

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)	Education Against Tobacco (EAT): a quasi-experimental prospective evaluation of a multinational medical-student-delivered smoking prevention programme for secondary schools in Germany
AUTHORS	Brinker, Titus; Stamm-Balderjahn, Sabine; Werner, Seeger; Klingelhöfer, Doris; Groneberg, David

VERSION 1 - REVIEW

REVIEWER	Matthis Morgenstern Institute for Therapy and Health Research, Kiel, Germany Geisel School of Medicine at Dartmouth, Lebanon, USA
REVIEW RETURNED	28-Apr-2015

GENERAL COMMENTS	<p>This is a paper on the effects of the medical-student-led program EAT on child or adolescent smoking behavior. While I think that the intervention is a great approach (doing two things at the same time: educating kids and engaging medical students) and that the implementation in Germany can be seen as a success story, I also believe that the paper needs significant improvement. This refers to the structure of the paper as well as the statistical analysis and the description of the results.</p> <p>Major points</p> <p>Introduction: I suggest to clearly separate the description of the intervention from the review of previous literature. The paragraph "Intervention" in the Methods section might be a good place to describe the intervention.</p> <p>Introduction: The overlap with the study protocol is quite high, the introduction should therefore be more concise.</p> <p>Methods: The study needs a clearer definition of the study sample. The authors might think of using the paragraph "Participants" and add the paragraph "Recruitment" from the Results. More important, they should early decide on the analytic sample, this gets completely mixed-up throughout the paper. Sometimes it is the 1474 of the baseline sample, sometimes it is the 1200 with complete data. The best way is to impute the missing data at t2 (using some imputation technique like MI or ML). If the authors decide to use listwise deletion, the weaker approach, then still all presented results, also the baseline results, should be on the same sample, i.e. 1200. The authors should then add a paragraph "attrition analysis" and describe the students lost to follow-up.</p> <p>Methods: There is also a mix-up in the description of the data assessment procedure and the actual measurement contents. The</p>
-------------------------	---

	<p>contents of the questionnaire are not well described, e.g., it is hard to get a feel for how the smoking behavior has been assessed.</p> <p>Statistical Analysis: The sample size calculation can be omitted, given the study protocol.</p> <p>Statistical Analysis: The longitudinal regression analysis is also not clearly described. How have the baseline differences between intervention and control group been adjusted for?</p> <p>Statistical Analysis: If a logistic regression is the adequate procedure depends on the scaling of the outcome, other regression models are also "state of the art", of course. In this case, the group allocation was not on the individual level, but on the class level. It is therefore necessary to consider this clustering statistically, either using cluster-robust standard errors or multilevel modeling.</p> <p>Results: Attrition Analysis. The proper test of differential drop-out is the interaction between group status and smoking status.</p> <p>Results: Table 1 should use the complete-cases sample (see above). Above that, it is not easy to understand. Why are the quitters + never smokers in sum not the non-smokers? Why using different denominators for the percentages? It seems, for example, more important to know what the proportion of daily smoking is in the sample than to know which proportion of the monthly-daily smokers are daily smokers.</p> <p>Results: The caption of Table 2 is hard to understand. And why are these "secondary outcomes"? The rate of smokers at t2 compared to t1 is the primary outcome, right? It would be nice to see the actual prevalences at t1 and t2 in a table, the ones that were used to test for significance.</p> <p>Results: There is too much emphasis on the sub-sample analysis of 67 smokers. While this is surely worthy of reporting, the authors should have a higher focus on properly handling the design issues in their primary analysis.</p> <p>Discussion and Intro: I think the distinction between primary and secondary prevention is overly technical here and it would be better and clearer to just describe the outcomes under study (smoking prevalence, smoking initiation, smoking cessation)</p> <p>Discussion: Further, the emphasis on "cessation" effects of the intervention is misleading and it is questionable if this is a proper term for 11-15 year-olds, anyway. The analysis shows that it is mostly the younger kids that "stopped" smoking, which for many of them means they stopped progressing from experimental to established smoking. To be clear about it: This comment is about terminology, I am not suggesting that the finding itself is not important.</p> <p>Conclusion: As far as I understand the analysis, there was an increase in the smoking rate in the control group, but not in the intervention group. Why do the authors claim they "only" found a secondary preventive effect? But maybe I am just confused. The test which produced the significant p-value is not well explained.</p> <p>Minor points:</p>
--	--

	<ul style="list-style-type: none"> - Title: "the largest" is a bit misleading as there are no other medical-student delivered interventions - Abstract uses two different age ranges - Abstract reads as if there were two measurements at baseline - Abstract: "No primary preventive effect was measured" means no effect was "found"? - "paper pencil" questionnaire is easier to understand than "written" - References: Some references do not match very well. This is true, for example, for the claim that most smokers start before adolescence. Also the references regarding logistic regression analysis.
--	---

REVIEWER	Laetitia Minary EA APEMAC, Lorraine University, France
REVIEW RETURNED	05-May-2015

GENERAL COMMENTS	<p>This manuscript describes a study which evaluates the effectiveness of a smoking prevention programme delivered by medical-students among a population of secondary schools student. Results revealed an effectiveness of the program at 6 months follow-up concerning smoking cessation.</p> <p>Some modifications to the manuscript could further strengthen it, as described below.</p> <p>Major compulsory revisions:</p> <p>Introduction: Page 4 line 22: in which population? Page 4 line 27: The transition between the two last sentences is difficult. Page 4 lines 34 to 43: I don't understand why only the smoke-free class competition program is cited. It is not representative of other existing programs Page 5 line 8: This paragraph should start with a brief description of the EAT program. Page 5 lines 28 to 48: It is very surprising that the dissemination process precedes the evaluation. This paragraph should rather be presented in the discussion section. Page 5 line 51 to page 6 line 26: These informations should be mentioned earlier in the text. Furthermore, the innovative aspect of the present study in comparison to the previous evaluative studies is insufficiently developed.</p> <p>Objectives: The objectives should be specified by defining the target population and the terms primary and secondary prevention (initiation prevention and smoking cessation). The definition of the primary endpoint is not clear. Effectiveness should be preferred to efficacy.</p> <p>Methods: Page 7 line 9: Before explaining the modalities of allocation of the intervention, the authors should describe the schools recruitment process: How many schools were eligible? How many participated to the program?</p>
-------------------------	--

	<p>Page 8 line 40: The authors should precise the setting of the data collection (in the classroom).</p> <p>Page 9 line 24: What is the data circuit between data collection and data entry?</p> <p>Sample size calculation: What are the underlying assumptions (expected difference?)?</p> <p>Analysis: The primary outcome analysis is not described.</p> <p>Results: This section should be reorganized to improve its reading.</p> <p>Page 10 line 54: What about other characteristics which could be associated to tobacco consumption (gender, age)?</p> <p>Page 13 lines 8 to 33: This paragraph is not clear. We don't understand where the 67 smokers are coming from (values are different in table 1).</p> <p>The primary outcome must be specified (what is associated to $p < 0,01$?): Is this the prevalence of smokers or prospective smoking status? If it is the prospective smoking status is this comparison adjusted on initial difference between the 2 groups (age, schooltype etc...)? If so, it must be describe in section analysis.</p> <p>This major results should be presented in the text (example: XX% were smokers in the intervention group at 6 months vs. XX% in the control group, $p = XXX$) and table 2 should be presented after the paragraph to avoid confusion.</p> <p>Page 13 line 31: p value should be presented after "no primary preventive effect was measured".</p> <p>Legal approval: This chapter should be integrated in the method section.</p> <p>Discussion Page 15 lines 8 to 19: this paragraph is redundant with introduction and should be removed.</p> <p>Page 16 : Interpretation : This section should appear earlier in the text (before limitations).</p> <p>P 16 line 53:"The discussed selection bias may have negatively affected our results for primary prevention": this assumption should be developed.</p> <p>Minor essential revisions: Page line 47: "At the same time" may be better as "in addition"</p> <p>Page 4 line 15: "These aspects [...] facilitate school-based prevention": In what way?</p> <p>Page 10 line 54: "For" smoking status?</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: Matthis Morgenstern

Thank you for helping us to improve our manuscript which is highly appreciated and valued.

Introduction: I suggest to clearly separate the description of the intervention from the review of previous literature. The paragraph "Intervention" in the Methods section might be a good place to describe the intervention.

Reply: The intervention is now described in the Methods section as suggested.

Introduction: The overlap with the study protocol is quite high, the introduction should therefore be more concise.

Reply: The introduction was shortened.

Methods: The study needs a clearer definition of the study sample. The authors might think of using the paragraph "Participants" and add the paragraph "Recruitment" from the Results.

Reply: We agreed with your suggestion and added the Recruitment paragraph to the participants section.

More important, they should early decide on the analytic sample, this gets completely mixed-up throughout the paper. Sometimes it is the 1474 of the baseline sample, sometimes it is the 1200 with complete data. The best way is to impute the missing data at t2 (using some imputation technique like MI or ML).

Reply: We see your point and thus went through the paper to add clarification to these numbers: 1,474=baseline sample; 274=loss-to-follow-up effect; 1,200=analytic sample. We used listwise deletion.

If the authors decide to use listwise deletion, the weaker approach, then still all presented results, also the baseline results, should be on the same sample, i.e. 1200. The authors should then add a paragraph "attrition analysis" and describe the students lost to follow-up.

Reply: To our knowledge, the cross-sectional baseline data analysis is conventionally performed on the basis of the full baseline data and independent from the longitudinal analysis (=analytic sample). This can be found in several peer-reviewed papers (1, 2). The attrition analysis (which we have performed) then investigates the drop-outs.

Methods: There is also a mix-up in the description of the data assessment procedure and the actual measurement contents. The contents of the questionnaire are not well described, e.g., it is hard to get a feel for how the smoking behavior has been assessed.

Reply: We took items from published questionnaires (Students in the Hospital and ohnekippe) and pretested it multiple times within our age group in accordance to the guidelines of Good Epidemiologic Practice and in consequence slightly adapted the questionnaire. We quantified cigarette and water pipe consumption by asking how many cigarettes are smoked daily/weekly/monthly/per year.

Statistical Analysis: The sample size calculation can be omitted, given the study protocol.

Reply: We omitted the sample size calculation.

Statistical Analysis: The longitudinal regression analysis is also not clearly described. How have the baseline differences between intervention and control group been adjusted for?

Reply: There were no systematic differences between the two groups at baseline in terms of smoking status. The robust panel logistic regression analysis analyzed the change in smoking status t1->t2 (smoker to non-smoker) which automatically adjusts for the baseline data We edited the statistical analysis chapter and added extra descriptions.

Statistical Analysis: If a logistic regression is the adequate procedure depends on the scaling of the outcome, other regression models are also "state of the art", of course. In this case, the group allocation was not on the individual level, but on the class level. It is therefore necessary to consider this clustering statistically, either using cluster-robust standard errors or multilevel modeling.

Reply: We agree with you and analyzed our data again via robust panel logistic regression (see manuscript).

Results: Attrition Analysis. The proper test of differential drop-out is the interaction between group status and smoking status.

Reply: Thank you for your advice which is very helpful. We did the test as suggested.

Results: Table 1 should use the complete-cases sample (see above).

Reply: We performed baseline analysis in accordance to the cited publications above.

Above that, it is not easy to understand. Why are the quitters + never smokers in sum not the non-smokers?

Reply: This is a limitation of the questionnaire/the target group: The missing quitters and ex-smokers are non-smokers who did not further define in the questionnaire if/when they quitted smoking. Thus, we can not put them in a subcategory of non-smokers.

Why using different denominators for the percentages? It seems, for example, more important to know what the proportion of daily smoking is in the sample than to know which proportion of the monthly-daily smokers are daily smokers.

Reply: This is a convention which can be found in many other publications of this type. The smokers are a subgroup and have been predefined in the study protocol as cigarette/waterpipe smokers who smoke at least once a month. The denominators are used to describe this subgroup most lucid. The same technique was used for the non-smokers subgroup and therefore the denominators are not different in the subgroup analysis.

Results: The caption of Table 2 is hard to understand. And why are these "secondary outcomes"? The rate of smokers at t2 compared to t1 is the primary outcome, right? It would be nice to see the actual prevalences at t1 and t2 in a table, the ones that were used to test for significance.

Reply: This is now Table 3 as the manuscript was edited. You are correct that it is possible to calculate the primary endpoint from Table 3 by simply adding up the numbers (changes in prevalence from baseline to endline in both groups) but it also shows how the prevalence is changing (longitudinal changes in smoking behavior = secondary outcomes). We added Table 2 for the primary endpoint analysis. We calculated the primary endpoint as requested separately in the manuscript.

Results: There is too much emphasis on the sub-sample analysis of 67 smokers. While this is surely worthy of reporting, the authors should have a higher focus on properly handling the design issues in their primary analysis.

Reply: Please specify which design issues you mean. It is a quasi-experimental study – multiple biases are discussed in the discussion section. In our opinion, the results part should exclusively report the results.

Discussion and Intro: I think the distinction between primary and secondary prevention is overly technical here and it would be better and clearer to just describe the outcomes under study (smoking prevalence, smoking initiation, smoking cessation)

Reply: We changed this as suggested.

Discussion: Further, the emphasis on "cessation" effects of the intervention is misleading and it is questionable if this is a proper term for 11-15 year-olds, anyway. The analysis shows that it is mostly the younger kids that "stopped" smoking, which for many of them means they stopped progressing from experimental to established smoking. To be clear about it: This comment is about terminology, I am not suggesting that the finding itself is not important.

Reply: We understand your point but can not think of a more appropriate term as the onset of addiction happens faster at young age.

Conclusion: As far as I understand the analysis, there was an increase in the smoking rate in the control group, but not in the intervention group. Why do the authors claim they "only" found a secondary preventive effect? But maybe I am just confused. The test which produced the significant p-value is not well explained.

Reply: Some pupils started and some pupils quit which is well depicted in Table 3. Overall there were no significant differences of students who started smoking. However, more pupils quitted in the intervention group and thus the percentage of smokers remains almost the same in the intervention group.

Minor points:

- Title: "the largest" is a bit misleading as there are no other medical-student delivered interventions

Reply: There are other medical student delivered interventions for example in the US but no evaluated ones and no multinational networks like Education Against Tobacco. We changed the term to multinational as this is more precise.

- Abstract uses two different age ranges

Reply: 10-15 year old participants were eligible when we planned the study. However, exclusively 11-15 year old pupils were recruited.

Abstract reads as if there were two measurements at baseline

Reply: We erased the word „two“ from the Abstract.

Abstract: "No primary preventive effect was measured" means no effect was "found"?

Reply: We exchanged the word "measured" with "found".

- "paper pencil" questionnaire is easier to understand than "written"

Reply: This was exchanged as suggested.

- References: Some references do not match very well. This is true, for example, for the claim that most smokers start before adolescence. Also the references regarding logistic regression analysis.

Reply: We went through the references again and exchanged or deleted them when they were deemed to be not specific enough or of low quality.

1.

Pbert L, Flint AJ, Fletcher KE et al.: Effect of a pediatric practice-based smoking prevention and cessation intervention for adolescents: a randomized, controlled trial. *Pediatrics* 2008, 121: 738-47.

2.

Lear, S.A., Ignaszewski, A., Linden et al.: The Extensive Lifestyle Management Intervention (ELMI) following cardiac rehabilitation trial. *Eur Heart J* 2003, 24:1920-27.

Reviewer: Laetitia Minary

Thank you for helping us to improve our manuscript which is highly appreciated and valued.

Page 4 line 22: in which population?

Reply: In German adolescents (added).

Page 4 line 27: The transition between the two last sentences is difficult.

Reply: We deleted the last sentence.

Page 4 lines 34 to 43: I don't understand why only the smoke-free class competition program is cited. It is not representative of other existing programs

Reply: This is true but it is the most popular program in Germany and to our knowledge even in the European Union. As the Journal asks us to be concise we could not list all the available programs. Despite proven ineffectiveness SFC is still broadly implemented.

Page 5 line 8: This paragraph should start with a brief description of the EAT program.

Reply: The other reviewer urged us to put the full description of the EAT program into the Methods section (Intervention).

Page 5 lines 28 to 48: It is very surprising that the dissemination process precedes the evaluation. This paragraph should rather be presented in the discussion section.

Reply: We embedded this paragraph into the discussion section as discussed.

Page 5 line 51 to page 6 line 26: Theses informations should be mentioned earlier in the text. Furthermore, the innovative aspect of the present study in comparison to the previous evaluative studies is insufficiently developed.

Reply: We present this information in a more concise way earlier in the text as suggested. The other

reviewer urged us to shorten the Background section as much of it was already mentioned in the study protocol.

Objectives:

The objectives should be specified by defining the target population and the terms primary and secondary prevention (initiation prevention and smoking cessation).

Reply: We specified these terms as suggested whenever they were mentioned in the text.

The definition of the primary endpoint is not clear.

Reply: We specified the definition of the primary endpoint: The primary endpoint was defined as the difference from t1 to t2 of the smoking prevalence in the control group vs. the difference from t1 to t2 in the intervention group of the smoking prevalence (difference of differences approach) with $p < 0.05$.

Effectiveness should be preferred to efficacy.

Reply: We exchanged the word.

Methods:

Page 7 line 9: Before explaining the modalities of allocation of the intervention, the authors should describe the schools recruitment process: How many schools were eligible? How many participated to the program?

Reply: 8 schools were eligible and were asked and all of them participated in the program. We added this information in the manuscript.

Page 8 line 40: The authors should precise the setting of the data collection (in the classroom).

Reply: We added this information in the text.

Page 9 line 24: What is the data circuit between data collection and data entry?

Reply: The class teachers sealed the envelopes and shipped them to the institute.

Sample size calculation:

What are the underlying assumptions (expected difference?)?

Reply: The sample size calculation was described in detail in our study protocol. Thus, the other reviewer urged us to omit it from this publication which we did.

Analysis:

The primary outcome analysis is not described.

Reply: We added the primary outcome analysis description (Methods section) and a table for result presentation (Results section).

Results: This section should be reorganized to improve its reading.

Reply: Thank you for helping us to improve our manuscript.

Page 10 line 54: What about other characteristics which could be associated to tobacco consumption (gender, age)?

Reply: We calculated both interaction effects via logistic regression and age was significant meaning that in the intervention group more young students dropped out whereas in the control group more older students dropped out ($p=0.045$). The calculation was added to the manuscript and interpreted in the Discussion section.

Page 13 lines 8 to 33: This paragraph is not clear. We don't understand where the 67 smokers are coming from (values are different in table 1). Is this the prevalence of smokers or prospective smoking status? If it is the prospective smoking status is this comparison adjusted on initial difference between the 2 groups (age, schooltype etc...)? If so, it must be describe in section analysis.

Reply: It is the prospective (analytical) sample. The baseline differences are automatically taken into account in robust panel logistic regression analysis of the primary endpoint: Other differences can be regarded as limitations of our quasi-experimental design and are discussed in the discussion section.

The primary outcome must be specified (what is associated to $p<0,01$):

Reply: The $p=0.01$ is associated to Table 2 (double-sided p-value). In our sample the group allocation was not on the individual level but on the class level. In order to take into account this clustering statistically we used robust panel logistic regression (xtlogit procedure with vce (cluster) option). This procedure was also used to calculate the difference from t1 to t2 of the smoking prevalence in the control group versus the difference from t1 to t2 in the intervention group (our primary endpoint) by the help of STATA 14 by StataCorp (Texas, USA).

This major results should be presented in the text (example: XX% were smokers in the intervention group at 6 months vs. XX% in the control group, $p=XXX$) and table 2 should be presented after the paragraph to avoid confusion.

Reply: We presented the major result before table 3 (former table 2) to avoid confusion as suggested.

Page 13 line 31: p value should be presented after "no primary preventive effect was measured".

Reply: We did not calculate a p-value as obviously the same percentage of participants started smoking in both groups (5.0%; see Table 3). However, we added this information in the text prior to Table 3.

Legal approval:

This chapter should be integrated in the method section.

Reply: We integrated this chapter into the method section as suggested.

Discussion

Page 15 lines 8 to 19: this paragraph is redundant with introduction and should be removed.

Reply: The other reviewer commented on this aspect in the Background section. The paragraph has been rephrased and is not redundant anymore.

Page 16 : Interpretation : This section should appear earlier in the text (before limitations).

Reply: We followed your suggestion and put the Interpretation chapter prior to the Limitations chapter.

P 16 line 53: "The discussed selection bias may have negatively affected our results for primary prevention": this assumption should be developed.

Reply: We developed this assumption at the beginning of the Limitations section.

Minor essential revisions:

Page line 47: "At the same time" may be better as "in addition"

Reply: Thank you for helping us to improve the language which we highly appreciate. We think the same way and exchanged the wording.

Page 4 line 15: "These aspects [...] facilitate school-based prevention": In what way?

Reply: Medical students train themselves in their didactic skills when participating in our program. The medical students who participate frequently were observed to be more confident in front of pupils and were observed to be better at explaining medical complexities in layman's terms which indeed facilitates qualitative primary prevention during and after medical school in physician-delivered programs.

Page 10 line 54: "For" smoking status?

Reply: We rephrased the whole sentence and think it is more understandable and precise now.

VERSION 2 - REVIEW

REVIEWER	Matthis Morgenstern Geisel School of Medicine at Dartmouth USA
REVIEW RETURNED	26-Jun-2015

GENERAL COMMENTS	<p>The authors really worked a lot on the revision, and I see clear improvement, especially regarding the description of the study sample and the analytical approach (which accounts for the clustering and differences at baseline). However, the paper is still very hard to follow, and I tried to figure out why this is, actually. I really had to think deeply why the authors state that there was no primary preventive effect, given that there was an increase in the smoking rate in the control group, but not in the intervention group. I needed to work my way until table 3 to understand the point: The two groups had the same increase, just that the intervention group had a bigger decrease. In the present form it reads as if the primary and secondary analysis are kind of independent from each other. There might be a way to describe the results more clearly:</p> <ul style="list-style-type: none"> - change the primary outcome from absolute prevalence differences to smoking initiation rate (number of students at t2 that smoke in the two groups, and who did not smoke at t1). This will produce a non-significant OR and avoids the confusion produced by table 2. - use Table 3 as the actual result table and present baseline frequencies and follow-up frequencies as well as change over time (the authors decided to leave all descriptives on the complete baseline sample, which I still think is not the best solution. But even if they leave it like this, they should report the smoking frequencies for the longitudinal students in the new table).
-------------------------	---

	<p>- combine tables 2 and 4 to one table</p> <p>Other issues:</p> <ul style="list-style-type: none"> - the paper needs a section which describes the smoking measures and the covariates ("culture") that were used in the panel logistic regression - I still think that numbers for the row "smoking behavior of smokers" (table 1) are hard to understand. Not possible to sum this up to 100%. - the term "time#group#endline#intervention group" should be changed/explained
--	--

REVIEWER	Laetitia Minary Centre hospitalier universitaire de Nancy, France
REVIEW RETURNED	25-Jun-2015

GENERAL COMMENTS	<p>The manuscript has been substantially modified, what reinforced the interest of this article.</p> <p>Some last minor comments:</p> <ul style="list-style-type: none"> - Page 4 line 17: "Less popular [...] available" "This paragraph should be moved after the next paragraph - Page 6 line 24: Maybe indicate the loss of follow-up in each group in brackets - Page 8 line 48: Are students who smoke less than once a month considered as non- smokers? Maybe the definition of non smokers should be adjusted - Page 14 line 20: "In this study, we report a significant effect to reduce smoking prevalence of a - widespread intervention delivered by volunteer medical students [...] of a medical student-delivered school-based tobacco intervention". This paragraph should be moved into line 4 page 14. After "[...] associated responsibilities within communities [21, 22] ». - Page 14 line 3 « About 3 years after medical student Titus J. Brinker founded EAT (January 2012), the programme [...] and is therefore less expensive than comparable programmes.": Maybe this paragraph could be presented later in the text (in a "dissemination of the intervention"section?) - Additional comment: Was a process evaluation also performed ?
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

The manuscript has been substantially modified, what reinforced the interest of this article.

Reply: Thank you for your kind guidance which we highly appreciate and value.

Some last minor comments:

Page 4 line 17: "Less popular [...] available" "This paragraph should be moved after the next paragraph

Reply: We followed your suggestion.

- Page 6 line 24: Maybe indicate the loss of follow-up in each group in brackets

Reply: We followed your suggestion.

Page 8 line 48: Are students who smoke less than once a month considered as non- smokers? Maybe the definition of non smokers should be adjusted

Reply: We added the following sentence as suggested „Non-smokers are defined as pupils who claimed to smoke less than “once a month” within the survey.”

- Page 14 line 20: “In this study, we report a significant effect to reduce smoking prevalence of a - widespread intervention delivered by volunteer medical students [...] of a medical student-delivered school-based tobacco intervention”. This paragraph should be moved into line 4 page 14. After “[...] associated responsibilities within communities [21, 22] ».

Reply: We followed your suggestion.

- Page 14 line 3 « About 3 years after medical student Titus J. Brinker founded EAT (January 2012), the programme [...] and is therefore less expensive than comparable programmes.”: Maybe this paragraph could be presented later in the text (in a “dissemination of the intervention”section?)

Reply: We think this is a great idea and followed your suggestion.

Additional comment: Was a process evaluation also performed ?

Reply: No, it was not.

Reviewer: 1

Reviewer Name Matthis Morgenstern
Institution and Country Geisel School of Medicine at Dartmouth, USA

Please leave your comments for the authors below

The authors really worked a lot on the revision, and I see clear improvement, especially regarding the description of the study sample and the analytical approach (which accounts for the clustering and differences at baseline). However, the paper is still very hard to follow, and I tried to figure out why this is, actually. I really had to think deeply why the authors state that there was no primary preventive effect, given that there was an increase in the smoking rate in the control group, but not in the intervention group. I needed to work my way until table 3 to understand the point: The two groups had the same increase, just that the intervention group had a bigger decrease. In the present form it reads as if the primary and secondary analysis are kind of independent from each other.

Reply: Thank you for helping us to improve our manuscript which is highly appreciated.

There might be a way to describe the results more clearly:

- change the primary outcome from absolute prevalence differences to smoking initiation rate (number of students at t2 that smoke in the two groups, and who did not smoke at t1). This will produce a non-significant OR and avoids the confusion produced by table 2.

Reply: We appreciate your suggestion. However, the primary outcome was defined as the change in prevalences in our study protocol and simply changing the primary outcome is thus not appropriate. In addition, we are convinced that the defined primary outcome should be the change in prevalences as it clearly is as important to eliminate experimentation behaviour as it is to prevent initiation in our young age group.

- use Table 3 as the actual result table and present baseline frequencies and follow-up frequencies as well as change over time (the authors decided to leave all descriptives on the complete baseline sample, which I still think is not the best solution. But even if they leave it like this, they should report the smoking frequencies for the longitudinal students in the new table).

Reply: As stated above, we are not willing to change our primary outcome as we do not think that this would improve our paper. It would arouse a lot of questions instead. The baseline numbers are presented in the abstract, in table 1 and various times in the text. We can not add them to table 3 without losing its format / understandability.

- combine tables 2 and 4 to one table

Reply: Thank you for this suggestion, but we do not think that this would increase the understandability of our manuscript as those two tables describe two different procedures.

Other issues:

- the paper needs a section which describes the smoking measures and the covariates ("culture") that were used in the panel logistic regression

Reply: We asked for „migrant background“ with a yes/no choice for the participants of the survey. The data/sample is described in our baseline table quite bluntly; there were no significant differences between the groups.

- I still think that numbers for the row "smoking behavior of smokers" (table 1) are hard to understand. Not possible to sum this up to 100%.

Reply: The row "cigarettes" (monthly/daily) adds up to a 100%. This is also the case for the row "waterpipe-smokers" (monthly/daily). The columns "daily, more than once per week, once per week" add up to a 100% as well.

- the term "time#group#endline#intervention group" should be changed/explained

Reply: We added a *sentence under the table to explain the terms. They are most precise for the procedure we performed as described in the methods section.