

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)	Cigarette smoking susceptibility among youth alternate tobacco product users: implications of flavored tobacco from a national cross-sectional Canadian sample (YSS 2012/2013).
AUTHORS	Minaker, Leia; Shuh, Alana; Nguyen, Nghia; Azagba, Sunday; Manske, Steve

VERSION 1 - REVIEW

REVIEWER	Hillel Alpert VITAL Science and Health, United States
REVIEW RETURNED	26-Aug-2015

GENERAL COMMENTS	<p>The full weight of the results depend on the adequacy of the measure of susceptibility to smoking. More description of the reliability, validity, and possible limitations of the measure used would be helpful. Sensitivity analysis could also be helpful if possible.</p> <p>E-cigarettes as a form of ATP are notably absent from this study. Considering high prevalence of e-cigarette use among youth, even during the study period, the importance of this limitation and potential relevance to the findings regarding other ATPs should be described. Also, the implications that the findings have for e-cigarettes as ATPs, especially since so many are flavored, should be discussed.</p> <p>More references could be cited regarding trends in ATP use among youth in North America.</p> <p>What are the effects of selection and missing data in this study? How do the non-selected (not meeting present study criteria) population and those with missing data for the key predictor and outcome variables differ from the sample whose data are analyzed?</p> <p>As many other influences on smoking susceptibility as possible should be taken into account in the analysis. Several important factors are included. The possible effects of other social influences such as teachers or other role models and exposure to advertising, for example, should be considered.</p> <p>The variables representing potential social influences on smoking were found to have substantial effects on the outcomes, and relative to the ATP use. The importance should be discussed.</p> <p>Table 2 is missing the results of bivariate analyses with categories of ATP use.</p>
-------------------------	---

	<p>Why not conduct the analysis on the three-category predictor variable, no ATP, non-flavored ATP, flavored ATP and capture all relevant comparisons?</p> <p>Analysis of flavored vs non-flavored ATP use and susceptibility to smoking is mentioned in the Discussion, but the results are not shown.</p> <p>Pearson Chi-Square statistic is not a test. What was the statistical test conducted (e.g. contingency table analysis, other)?</p> <p>Please provide more detail of how the regression analyses were conducted. In what order was the stepwise procedure? Which statistical inclusion and exclusion criteria were used? "PROC LOGISTIC in SAS" is not a method.</p> <p>Please explain why Model 2 was the same as Model 1 except for inclusion of the living area variable?</p> <p>In Discussion, do the authors have a theory for why youth smoking susceptibility seems to be relatively high in Canada?</p> <p>The limitations described of the study are not inherent to survey research, as implied on Page 13, but to the survey data used and the present study. Survey research can be longitudinal, include biochemical or other objective measures, and can include nicotine dependence data if gathered.</p> <p>Page 2: "The results of this study are generalizable...." to which populations?</p>
--	--

REVIEWER	Nadra E. Lisha, Ph.D. UCSF USA
REVIEW RETURNED	04-Oct-2015

GENERAL COMMENTS	<p>This manuscript is well written and has interesting policy implications regarding the use of flavorings in tobacco products. The authors used a large national sample from Canada using the Youth Smoking Survey data from 2012/2013. Students in the sample were from grades 9-12. Smoking susceptibility in non-smokers was examined in a series of logistic regression models. Those who tried alternative tobacco products (ATP) either flavored or not were more susceptible to smoking compared to those who never tried any types of tobacco among never smokers.</p> <p>Some relatively minor revisions are necessary:</p> <ol style="list-style-type: none"> 1. The introduction provides data from between 2004 and 2009 on the increase in ATP; more recent data is available and should be provided. 2. The second paragraph of the introduction, the last two sentences (cigarillo and hookah) should be switched to be consistent with the list in the beginning of the paragraph. 3. The sentence in the 3rd paragraph of the introduction "Some studies, however found evidence to support..." feels out of place. Please explain this in the context of the other findings and provide some ideas as to why these results might differ between studies. 4. The last paragraph of the introduction poses two research questions. Based on the lit review it appears that these should really
-------------------------	---

	<p>be directional hypotheses. If that is indeed the case, please write them as such.</p> <p>5. In the measures section, (p. 6, ln 47-57) it is not necessary to repeat the wording of all the tobacco products using the new question stem as it was previously listed in the prior paragraph. Please shorten this and refer to the first list.</p> <p>6. Table 1 is confusing as the percentages are presented across category rather than within. Please fix this. For example, instead provide the percent of females that are not susceptible (6485/12295 = 52%).</p> <p>7. The analyses presented in Table 2 are very interesting but sort of feel out of place as they are not the central focus of the manuscript. ATP use should be integrated into this analysis – most likely as an additional predictor in the model.</p> <p>8. The analyses presented in Table 2 need to be cleaned up – the reference group should always be either first or last, and make sure that a “1.00” appears in all the places it should (missing from gender and grade).</p> <p>9. Stepwise regression is an unnecessary step here unless the authors are interested in the R-squares associated with the two blocks of variables. If so, please add this and explain. If not, please just present the final analysis.</p> <p>10. It is unclear what variables were used as covariates for the analyses presented in Table 3.</p> <p>11. Discussion – please delete the sentence “Each of these key findings is described below.”</p> <p>12. In the discussion (p. 12, ln. 13-30) the findings from Table 2 are discussed. This paragraph does not fit well and needs to be edited. For example, why are the findings for female gender not presented? This seems like an interested finding that goes counter to typical patterns. Second, the finding that younger grades are actually MORE susceptible to smoking is another important and counter to the literature finding. Please elaborate on reasons this might exist. One possibility is that the people who have never smoked by the 12th grade are already the people who are extremely resistant to smoking and thus restricting the sample in this way creates a finding that goes counter to what we might typically think. Same paragraph, the words “also in line” do not fit as the previous finding was not “in line” with the literature.</p> <p>13. The following paragraph makes a reference to policy in Canada, the EU and the US. Please describe this more.</p> <p>14. The followup analysis on the bottom of page 12 is interesting, but should also be in the results section (flavored ATP vs. non).</p> <p>15. The conclusio</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1	Response to reviewers
The full weight of the results depend on the adequacy of the measure of susceptibility to smoking. More description of the reliability, validity, and possible limitations of the measure used would be helpful. Sensitivity analysis could also be helpful if possible.	We have added the following statement to the methods section, “Methods of characterizing “susceptible non-smokers” were developed by Pierce et al ³⁴ and are consistent with those used in the Global Youth Tobacco Survey. ³⁶
E-cigarettes as a form of ATP are notably absent from this study. Considering high prevalence of e-cigarette use among youth, even during the study	The following statements were added to the Discussion section: “Finally, this survey did not contain information on electronic cigarette use,

<p>period, the importance of this limitation and potential relevance to the findings regarding other ATPs should be described. Also, the implications that the findings have for e-cigarettes as ATPs, especially since so many are flavored, should be discussed.</p>	<p>which is a limitation since electronic cigarettes may be considered an important and increasing form of ATP use.^{46,47} In addition, the liquid vaporized by electronic cigarettes is often flavored,⁴⁸ which has implications for regulation of flavored ATPs not captured by the current study.”</p>
<p>More references could be cited regarding trends in ATP use among youth in North America.</p>	<p>We have added several new references to the introduction and have updated our literature review.</p>
<p>What are the effects of selection and missing data in this study? How do the non-selected (not meeting present study criteria) population and those with missing data for the key predictor and outcome variables differ from the sample whose data are analyzed?</p>	<p>Comparisons to check selection bias are typically feasible with small and experimental studies and have not, to our knowledge, been performed in a large survey. Therefore, we cannot answer how youth who were missing data differ from the sample who had complete data. In terms of how non-selected participants may have differed from participants, we note that non-select participants represent only about 5% of the Canadian population. For analyses in this paper, there were no missing values for smoking status, and less than 1% missing values for the outcome of susceptibility to smoking. For other variables, the highest percent of missing values was for the variable “number of closest friend smoking” (9%). All other variables had less than 5% of missing values as noted in table 1.</p>
<p>As many other influences on smoking susceptibility as possible should be taken into account in the analysis. Several important factors are included. The possible effects of other social influences such as teachers or other role models and exposure to advertising, for example, should be considered.</p>	<p>We agree with the reviewer and therefore we have included number of closest friends smoking and family member smoking as covariates in the analyses. Unfortunately the YSS survey does not collect any information on teachers or other role models nor on exposure to advertising. We have added a statement to this effect in the limitations section.</p>
<p>The variables representing potential social influences on smoking were found to have substantial effects on the outcomes, and relative to the ATP use. The importance should be discussed.</p>	<p>We have included a line in the discussion section, “Our findings that family member and close friend smoking are associated with smoking susceptibility are in line with previous research finding that parental and peer smoking is a well-established factor associated with smoking susceptibility.^{39,46} “</p>
<p>Table 2 is missing the results of bivariate analyses with categories of ATP use.</p>	<p>We have intentionally excluded ATP use from Table 2 because details about ATP use (different exposures) are presented in table 3.</p>

<p>Why not conduct the analysis on the three-category predictor variable, no ATP, non-flavored ATP, flavored ATP and capture all relevant comparisons?</p>	<p>In table 3, we present five nested case-control comparisons that included non-flavored ATP, flavored ATP, other types of tobacco, and last 30 day use. In addition, each comparison has relevant populations for case and control.</p>
<p>Analysis of flavored vs non-flavored ATP use and susceptibility to smoking is mentioned in the Discussion, but the results are not shown.</p>	<p>We have added these results to table 3, and we have also added a description of the analyses and outcomes in the methods and results sections, respectively.</p>
<p>Pearson Chi-Square statistic is not a test. What was the statistical test conducted (e.g. contingency table analysis, other)?</p>	<p>This was a typo that we have corrected. We did not use the Pearson Chi-Square statistic. For this survey analysis, Rao-Scott Chi-Square statistics were applied, as noted in the bottom of table 1.</p>
<p>Please provide more detail of how the regression analyses were conducted. In what order was the stepwise procedure? Which statistical inclusion and exclusion criteria were used? "PROC LOGISTIC in SAS" is not a method.</p>	<p>We agree that "PROC SURVEYLOGISTIC in SAS" is not a method. This is a procedure to generate logistic regression. We have added additional information about our backward elimination strategy. In table 2, two models intentionally include all specified covariates. In table 3, the five models are intentionally "backward elimination".</p>
<p>Please explain why Model 2 was the same as Model 1 except for inclusion of the living area variable? In Discussion, do the authors have a theory for why youth smoking susceptibility seems to be relatively high in Canada?</p>	<p>Model 1 and 2 both included the living area variable. We have corrected this mistake in the methods section.</p> <p>We have added a statement with our thoughts on why smoking susceptibility may be high in Canada: "It may be that the high prevalence of cigarette smoking susceptibility is related to the overall low smoking prevalence in Canada. To be considered "susceptible" to smoking, respondents must report being non-smokers. With a low smoking rate, therefore, the pool of potentially smoking susceptible respondents is larger."</p>
<p>The limitations described of the study are not inherent to survey research, as implied on Page 13, but to the survey data used and the present study. Survey research can be longitudinal, include biochemical or other objective measures, and can include nicotine dependence data if gathered.</p>	<p>We have corrected this mistake. The reviewer is correct that our statement about limitations common to survey research is too broad.</p>
<p>Page 2: "The results of this study are generalizable...." to which populations?</p>	<p>We have changed this sentence to clarify the meaning: "Despite these limitations, the results from this study indicate that ATP use is indeed associated with smoking susceptibility, which is</p>

	of concern given the growing prevalence of ATP use in Canada and globally. ^{2,3,5,6,21»}
Reviewer 2	
The introduction provides data from between 2004 and 2009 on the increase in ATP; more recent data is available and should be provided.	We have updated our literature and have provided more recent references.
The second paragraph of the introduction, the last two sentences (cigarillo and hookah) should be switched to be consistent with the list in the beginning of the paragraph.	We have changed the order of the paragraph to be consistent with the first line.
The last paragraph of the introduction poses two research questions. Based on the lit review it appears that these should really be directional hypotheses. If that is indeed the case, please write them as such.	We have changed the research questions section to: "Specifically we hypothesize that ATP use will be significantly positively associated with smoking susceptibility among youth never smokers. Second, we hypothesize that flavored ATP use will also be strongly and positively associated with smoking susceptibility among youth never smokers."
In the measures section, (p. 6, ln 47-57) it is not necessary to repeat the wording of all the tobacco products using the new question stem as it was previously listed in the prior paragraph. Please shorten this and refer to the first list.	We have made this change as requested.
Table 1 is confusing as the percentages are presented across category rather than within. Please fix this. For example, instead provide the percent of females that are not susceptible (6485/12295 = 52%).	We prefer to leave the table as it is, since we believe it is clearer to show that 68.5% of females are non-susceptible and 31.5% are susceptible rather than what proportion of non-susceptible were male vs. female. In other words, each row adds to 100%, which we think is actually very intuitive. We will make the change the reviewer has suggested if the editor also agrees that we should change the percentages to be presented by column rather than by row.
The analyses presented in Table 2 are very interesting but sort of feel out of place as they are not the central focus of the manuscript. ATP use should be integrated into this analysis – most likely as an additional predictor in the model.	We have included the results in Table 2 since they correspond to our objective outlined in the last paragraph of the introduction: "The purpose of this study was to examine smoking susceptibility and ATP use in a national sample of Canadian grades 9-12 students." Different ATP uses are presented by five different models with different nested case-controls analyses. We have intentionally

	<p>excluded ATP use from table 1 and 2.</p> <p>In summary: Table 1 is merely a descriptive analysis covariates. Table 2 provides findings from logistic regressions to align with our study objective. Table 3 examines ATP use with all relevant covariates.</p>
<p>The analyses presented in Table 2 need to be cleaned up – the reference group should always be either first or last, and make sure that a “1.00” appears in all the places it should (missing from gender and grade).</p>	<p>We have fixed these mistakes and have changed the order so that the last category of each independent variable and covariate is the reference category.</p>
<p>Stepwise regression is an unnecessary step here unless the authors are interested in the R-squares associated with the two blocks of variables. If so, please add this and explain. If not, please just present the final analysis.</p>	<p>Table 2 did not use any strategy and instead includes all co-variates.</p> <p>Table 3 used backward elimination strategy for logistic regression. We have added a method description of confounding selection to the method section.</p>
<p>It is unclear what variables were used as covariates for the analyses presented in Table 3.</p> <p>Discussion – please delete the sentence “Each of these key findings is described below.”</p>	<p>We have added a description of all covariates to the bottom of Table 3 (see explanations after superscripted letters).</p> <p>We have deleted this sentence.</p>
<p>In the discussion (p. 12, ln. 13-30) the findings from Table 2 are discussed. This paragraph does not fit well and needs to be edited. For example, why are the findings for female gender not presented? This seems like an interested finding that goes counter to typical patterns. Second, the finding that younger grades are actually MORE susceptible to smoking is another important and counter to the literature finding. Please elaborate on reasons this might exist. One possibility is that the people who have never smoked by the 12th grade are already the people who are extremely resistant to smoking and thus restricting the sample in this way creates a finding that goes counter to what we might typically think. Same paragraph, the words “also in line” do not fit as the previous finding was not “in line” with the literature.</p>	<p>We have expanded the discussion around female gender and younger age as predictors of smoking susceptibility.</p>
<p>The following paragraph makes a reference to policy in Canada, the EU and the US. Please describe this more.</p>	<p>We have briefly described all of the recent relevant policy changes.</p>
<p>The followup analysis on the bottom of page 12 is interesting, but should also be in the results section</p>	<p>We have taken this section out and have included it instead in the methods and results</p>

(flavored ATP vs. non).	section.
-------------------------	----------

VERSION 2 – REVIEW

REVIEWER	Nadra E. Lisha, Ph.D. UCSF, USA
REVIEW RETURNED	25-Nov-2015

GENERAL COMMENTS	For Table 1 please make the changes that were requested previously. The numbers will still add up to 100%, but it will be by column which makes more sense. By doing the percentages the way you have done you are implying a different comparison than what is being made. For Table 2, there is still no word “ref” for ethnicity.
-------------------------	---

VERSION 2 – AUTHOR RESPONSE

We have made all revisions.