achievement (even when adolescents apply for College/University the achievements at the “end-of-the-school-year(s)” are evidenced (this is clearly stated now; please see page 9 – line 9). Therefore, we believe that this particular GPA accentuates overall scholastic achievement at the moment. Also, GPA is one of three scholastic factors we have observed

4- Smoking behavior measurements are not standard.
RESPONSE: Yes, we agree that the expression of the smoking prevalence is not standard, but we have used the smoking behavior variable used in previous studies done on the territory of former Yugoslavia to assure meaningful comparison with previous data (especially those done 5 years ago in Bosnia and Herzegovina). This is now specified at the beginning of the Variables subsection (text now reads: “To extend current knowledge and allow meaningful comparison with previous cross-sectional reports from countries belonging to the former Yugoslavia, such as Bosnia and Herzegovina, the variables were collected using the Questionnaire of Substance Use, which was previously reported to be a reliable and valid measuring tool in similar samples of participants [6 7 27].”; please see page 9; Ln 7-10).

5- The study has limited generalizability: Bosnian students 16-18 years of age.
RESPONSE: Yes, we agree. It is now stated in the manuscript (text now reads: “Therefore, the generalizability of the findings is somewhat limited (mostly for Bosnian and Herzegovinian adolescents 16-to-18 years of age” ; please see page 19- end, and 20 - beginning)

Other points;
1- Page 9, line 19: this is a multi-stage cluster sample, not a simple random one.
RESPONSE: Corrected, thank you (text reads: “A multi-stage cluster sampling method was used to select the participants.”; please see page 7; line 6)

2- Page 9, line 20: use” random” instead of “lottery”
RESPONSE: Corrected, thank you

3- Page 9, line 38: what is “passive IC means?”
RESPONSE: Corrected, thank you.

4- Page 9, the part “this was one of the first studies…previously reported.” belongs to discussion.
RESPONSE: Avoided in this part of the text as you suggested.

5- Page 10, lines 36-53: the description of the study sample is long and confusing. A suggestion: “of the 1213 eligible students, 1059 (87%) had complete data in both waves. Of the latter, 872 who identified as Bosnians were included in the study (72% of the total eligible).
RESPONSE: Thank you for your suggestion. Actually, we have tried to be as specific as possible, but we agree that it resulted in “overexplanation”. Text now reads as you suggested (please see page 8; line 20-23)

6- Page 17: Calling estimates in Model 1 as crude is misleading. Crude estimates are already provided in the bivariate analysis. In fact, there is no need for Model 1—bivariate analyses and Model 2 are enough.
RESPONSE: Corrected as suggested (please see Legend of Table 2). Also, as 2nd reviewer suggested to include socio-economic-status as potentially confounding variable, the adjusted Model is somewhat changed (Socioeconomic status is included as additional covariate)

7- A related issue is Table 3; age and gender seem to be misplaced.
RESPONSE: Thank you, we corrected it.

Staying at your disposal for any further changes and improvements.
Authors
Reviewer: 2

Reviewer Name: Filippos Filippidis
Institution and Country: Imperial College London, United Kingdom

Thank you for the opportunity to review this manuscript. This is a potentially useful addition to the literature, especially with regards to smoking initiation. The longitudinal aspect of the study is its main strength, so I would suggest that the results on smoking initiation are particularly highlighted. I have some specific comments (see below).
RESPONSE: Thank you for recognizing the potential of our work. We have tried to follow your comments and suggestions. Please see bellow how did we respond to your critics and suggestions.

The introduction is rather long and could use with some editing in order to improve the flow. The authors do give appropriate background information in general, but the paragraphs are not connected well and therefore the overall argument why this research is important does not come out strongly.
RESPONSE: Yes, we must agree that Intro was originally too long. We have tried to explain all aspects of our study, since it was the first prospective investigation of smoking initiation at the territory of former Yugoslavia. We have tried to improve some parts of Introduction while being more focused (please see highlighted part of the text within Introduction).

Page 7, line 8. Why “theoreticians”? Is this opinion based more in theory compared to others? Is it supposed to mean that this is not based on empirical evidence? Not clear what the authors mean.
RESPONSE: Thank you for noticing it, it was a mistake. Term is changed, and some additional references that directly investigated this problem are added. Please see 2nd paragraph of the Introduction, page 4 (end) – page 5 (beginning).

Page 7, lines 10-15. Is this the only possible explanation? For example, not doing well in school may

be an indication of other issues that might also affect smoking initiation. Or failing at school may reduce self-esteem and make adolescents more vulnerable to peer pressure etc.
RESPONSE: We are particularly grateful for this comment! Actually, in writing Introduction we have been focused on previous reports and therefore named only those “theories” mentioned in research done on the territory. However, in revised version of the manuscript the problem is explained more profoundly.Text reads: “However, it is also possible that other factors, such as parental conflict, and/or poor familiar control, result in both educational failure and smoking. Additionally, the association is in some cases explained by the “theory of problem behaviour” (i.e. that the problem behaviours such as failure in school and smoking often appear in tandem because some people have a psychosocial tendency for unconventionality) [17].” (please see page 5; lines 1-5)

Page 9, line 24. What does “typical” mean here? Is it fair to assume that they are representative of the country’s population? If so, in what sense? It would be important to demonstrate that findings from these two regions could be generalizable to the entire country. For example, a supplementary table with demographics or socio-economic characteristics comparing the two regions with the entire country could serve this purpose.
RESPONSE: Thank you for noticing it. It actually allowed us to explain one of the most issues of selecting these two Cantons as “typical” for Bosnia and Herzegovina. It is highly specific problem and mostly is connected with 90s wars on the territory. We have provided a brief explanation on it. Please see page 7; line 9-12 (“Bosnia and Herzegovina is a multi-ethnic country, home to three constitutive ethnicities (Bosniaks, Serbs and Croats). Devastating wars that occurred in early 90s resulted in massive emigrations of minority ethnic groups (specifically for different parts of the country), and overall material devastation [29]. For the two Cantons studied, pre-war ethnic figures did not change drastically. Therefore, these two Cantons should be observed as two typical regions in Bosnia and Herzegovina.”) (please see page 7; lines 9-14).

Page 9, line 41-54. This section seems out of place. It might fit better in the discussion.
RESPONSE: The section is removed

Page 10, line 32. I think there is something missing in this sentence. It needs some rephrasing in any case.
RESPONSE: Sentence is rephrased. Text now reads: “The study fulfilled all ethical guidelines and received the approval of the Ethical Boards from the University of Mostar, Bosnia and Herzegovina, and the University of Split, Croatia. After obtaining ethical approvals, the study was officially authorised by the Ministries of Education in Zenica-Doboj Canton and Tuzla Canton, the two areas of Bosnia and Herzegovina where the research was taking place.” (please see page 8; ln 11-16)

Page 10. Even though I understand the rationale for including only ethnic Bosniaks in the analysis, it would be interesting to run a sensitivity analysis with all ethnicities. It would make results more generalizable both within B&H and in other countries.
RESPONSE: Thank you for recognizing the importance of stratification of participants. We certainly agree with your observation that analysis including all ethnicities will be more representative and generalizable. For that purpose we are currently performing similar investigation in other parts of country. Namely, we will try to collect data on adolescents relative to overall population according to recent population census, and hope to present it soon.

Table 1 could be a supplementary
RESPONSE: As you suggested Table 1 is provided as supplementary material

Page 12, lines 39-44. Not clear how smokers were defined. Those who smoke from time to time were considered smokers? Please clarify.
RESPONSE: Corrected and clarified. Text now reads: Cigarette smoking was assessed on a four-

point scale with the following responses: “No, I don’t smoke”, “From time to time, but not daily”, “Less than 10 cigarettes daily”, and “More than 10 cigarettes daily”. Participants were later classified as non-smokers (those who responded “No, I don’t smoke)”or smokers (the remaining three answers).(Page 9, line 21-24)

Page 14, line 27. The MW value mentioned for sport achievement is not the same as the one shown in table 2.
RESPONSE: Corrected, thank you, it was technical mistake

Page 16. I think it is good to show the results of Model 1 in the table, but I wouldn’t mention them in the main text. Adjusted results are more relevant, so I’d suggest the authors focus on these in text.
RESPONSE: Corrected as suggested. We have omitted results of the “Model 1” in the Results, but retained it in the Table (now Table 1). Also, and as you suggested controlling the logistic regression for additional confounders (socio economic status) the logistic regressions are re-calculated (please see later comments and responses).

Results/Table 3. It is not clear what these ORs refer to. For example, are they shown for a 1-year change in age? Also, for categorical values, I don’t think it is appropriate to present one OR. For example, school absence was assessed as almost never, rarely, etc. The change from almost never to rarely is not conceptually the same as the change from “from time to time” to “often”. Please clarify what the OR refer to and use appropriate categories for categorical values.
RESPONSE: As you suggested, in this version of the manuscript we have re-calculated and presented ORs for “school absence” as categorical variable. Accordingly, the Results are rewritten in some parts. Please see Table 1 and Results (beginning of the Section) for more details.

Discussion. While the authors adequately explore potential causal pathways (including reverse causality), they do not discuss confounding as a potential explanation for the detected associations. Quitting sports, poor academic performance or/an absence from school might be symptoms of adverse personal or family conditions. They might also show poor parental involvement. These factors, as well as others that were not controlled for (e.g. family income) could be the causes of both absence from school and smoking initiation. I believe that this should be an important part of the discussion.
RESPONSE: Yes, Discussion originally lack interpretation of potential confounding factors. However, as you suggested to include socioeconomic status (SES), while we actually observed this variable (did not presented it in the first version of the paper) we have included SES as confounding factor and re-calculated logistic regression (please see text on Statistics, and Table 2). Also, smokers and nonsmokers were compared on SES (please see Table 1 – highlighted text). Namely, in one of our first calculations, the SES was observed as confounding variable, but since it did not change overall regression models we decided not to present it. We agree that it is an interesting issue, so consequently followed your suggestion and included it in this version of the manuscript. Also, SES is (now) included in the data set provided as supplementary material.

Discussion. Also, it is important for the readers to understand if results can be generalizable to adolescents in other countries. For that reason, it would be useful if some more contextual information is given. The authors do mention social acceptability and overall prevalence of smoking, but I think some more could be useful. For example, is smoking prohibited in schools? Are sales of tobacco products prohibited <18 years? How does the mean age of smoking initiation compare to other European countries? How does the prevalence of smoking among adolescents compare to other countries? Data from GYTS studies, ESPAD surveys and this study (https://www.ncbi.nlm.nih.gov/pubmed/25829499) may provide comparable information.
RESPONSE: Thank you for this suggestion since it “allowed” us to explain socio-cultural context the results more specifically (text reads: “Meanwhile, in order to objectively overview the findings, some

specific contextual information on the socio-cultural environments is necessary. Bosnia and Herzegovina is a country which is traditionally oriented toward tobacco consumption since the country was part of the former Ottoman Empire, and tobacco farming has been an important part of the economy in some cantons for more than 300 years (e.g. particularly in the Herzegovina-Neretva Canton, which is the Mediterranean part of the country) [6]. As a result, smoking is socially accepted in public, and cigarettes are relatively cheap. Next, although smoking is prohibited in schools, such regulations are really only imposed for closed high-school buildings. It is probably even more important that smoking is not prohibited in places of social gatherings (e.g. pubs, cafe bars, and disco clubs). Finally, although is formally allowed only for those over 18 years, there is no ID control for purchasing cigarettes, while cigarette vending machines are also common."; please see page 19, line 12-22). Unfortunately, to the best of our knowledge, exact age of smoking initiation for BH is not known (at least, it is not empirically evidenced). Therefore, we were able to report it roughly (according to our study; 2/3 of adolescents started to smoke before 16 years of age).

Discussion. I found some results surprising. For example, 2-5 years' experience in sport seemed to be worse than no involvement at all. The authors don't discuss this at all. How would they explain this? If we believe this result, then discouraging adolescents from doing sports might be effective in smoking prevention –which is hard to believe. I would like to see some discussion on this.
RESPONSE: Indeed, the finding on higher odds for smoking initiation in those who were engaged in sport for less than 5 years (when compared to adolescents who were never involved in sports) is somewhat surprising. However, it generally fits into other findings, especially the one that "quitting sports" is predictor of smoking initiation between 16 and 18 years of age. Namely, those children who quitted sports by the age of 16 have been involved in sports for some time, but either (1) started to practice sports relatively late, or (2) quitted early. Both resulted in short period of time being involved in sports (less than 5 years). This is more precisely explained in the text now. Please see paragraph of text starting at page 17 (line 5), and ending at page 18 (line 14).

References. References 3 and 4 don't seem particularly relevant to the actual analysis. Also reference 6 (ESPAD) is not properly presented.
RESPONSE: Corrected as suggested (now Reference 4)

Thank you for your profound review.
Staying at disposal for any further changes and improvements

VERSION 2 – REVIEW

REVIEWER	Fawaz Mzayek, MD, PhD University of Memphis, USA
REVIEW RETURNED	20-Jan-2017

GENERAL COMMENTS	The manuscript is improved considerably and the authors responded adequately to the majority of the points raised by this reviewer. However, two issues are not resolved completely. 1- "No information was gathered on important determinants of smoking (e.g., peer smoking, smoking environment in the school, nicotine dependence)" The authors responded that the smoking atmosphere and enforcement of anti-smoking laws are rather homogenous across the study sample. While this may be true on the group-level, individual-level information is still important for such important predictors as peer-smoking and nicotine dependence.

	The authors go on to state that they did not collect such information because they wanted to stay focused on their research question "competitive sports and scholastic performance as predictors of smoking initiation". While a focused research question is a strength in any study, the goal of epidemiological research is to characterize the association between the hypothesized exposure/s and the outcome as close to the truth as possible. This requires exploring the possible effect of potential confounding and/or effect-modifying factors on the studied association. The authors state that they did not include factors such as peer- and parental-smoking "due to their known association with smoking in adolescents". One can argue that this actually a reason for measuring and adjusting for these factors in this study, especially if they are also associated with sports participation and scholastic achievement. 2- The independent variables are crude and may not actually reflect what intended to be measured (GPA on the last semester only, "no involvement in sport" a student maybe not be involved in formal sports but goes to the gym 5 times a week) The authors responded that they concentrated only on competitive sports and not physical activity in general, and that last-semester GPA in this case is the end-of-schoolyear GPA and that the educational authorities in Bosnia-Herzegovina put special emphasis on this score. The fact that the authorities put especial emphasis on last semester GPA, does not make it a precise and accurate marker of the student scholastic performance. The same can be said about other aspects of physical activity, like regularly engaging in some sort of training, despite non-formally. So, while these factors were measured, they can still be considered crude measurements that may have resulted in misclassification bias. In conclusion, these two points remain as potential weaknesses in the study

REVIEWER	Filippos Filippidis Imperial College London, UK
REVIEW RETURNED	19-Dec-2016

GENERAL COMMENTS	The authors have done a good job addressing the reviewers' comments. Some limitations could not be addressed, such as unmeasured confounding and limited generalizability of the results to other populations, which somewhat limits the value of the study. However, these limitations have been clearly explained in the text. Some minor comments follow: Note: the page numbers refer to the version with the highlighted text and the line numbers refer to the first set of numbers (on the left hand side of the page). Introduction, Page 39, lines 8-13. The language is a bit strong here. I would suggest using phrases such as "the association could be explained by the theory…" or "the theory of planned behaviour has been used to explain…". Methods, Page 44, lines 13-18. I am not convinced by this

	argument. Peer-smoking and parental-smoking may be potential confounding factors here if they are also associated with sports participation and scholastic performance. In that case, they should be included and the fact that they are not is a study limitation. I would expect an explanation showing that these factors are not associated with sports participation and/or scholastic achievement (i.e. are not confounding factors). If there is no such explanation, then this comment should be moved to the limitations of the study. Table 2. Move the reference category on top in each variable.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Filippos Filippidis
Institution and Country: Imperial College London, UK
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
The authors have done a good job addressing the reviewers' comments. Some limitations could not be addressed, such as unmeasured confounding and limited generalizability of the results to other populations, which somewhat limits the value of the study. However, these limitations have been clearly explained in the text. Some minor comments follow:

Introduction, Page 39, lines 8-13. The language is a bit strong here. I would suggest using phrases such as "the association could be explained by the theory…" or "the theory of planned behaviour has been used to explain…".
RESPONSE: Thank you for this suggestion. We have used the phrasing suggested. Please see highlighted text.

Methods, Page 44, lines 13-18. I am not convinced by this argument. Peer-smoking and parental-smoking may be potential confounding factors here if they are also associated with sports participation and scholastic performance. In that case, they should be included and the fact that they are not is a study limitation. I would expect an explanation showing that these factors are not associated with sports participation and/or scholastic achievement (i.e. are not confounding factors). If there is no such explanation, then this comment should be moved to the limitations of the study.
RESPONSE: To the best of our knowledge no study directly examined the association between peer/parental-smoking with smoking and/or sport participation. Therefore, we have not been able to directly specify it, and as you suggested, your comment is included in the study limitations (please see highlighted text at the end of the manuscript, but also section "Strength and limitations of the study" at the beginning of the paper). Also, limitations are broaden according to suggestions of 1st Reviewer.

Table 2. Move the reference category on top in each variable.
RESPONSE: REF category is placed on top for each categorical variable

Reviewer: 1
Reviewer Name: Fawaz Mzayek, MD, PhD
Institution and Country: University of Memphis, USA
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

The manuscript is improved considerably and the authors responded adequately to the majority of the points raised by this reviewer. However, two issues are not resolved completely.

1- “No information was gathered on important determinants of smoking (e.g., peer smoking, smoking environment in the school, nicotine dependence)”
The authors responded that the smoking atmosphere and enforcement of anti-smoking laws are rather homogenous across the study sample. While this may be true on the group-level, individual-level information is still important for such important predictors as peer-smoking and nicotine dependence.
The authors go on to state that they did not collect such information because they wanted to stay focused on their research question “competitive sports and scholastic performance as predictors of smoking initiation”. While a focused research question is a strength in any study, the goal of epidemiological research is to characterize the association between the hypothesized exposure/s and the outcome as close to the truth as possible. This requires exploring the possible effect of potential confounding and/or effect-modifying factors on the studied association. The authors state that they did not include factors such as peer- and parental-smoking “due to their known association with smoking in adolescents”. One can argue that this actually a reason for measuring and adjusting for these factors in this study, especially if they are also associated with sports participation and scholastic achievement.
RESPONSE: Thank you for your observation. We truly appreciate it a low suggestion. Please see bellow next comment for response.

2- The independent variables are crude and may not actually reflect what intended to be measured (GPA on the last semester only, “no involvement in sport” a student maybe not be involved in formal sports but goes to the gym 5 times a week)

The authors responded that they concentrated only on competitive sports and not physical activity in general, and that last-semester GPA in this case is the end-of-schoolyear GPA and that the educational authorities in Bosnia-Herzegovina put special emphasis on this score.

The fact that the authorities put especial emphasis on last semester GPA, does not make it a precise and accurate marker of the student scholastic performance. The same can be said about other aspects of physical activity, like regularly engaging in some sort of training, despite non-formally. So, while these factors were measured, they can still be considered crude measurements that may have resulted in misclassification bias.

In conclusion, these two points remain as potential weaknesses in the study

RESPONSE: First of all, we must thank you for this and previous review. It allowed us to improve the manuscript, but also (and even more important) – to learn a lot. We have no doubt that your comments are appropriate, and must say that in future investigations (including the one we are starting right now, with younger adolescents, 14-16 yr old) we will follow your suggestions with regard to evaluate factors you have specified. For a moment, these factors were not studied herein, and we have no other option but to mention it as a study limitations. Therefore, in this version of the manuscript these weaknesses were specified in a study limitations (end of Discussion), but also at the beginning of the manuscript (heading “Study limitations and strengths”).

VERSION 3 – REVIEW

REVIEWER	Filippos Filippidis Imperial College London, United Kingdom
REVIEW RETURNED	30-Jan-2017

GENERAL COMMENTS	No further comments.