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)	Alcohol-related harm to others in England: a cross-sectional analysis of National survey data
AUTHORS	Beynon, Caryl; Bayliss, David; Mason, Jenny; Sweeney, Kate; Perkins, Clare; Henn, Clive

VERSION 1 – REVIEW

REVIEWER	Lauren M. Kaplan University of California, San Francisco USA
REVIEW RETURNED	18-Dec-2017

GENERAL COMMENTS	This manuscript addresses a timely and important public health concern using a nationally representative sample. Key elements that add to the literature are 1) the significant associations between renter status, disability and socioeconomic factors associated with alcohol's harm to others; 2) the focus on preparatory of harm, and 3) the frequency of harm. Below are comment, which if addressed, would add clarity and strengthen this manuscript: Abstract: 1) Specify the items used for aggressive harm. 2) Define family stage of life briefly -- and do so in the main text. 3) Define "serious" harm. 4) The authors reference the WHO THAI project and do cite Dr. Thomas K. Greenfield's work; however they do not fully describe why the harms measures used in 2014 in their survey did not more closely fit the WHO-THAI project or the relevant harms work published before 2015. Introduction: 1) Removing extraneous text and writing more concisely would allow the authors to delve more deeply into more specific statements regarding prior research. 2) Page 3; Lines 22-26; Specify more what these inconsistent findings are, a more throughout description of prior findings would better frame the current study, this can be done concisely and would add precision and relevance to the introduction. 3) Page 3; Lines 28-31 - The authors do acknowledge gender differences in alcohol's harm to others, yet these differences are kept at a general level -- adding more specificity in the differences (e.g. women being more likely to experience harms from a partner of family member and men more likely to experiencing harm in the US as well as they types of harms more prevalent among men and women). I see that gender difference in harm are discussed briefly

	later in the discussion but in a national study it seems important to provide more background on prevalence of different harms and perpetrators among men and women. The authors do not cite a relevant study examining the role of context in harms among men and women in the US. Please see the below reference for more on drinking context, type of harm, and gender: Drinking context and alcohol's harm from others among men and women in the 2010 US National Alcohol Survey. Kaplan LM, Karriker-Jaffe KJ, Greenfield TK. J Subst Use. 2017;22(4):412-418. doi: 10.1080/14659891.2016.1232758. Epub 2016 Nov 16. PMID:28757805 Methods: 1) I suggest that a methodologist further examine the justifications for the lack of formal sample size calculation. 2) 2) The authors grouped some harms together and defined them as “aggressive”, yet I do not see a discussion of any analysis done to examine whether these groupings were consistent with their data and with prior research. For example, why were potential harms to children not defined as aggressive harm? Potential harm to children is an important topic – why do the authors not report more on this? 3) Page 4; Line 18 – There appears to be a typo in the definition of defining financial harm. Please clarify this item. 4) Page 4; Family stages (lines 46-47) are not clearly defined. No citations are listed for family stages. Please add definitions and citations to clarify this measure. Analysis: 1) There are gender differences in harm -- why was this analysis not stratified by gender? I suggest running additional models stratifying by gender to add more context to the results and relate the results to prior research. 2) It seems that multiple groupings could be used (e.g. property harm, residential instability or neighborhood disorder related to alcohol’s harm. Emotional abuse and neglect etc). I recommend conducting more analysis (perhaps cluster analysis, could also calculate Cronbach’s alphas to see how measures may or may not be reliable for different groupings). 3) Page 5; Lines 6-9 – Please report on any significant differences among those excluded from the analysis. Results: 1) Consider the use of subheadings if permitted by formatting to organize the text. 2) The significant findings for being a hazardous/harmful drinker should vary by gender. I recommend additional analysis by gender as men’s drinking may be significant but not women’s. Discussion: 1) Page 12; lines 26-29 – Clarify whether the Wales study used lifetimes and not past 12 month measures – that may explain the different prevalence. 2) Page 12; lines 52-56 – Add citations to drinking context and harm. Add any relevant findings for gender here. Cite literature finding that men’s, but not women’s drinking is associated with harm and need for more work on women’s drinking and harm. 3) Page 14; lines 43-47; Research on drinking patterns and context in relation to harm is extant, please add and cite. 4) The renter status and disability findings are key findings and warrant a more thoughtful discussion. I recommend a more detailed interpretation of these results. I recommend relating the renter finding more to potential policies (for example a parallel might be

	smoke free housing polices used to reduce harm from smoking --- how might these inform alcohol and housing policy?) 5) Higher education and being White British increased risk for harm – there is research on affluence and wealth and how higher income drinkers may experience more harms. I recommend discussion these findings in a more contextualized way and citing relevant research.

REVIEWER	Betsy Thom Middlesex University, London, UK
REVIEW RETURNED	04-Jan-2018

GENERAL COMMENTS	This is a well conducted and well reported study. The authors acknowledge that it is exploratory and I was pleased to see that they commented on the issue of assumed causality in the limitations section. It was also good that a distinction was made between 'aggressive harm' and other harms as this is sometimes ignored in studies of alcohol- related harms. I would have liked more discussion of the harm items especially some ideas about how these measures might be categorised or grouped in future in order to reflect not only seriousness of the harm but possibly also the issue of perceptions of causality (e.g. noise at night compared to put at risk in a car when someone had been drinking). Also, how can policy respond to these issues? Perhaps matter for another paper. Statistics is not my strong point but as far as I can judge the analysis is sound. This is a useful contribution to the literature.

REVIEWER	Philip Clare National Drug & Alcohol Research Centre, Australia
REVIEW RETURNED	06-Feb-2018

GENERAL COMMENTS	The article is well written and presented, and considers an important issue in alcohol-related harm. However, there are some issues with the statistical methods used that undercut the strength of the findings: 1. **Type-I error.** The authors appear to have conducted a relatively large number of bivariate and multivariate statistical tests, without adjusting to control the type-I error rate. I would suggest that some p-value correction be used, in order to make the findings more robust. 2. **Missing data.** At the least, some consideration of the missing data should be presented. Given the (it appears) relatively small proportion of missing data, it may not be a major issue, but even small amounts of missing data can introduce bias, so it is worth further consideration. There appears to be 5%-10% missing data, overall (including non-response to the AHTO questions). The amount of missing data should be reported (univariately and multivariately) and examined, without requiring the reader to do the work. For example, tests could be conducted demonstrating that the missingness is MCAR (e.g. Little's test), in which case complete-case analysis is validwithout further consideration. This could be largely presented in supplementary material, with a brief reference in the manuscript. If the missing data

	is not MCAR however, then analysis should be reconducted using multiple imputation or full-information maximum likelihood. In addition, I have a number of queries/suggestions regarding the analyses and presentation of findings: · Some more information regarding the sample should be provided – response rate etc. · The manuscript specifies a weighted sample of 4874, but some of the bivariate Ns add to 4875. Please comment/address this · Regarding the bivariate tests, I wonder why logistic regression was not used. This would more readily allow a side-by-side comparison with the multivariate analysis. · On a related note, I wonder if perhaps log-binomial regression to estimate the relative risk ratios (as opposed to logistic regression estimating ORs) might be more appropriate – the proportion of each group experiencing an event is a risk (the odds would be the ratio harm/no harm as opposed to harm/total). · I may have simply missed it, but I could not see an explanation for why social grade was not entered in the multivariate models. If it is not already explained, it should be. · Please replace p-value indicators with the precise p-value (i.e., replace asterisks with ‘p=0.xxx’ or p<0.001 if applicable). · There are a number of instances where causal language is used – including the abstract, where the objectives discuss “predicting” harm. Although covariates in models are commonly referred to as “predictors”, in cross-sectional data models/variables cannot truly be predictive, protective, etc. · It would be interesting to see an analysis of the *number* of different harms reported, not just the binary analysis presented. Something like a Poisson or Negative Binomial model? · I am concerned that the results may be driven by the most commonly reported ‘harms’, which also (in my opinion, at least) appear to be the least problematic. This is not to say they are not harmful, but it seems questionable equating being kept awake with being physically assaulted (for example). I realise this might be complicated – there is likely to be disagreement about what constitutes a serious harm. But perhaps some additional analyses of “groups” of harms could be conducted (one such area could be ‘social’ or ‘emotional’). I would also be very interested to see how the items performed in something like a factor analysis, or even more basic psychometric properties like internal consistency (it would be great to see a structural equation model where instead of ‘any harms’, the latent variable underlying harms

	is regressed on the covariates – but I realise that is somewhat beyond the scope of this article).

REVIEWER	Grace Chan University of Connecticut School of Medicine, USA
REVIEW RETURNED	16-Feb-2018

GENERAL COMMENTS	Based on data collected as part of the UK Alcohol Toolkit Study (ATS) and 18 alcohol harms to others (AHTO) additional items, the authors conducted an exploratory study and reported on prevalence and frequency of these AHTO. This paper highlighted an important public health issue in UK that has also been observed in other countries. In general, it is a well-written manuscript and appropriate statistical analyses had been carried out. However, the strength of this paper may be improved by considering the following comments. 1. While the whole ATS sample (over 20,000 per year since March 2014) is representative of the UK population, it is unclear if the relatively small subsample (less than 5,000) used in this exploratory study does as well even after incorporating the ATS generated sample weights into the analyses. Specifically, this exploratory subsample's characteristics were not reported (neither in the main manuscript nor among the supplemental materials). Also different (weighted/raw) sample sizes were stated in different sections of the manuscript (e.g., n = 4874 in Abstract, n = 4881 in Results). Moreover, why only three months' AHTO data were included in the analysis given they were collected data over a 5-month period (Nov 2015 – Mar 2016)? What methods and/or criteria were used to select this subset? 2. What is the psychometric property of this self-report 18-item AHTO "instrument"? Please comment on the potential discrepancies between reported AHTO and actually experienced AHTO. In particular, how may these discrepancies associate with respondents' characteristics such as sex, race, age, socio-economic status and life stage, and type of AHTO? 3. Since different types of AHTO are likely to have subgroup specific prevalence, is it possible to at least added sex-specific prevalence to Table 1? 4. The binary outcomes in both logistic regression analyses combined very different types of AHTO, it would be insightful to comment on or report some AHTO-specific associations. 5. Apart from how the perpetrator "related" to the respondent (e.g., family, friend, stranger), is there any data on other characteristics of perpetrators such as sex (same or opposite as respondent) and age (younger, about the same, or older comparing to respondent)? These information will be helpful in designing AHTO prevention program in the future.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Lauren M. Kaplan

Institution and Country: University of California, San Francisco, USA Competing Interests: None declared.

This manuscript addresses a timely and important public health concern using a nationally representative sample. Key elements that add to the literature are 1) the significant associations between renter status, disability and socioeconomic factors associated with alcohol's harm to others; 2) the focus on preparatory of harm, and 3) the frequency of harm.

Below are comment, which if addressed, would add clarity and strengthen this manuscript:

Abstract:

1) Specify the items used for aggressive harm.
We have added this to the abstract.

2) Define family stage of life briefly -- and do so in the main text.
We have added this to the abstract and to the methods.

3) Define "serious" harm.
We have added a definition of 'aggressive harm' to the abstract and in the conclusions we define serious harms as those defined as aggressive.

4) The authors reference the WHO THAI project and do cite Dr. Thomas K. Greenfield's work; however they do not fully describe why the harms measures used in 2014 in their survey did not more closely fit the WHO-THAI project or the relevant harms work published before 2015.
We do not refer to this in the abstract. We refer to this in the final paragraph of the paper and here we have indicated why we did not use the WHO THAI survey (i.e. we were not aware of the WHO ThaiHealth project when we began our work).

Introduction:

1) Removing extraneous text and writing more concisely would allow the authors to delve more deeply into more specific statements regarding prior research.
Morespecific statements regarding prior research have been added in accordance with the specific comments detailed in the points below.

2) Page 3; Lines 22-26; Specify more what these inconsistent findings are, a more throughout description of prior findings would better frame the current study, this can be done concisely and would add precision and relevance to the introduction.
We have added information in relation to the inconsistencies (introduction: end of paragraph 3).

3) Page 3; Lines 28-31 - The authors do acknowledge gender differences in alcohol's harm to others, yet these differences are kept at a general level -- adding more specificity in the differences (e.g. women being more likely to experience harms from a partner of family member and men more likely to experiencing harm in the US as well as they types of harms more prevalent among men and women). I see that gender difference in harm are discussed briefly later in the discussion but in a national study it seems important to provide more background on prevalence of different harms and perpetrators among men and women. The authors do not cite a relevant study examining the role of context in harms among men and women in the US. Please see the below reference for more on drinking context, type of harm, and gender:
Drinking context and alcohol's harm from others among men and women in the 2010 US National Alcohol Survey.
Kaplan LM, Karriker-Jaffe KJ, Greenfield TK.
J Subst Use. 2017;22(4):412 PubMed -418. doi: 10.1080/14659891.2016.1232758. Epub 2016 Nov 16.
PMID:28757805
We acknowledge that differences between men and women may be important. We have therefore added additional information about this in the results and the discussion. However we have not added this to the introduction, given that differences by sex are not the focus of the study.

Methods:

1) I suggest that a methodologist further examine the justifications for the lack of formal sample size calculation.

As this is an exploratory study we had no data on the effect size and otherparameters on which to base this. We opted for a relatively large sample size (much larger than other AHTO studies conducted in the UK). We have made no change to the paper in response to this comment.

2) The authors grouped some harms together and defined them as “aggressive”, yet I do not see a discussion of any analysis done to examine whether these groupings were consistent with their data and with prior research. For example, why were potential harms to children not defined as aggressive harm? Potential harm to children is an important topic – why do the authors not report more on this?

Our categorisation of ‘aggressive harm’ was based on previous literature. We have added this justification. We felt that these three harms were conceptually related. We have not considered other important harm-groupings given the length of the paper.

3) Page 4; Line 18 – There appears to be a typo in the definition of defining financial harm. Please clarify this item.

We have added the following line from the questionnaire to the top of the harm questions to make this clearer: ‘Because of someone else’s drinking I have…’

4) Page 4; Family stages (lines 46-47) are not clearly defined. No citations are listed for family stages. Please add definitions and citations to clarify this measure.

This is a good point thanks. We have added this definition to the methods. There is no citation.

Analysis:

1) There are gender differences in harm -- why was this analysis not stratified by gender? I suggest running additional models stratifying by gender to add more context to the results and relate the results to prior research.

The analysis includes sex as an independent variable and we do find differences by sex. We agree that differences by sex are important but were not the focus of this paper. However, we have added additional analyses by sex to the results and discussion.

2) It seems that multiple groupings could be used (e.g. property harm, residential instability or neighborhood disorder related to alcohol’s harm. Emotional abuse and neglect etc). I recommend conducting more analysis (perhaps cluster analysis, could also calculate Cronbach’s alphas to see how measures may or may not be reliable for different groupings).

Thank you for this suggestion. We are unable to include further analysis given that the paper is already over the word limit. Also, no other immediate groups stand out (property harm, for example, is only covered by one harm type anyway). However we have added this suggestion to the final paragraph of the discussion.

3) Page 5; Lines 6-9 – Please report on any significant differences among those excluded from the analysis.

We have included a section on missing data to the results and compared people who were included/excluded from the analyses. We have also added a comment on this in the limitations.

Results:

1) Consider the use of subheadings if permitted by formatting to organize the text.

We have made this change.

2) The significant findings for being a hazardous/harmful drinker should vary by gender. I recommend additional analysis by gender as men’s drinking may be significant but not women’s.

We agree that this would be interesting but are unable to add additional analyses as we are already over the word limit and identifying differences by sex was not a study objective. However, the multivariate analysis will control for the joint effects of sex and hazardous/harmful drinking and so the effect of hazardous/harmful is over and above any effect of sex.

Discussion:

1) Page 12; lines 26-29 – Clarify whether the Wales study used lifetimes and not past 12 month measures – that may explain the different prevalence.
The study in Wales also collected data on harms in the last 12 months. We have added this information.

2) Page 12; lines 52-56 – Add citations to drinking context and harm. Add any relevant findings for gender here. Cite literature finding that men's, but not women's drinking is associated with harm and need for more work on women's drinking and harm.
Differences in harm by sex were not the focus of this paper so we are unable to add much on this. However, we agree that this is an important issue and have added some additional text. We have also highlighted that differences by sex should be explored in future research on AHTO in the UK.

3) Page 14; lines 43-47; Research on drinking patterns and context in relation to harm is extant, please add and cite.
Our discussion includes a discussion on drinking patterns (e.g. harmful, hazardous, binge) and we have added text about context in relation to male/female differences. In the text you highlight we have clarified that we are referring to the UK as data here is lacking.

4) The renter status and disability findings are key findings and warrant a more thoughtful discussion. I recommend a more detailed interpretation of these results. I recommend relating the renter finding more to potential policies (for example a parallel might be smoke free housing polices used to reduce harm from smoking --- how might these inform alcohol and housing policy?)
We have added additional text in relation to disability and housing to the discussion.

5) Higher education and being White British increased risk for harm – there is research on affluence and wealth and how higher income drinkers may experience more harms. I recommend discussion these findings in a more contextualized wayand citing relevant research.
We did not measure wealth per se. We include a discussion on educational attainment, social grade and employment and indicate that these are proxy measures for SES. We feel it is impossible to summaries all the literature on wealth/SES given the multitude of measures used and we have added a comment to this effect to the discussion.

Reviewer: 2
Reviewer Name: Betsy Thom
Institution and Country: Middlesex University, London, UK Competing Interests: None declared

This is a well conducted and well reported study. The authors acknowledge that it is exploratory and I was pleased to see that they commented on the issue of assumed causality in the limitations section. It was also good that a distinction was made between 'aggressive harm' and other harms as this is sometimes ignored in studies of alcohol- related harms. I would have liked more discussion of the harm items especially some ideas about how these measures might be categorised or grouped in future in order to reflect not only seriousness of the harm but possibly also the issue of perceptions of causality (e.g. noise at night compared to put at risk in a car when someone had been drinking). Also, how can policy respond to these issues? Perhaps matter for another paper. Statistics is not my strong point but as far as I can judge the analysis is sound. This is a useful contribution to the literature.
Thank you. In the final paragraph of the discussion we have suggested that other harm groupings are considered in future research. We did not understand your comment about groupings by 'perceptions of causality' sorry. We have also added a couple of sentences about the policy response.

Reviewer: 3
Reviewer Name: Philip Clare
Institution and Country: National Drug & Alcohol Research Centre, Australia Competing Interests: None declared

Please see comments in file, attached.

The article is well written and presented, and considers an important issue in alcohol-related harm. However, there are some issues with the statistical methods used that undercut the strength of the findings:

1. **Type-I error.** The authors appear to have conducted a relatively large number of bivariate and multivariate statistical tests, without adjusting to control the type-I error rate. I would suggest that some p-value correction be used, in order to make the findings more robust.

We do not think it is appropriate to make corrections in the context of this paper, but we do accept it would be helpful to say why this is the case to aid interpretation and we have added a sentence on this to the analysis section. Our paper is exploratory and this is clearly stated from the beginning. We do not set out a formal hypothesis which is tested and used to validate/refute a theory. Instead we look to describe the data collected and present it in such a way as to stimulate further development of this topic area. Corrections rely on an informative universal null hypothesis (Perneger, TV. BMJ. 1998, Apr 18;316[7139]), but such a hypothesis is of no interest in this case. Different descriptive results (e.g. experience of harm, perpetrators or harm, frequency of harm) point to different substantive questions of interest.

The bivariate tests are used to inform the multivariate analyses (only the latter informs the substantive discussion) and therefore do not inflate the chance of finding substantively significant results. If the bivariate tests used a deflated alpha then we may miss important variables in the multivariate model. In the exploratory context of this paper, avoiding type II error may be considered more important, to avoid limiting further avenues of potentially important research (Rothman, KJ. Epidemiology. 1990 Jan;1[1]). As there is no overall hypothesis, we feel that making a correction in this context would introduce interpretation difficulties for readers.

2. **Missing data.** At the least, some consideration of the missing data should be presented. Given the (it appears) relatively small proportion of missing data, it may not be a major issue, but even small amounts of missing data can introduce bias, so it is worth further consideration. There appears to be 5%-10% missing data, overall (including non-response to the AHTO questions). The amount of missing data should be reported (univariately and multivariately) and examined, without requiring the reader to do the work. For example, tests could be conducted demonstrating that the missingness is MCAR (e.g. Little's test), in which case complete-caseanalysis is valid without further consideration. This could be largely presented in supplementary material, with a brief reference in the manuscript. If the missing data is not MCAR however, then analysis should be reconducted using multiple imputation or fullinformation maximum likelihood.

We have now added a section on missing data into the paper. We have included a table which compares those included/excluded from the analyses because of missing responses to the AHTO questions (as requested by another reviewer). We have included details of the number and proportion of people who were excluded from the multivariate analyses. We did not consider imputation because the proportion of missing data is small. However, we agree with your comments and we have added this in the limitations. We hope this is satisfactory.

In addition, I have a number of queries/suggestions regarding the analyses and presentation of findings:

• Some more information regarding the sample should be provided – response rate etc.

Unfortunately IPSOS Mori did not collect this information. We have added this to the limitations.

• The manuscript specifies a weighted sample of 4874, but some of the bivariate Ns add to 4875. Please comment/address this

This is due to rounding as we are using weighted data. We have explained this in the table footnotes.

• Regarding the bivariate tests, I wonder why logistic regression was not used. This would more readily allow a side-by-side comparison with the multivariate analysis.

The purpose of the bivariate analyses were to describe the response patterns and test variation (i.e. to identify the variables significantly related to the outcome in order to build the multivariate model) rather than test strata level associations. We felt there was no need to directly compare bivariate and multivariate results.

• On a related note, I wonder if perhaps log-binomial regression to estimate the relative risk ratios (as opposed to logistic regression estimating ORs) mightbe more appropriate – the proportion of each group experiencing an event is a risk (the odds would be the ratio harm/no harm as opposed to harm/total).

We have left the statistics as they are given that other reviewers have said they are happy with our approach. We appreciate that there are other approaches to the analysis but our approach is valid and is consistent with previous literature (e.g. Bellis et al., 2015 and Moan et al., 2015).

• I may have simply missed it, but I could not see an explanation for why social grade was not entered in the multivariate models. If it is not already explained, it should be.

Social grade was not entered into the multivariate model because it was not significantly associated with the outcome in the bivariate analyses. Only variables that were significantly associated to the outcome in the bivariate analyses were included in the multivariate model. We have reinforced this point by adding the word 'significantly' into the relevant sentence in the analysis section of the methods.

• Please replace p-value indicators with the precise p-value (i.e., replace asterisks with

'p=0.xxx' or p<0.001 if applicable).

We have added the actual p values.

• There are a number of instances where causal language is used – including the abstract, where the objectives discuss “predicting” harm. Although covariates in models are commonly referred to as “predictors”, in cross-sectional data models/variables cannot truly be predictive, protective, etc.

Thank you for bringing this to our attention; we have changed this throughout the paper.

• It would be interesting to see an analysis of the *number* of different harms reported, not just the binary analysis presented. Something like a Poisson or Negative Binomial model?

Thank you for this suggestion. However we are unable to extend this paper any further given we are already over the word limit. We agree that the number of harms would be of interested but this was not a specific objective of the paper.

• I am concerned that the results may be driven by the most commonly reported ‘harms’, which also (in my opinion, at least) appear to be the least problematic. This is not to say they are not harmful, but it seems questionable equating being kept awake with being physically assaulted (for example). I realise this might be complicated – there is likely to be disagreement about what constitutes a serious harm. But perhaps some additional analyses of “groups” of harms could be conducted (one such area could be ‘social’ or ‘emotional’). I would also be very interested to see how the items performed in something like a factor analysis, or even more basic psychometric properties like internal consistency (it would be great to see a structural equation model where instead of ‘any harms’, the latent variable underlying harms is regressed on the covariates – but I realise that is somewhat beyond the scope of this article).

Because the research on this topic is relatively sparse we took a broad scope with descriptive aims. Now this has been achieved, future research can progress with more specific enquiries, with a larger number of reports of individual harm types.

We agree that grouping harms would be interesting but we had no further space to include additional analyses in this initial exploratory paper. We felt that (after harm/no harm) aggressive harms were the key group to focus on. We have added a justification for why we focused on aggressive harms specifically (i.e. it is in line with previous research) and we have added to the final paragraph that other harm groups should be considered in the future. Furthermore, we do not suggest or discuss the idea that the specific harms we measured are observable manifestations of an underlying but unmeasured construct. We wanted to understand prevalence of these specific harms.

Reviewer: 4
Reviewer Name: Grace Chan
Institution and Country: University of Connecticut School of Medicine, USA Competing Interests: none

Based on data collected as part of the UK Alcohol Toolkit Study (ATS) and 18 alcohol harms to others (AHTO) additional items, the authors conducted an exploratory study and reported on prevalence and frequency of these AHTO. This paper highlighted an important public health issue in UK that has also been observed in other countries. In general, it is a well-written manuscript and appropriate statistical analyses had been carried out. However, the strength of this paper may be improved by considering the following comments.

1. While the whole ATS sample (over 20,000 per year since March 2014) is representative of the UK population, it is unclear if the relatively small subsample (less than 5,000) used in this exploratory

study does as well even after incorporating the ATS generated sample weights into the analyses. Specifically, this exploratory subsample's characteristics were not reported (neither in the main manuscript nor among the supplemental materials). Also different (weighted/raw) sample sizes were stated in different sections of the manuscript (e.g., n = 4874 in Abstract, n = 4881 in Results). Moreover, why only three months' AHTO data were included in the analysis given they were collected data over a 5-month period (Nov 2015 – Mar 2016)? What methods and/or criteria were used to select this subset?

- *Each month the ATS takes a representative sample of the population. We have clarified this under 'the survey' section of the methods.*
- *The characteristics of the sample are presented in Table 2 and in we have added Supplementary Table 1 to provide more detail on this.*
- *The unweighted sample size was 4,881 and the weighted sample was 4,874; both of these values are reported in the results but only the weighted value is reported in the abstract as the analyses were conducted using weighted data.*
- *Only three months of data on AHTO were collected and all of this was included in the analyses. This was an error and has been corrected under 'the survey' section of the method.*

2. What is the psychometric property of this self-report 18-item AHTO"instrument"? Please comment on the potential discrepancies between reported AHTO and actually experienced AHTO. In particular, how may these discrepancies associate with respondents' characteristics such as sex, race, age, socio-economic status and life stage, and type of AHTO?
We have added that the internal validity of the AHTO questions that we used has not been tested and that we failed to identify any validated measure that we could use.

3. Since different types of AHTO are likely to have subgroup specific prevalence, is it possible to at least added sex-specific prevalence to Table 1?
We have added the sex-specific prevalence to the paper.

4. The binary outcomes in both logistic regression analyses combined very different types of AHTO, it would be insightful to comment on or report some AHTO-specific associations.
We would like to investigate types of AHTO in more detail. We chose aggressive/non-aggressive harms. Unfortunately breaking the types of harm into further sub-groups is out of the scope of this paper.

5. Apart from how the perpetrator "related" to the respondent (e.g., family, friend, stranger), is there any data on other characteristics of perpetrators such as sex (same or opposite as respondent) and age (younger, about the same, or older comparing to respondent)? These information will be helpful in designing AHTO prevention program in the future.
This information was not collected.

VERSION 2 – REVIEW

REVIEWER	Lauren M. Kaplan, PhD Division of General Internal Medicine, University of California, San Francisco
REVIEW RETURNED	07-Jun-2018

GENERAL COMMENTS	The authors have provided a clear summary of their revisions. I thank the authors for their substantial changes, which have improved their manuscript. I believe the authors can, and should, address the minor comments below. Abstract: Please revise outcome to define the harms which are not defined as aggressive. Discussion: Please amend this statement in the Discussion to briefly describe which harms are referred to: "Women have also been identified as being at higher risk of harm in

	the USA" (page 14; lines 32-33). The authors have made substantial revisions to include their results for men and women. In Supplemental Table 2, they find a large difference between men and women in "Been emotionally hurt or neglected" -- this could be added to the text on Supplemental Table 2 in the Discussion (page 14; lines 36-38) and to their Results (page 6; lines 18-21). The authors have added an important section on renter status to their Discussion (page 15 lines 16-24). However, the authors do not cite any research on smoking policy and housing. An important point missing from this section is that of Housing First approach. I advise that the authors cite smoking policy implementation research (references provided below) to ensure that their recommendations are not framed in such a way as to create more harm among vulnerable populations (e.g. making abstinence a criteria for housing). Further, there is research focused on victimization and housing that is not cited in this section (provided below). While the harms experienced may not be explicitly alcohol-related, it is not accurate to state that no research on housing and harm exists. This is a well-established research area. Those with more unstable housing (e.g. sheltered, living temporarily in shelters) are more likely to be victimized than those with private housing. Below are references for the authors, which can be cited in regards to: 1) housing and victimization & 2) smoke free policies and their implementation in living environments among vulnerable populations: Kushel MB, Evans JL, Perry S, Robertson MJ, Moss AR. No Door to Lock: Victimization Among Homeless and Marginally Housed Persons. Archives of Internal Medicine. 2003;163(20):2492-2499. Nyamathi AM, Leake B, Gelberg L. Sheltered Versus Nonsheltered Homeless Women: Differences in Health, Behavior, Victimization, and Utilization of Care. Journal of General Internal Medicine. 2000;15(8):565-572. Dietz TL, Wright JD. Age and Gender Differences and Predictors of Victimization of the Older Homeless. http://dx.doi.org/10.1300/J084v17n01_03. 8 Sep 2008 2008. Meinbresse M, Brinkley-Rubinstein L, Grassette A, et al. Exploring the experiences of violence among individuals who are homeless using a consumer-led approach. Violence Vict. 2014;29(1):122-136. Vijayaraghavan M, Pierce JP. Interest in Smoking Cessation Related to a Smoke-Free Policy Among Homeless Adults. J Community Health 2015;40(4):686-91 Vijayaraghavan M, Hurst S, Pierce JP. Implementing Tobacco Control Programs in Homeless Shelters: A Mixed-Methods Study. Health Promot Pract 2016;17(4):501-511. Vijayaraghavan M, Guydish J, Pierce JP. Building Tobacco Cessation Capacity in Homeless Shelters: A Pilot Study. J Community Health 2016;41:998-1005. I commend the authors for adding a relevent section on disability and harm. However, the authors fail to cite research on disability and exposure to harm, such harms may not be alcohol-related but exposure to financial, physical, and sexual victimization has been linked to disability (e.g. functional and cognitive impairment and among older, impaired populations). I advise the authors cite this research and clarify that little is known about alcohol related harm and disability, despite research indicating that those with disabilities are at a higher risk of harm.

	Page 14 (lines 50-51) Please add citations here: "Literature on the effect of socio-economic status is mixed and comparisons are hindered by the multitude of different measures used." The authors have decided to not add a statement on the possibility of wealth and culture when discussing their findings for SES due to word count. I advise the authors to consider a brief, one sentence statement, which is doable with editorial revisions. I advise the same regarding the finding for British exposure to harms not being attirbutable to AUDIT score as both findings suggest potential socio-cultural factors may play a role in harm. This can be said quite briefly.

REVIEWER	Philip Clare UNSW Sydney
REVIEW RETURNED	30-May-2018

GENERAL COMMENTS	The revision present a marked improvement over the original submission. At this stage, my main concern lies with the requests made by reviewers for additional analyses which would add to the value of this paper, but which the authors have chosen not to perform because of the 'word limit'. Given the fact that results can be presented in supplementary materials, and discussed in text in very few words (if the results are consistent with the primary analysis), I do not feel this is a valid excuse. Where the analyses requested are outside the scope of the study, I accept the decision to not include them. In particular, the discussion of gender differences is complex enough to warrant its own investigation. But (for example), failing to examine the complexity of harm severity because this study is exploratory leaves open the possibility that the findings do not generalise beyond common but low-impact 'harms'. Similarly, discussion of SES being curtailed because the study lacks a 'good' measure, is somewhat weak - there is a wealth of studies using education in particular as a proxy for SES, so a nuanced consideration of SES is certainly possible at not, as the authors state, "impossible". Finally, regarding the statistical methodology: I disagree with the strong claim that exploratory studies need not concern themselves with Type I error, but am willing to be forgiving given the otherwise important nature of the work. However, with regards to the sample size: R1 was correct to point out that a rationale for the sample size should be presented. That no formal power analysis was conducted is unfortunate, but common in this kind of research (although I question the authors statement that they could not at all estimate effect sizes - if nothing else, power could have been calculated based on a threshold of minimum relevant difference). But the article should at least state that the sample size is in line with other similar studies. I also question the authors' rebuttal of my suggestion of log-binomial regression for relative risk ratios instead of odds ratios - for one, only 1 of the other 3 authors commented on the validity of the approach, and they confessed to not be a statistician. Relying on the literature only works if the literature is clearly correct - but there are a number of studies discussing the limitations of using logistic regression when what is really being discussed is incidence not odds (which you do by presenting the prevalence as a proportion of the group). See: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3399983/ https://academic.oup.com/aje/article/157/10/940/290159 As a minor point - I'm still not clear how precisely social grade was evaluated as non-significant. In table 2, the 95% CIs suggest C1 and

	E have sig higher harm than AB, C2 and D.

REVIEWER	Grace Chan University of Connecticut School of Medicine, USA
REVIEW RETURNED	22-May-2018

GENERAL COMMENTS	This revision had cleared up many things. There are at least few more minor modifications that may enhance the manuscript further. 1. Weighted or non-weighted (i.e., raw) data? Please add this information to the title of Figures 1 – 2, and Supplementary Tables 3 – 4. 2. Please add n and/or % to Figure 2. 3. Order of harm type in Supplementary Tables 3 and 4: Please consider matching the row orders with those in Table 1 and Supplementary Table 2. 4. Order of perpetrator type in Supplementary Table 4: Please consider starting with the most frequent, i.e., “a friend”, to the least frequent, i.e., “A work colleague” according to Figure 1.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Lauren M. Kaplan, PhD

Institution and Country: Division of General Internal Medicine, University of California, San Francisco, USA Competing Interests: None declared.

The authors have provided a clear summary of their revisions. I thank the authors for their substantial changes, which have improved their manuscript. I believe the authors can, and should, address the minor comments below.

Abstract:

Please revise outcome to define the harms which are not defined as aggressive.

We have added a definition to the abstract.

Discussion:

Please amend this statement in the Discussion to briefly describe which harms are referred to:

"Women have also been identified as being at higher risk of harm in the USA" (page 14; lines 32-33).

We have added this information.

The authors have made substantial revisions to include their results for men and women. In Supplemental Table 2, they find a large difference between men and women in "Been emotionally hurt or neglected" -- this could be added to the text on Supplemental Table 2 in the Discussion (page 14; lines 36-38) and to their Results (page 6; lines 18-21).

This is a good addition, thank you. We have added this information to both the results and discussion.

The authors have added an important section on renter status to their Discussion (page 15 lines 16-24). However, the authors do not cite any research on smoking policy and housing. An important point missing from this section is that of Housing First approach. I advise that the authors cite smoking policy implementation research (references provided below) to ensure that their recommendations are not framed in such a way as to create more harm among vulnerable populations (e.g. making abstinence a criteria for housing). Further, there is research focused on victimization and housing that is not cited in this section (provided below). While the harms experienced may not be explicitly alcohol-related, it is not accurate to state that no research on housing and harm exists. This is a well-established research area. Those with more unstable housing (e.g. sheltered, living temporarily in shelters) are more likely to be victimized than those with private housing. Below are references for the authors, which can be cited in regards to: 1) housing and victimization & 2) smoke free policies and their implementation in living environments among vulnerable populations:

Kushel MB, Evans JL, Perry S, Robertson MJ, Moss AR. No Door to Lock: Victimization Among Homeless and Marginally Housed Persons. Archives of Internal Medicine. 2003;163(20):2492-2499.

Nyamathi AM, Leake B, Gelberg L. Sheltered Versus Nonsheltered Homeless Women: Differences in Health, Behavior, Victimization, and Utilization of Care. Journal of General Internal Medicine. 2000;15(8):565-572.

Dietz TL, Wright JD. Age and Gender Differences and Predictors of Victimization of the Older Homeless. http://dx.doi.org/10.1300/J084v17n01_03. 8 Sep 2008 2008.

Meinbresse M, Brinkley-Rubinstein L, Grassette A, et al. Exploring the experiences of violence among individuals who are homeless using a consumer-led approach. Violence Vict. 2014;29(1):122-136.

Vijayaraghavan M, Pierce JP. Interest in Smoking Cessation Related to a Smoke-Free Policy Among Homeless Adults. J Community Health 2015;40(4):686 PubMed -91

Vijayaraghavan M, Hurst S, Pierce JP. Implementing Tobacco Control Programs in Homeless Shelters: A Mixed-Methods Study. Health Promot Pract 2016;17(4):501 PubMed -511.

Vijayaraghavan M, Guydish J, Pierce JP. Building Tobacco Cessation Capacity in Homeless Shelters: A Pilot Study. J Community Health 2016;41:998-1005 PubMed .

While the victimisation and housing research is helpful and relevant (we have added this, thank you), we do not feel the smoking policy implementation research fits within this paper. Although some parallels can be drawn between harm caused by others' smoking and others' alcohol consumption, the paper is not policy focussed. The inclusion of this work, especially around such a specific issue, comes across as out of place when read in the context of the paper.

I commend the authors for adding a relevant section on disability and harm. However, the authors fail to cite research on disability and exposure to harm, such harms may not be alcohol-related but exposure to financial, physical, and sexual victimization has been linked to disability (e.g. functional and cognitive impairment and among older, impaired populations). I advise the authors cite this research and clarify that little is known about alcohol related harm and disability, despite research indicating that those with disabilities are at a higher risk of harm.

We have stated that while there is no evidence on the association of disability and alcohol-related harm there is good evidence that people with a disability are the victims of harms such as physical, sexual and intimate partner violence and financial hardship.

Page 14 (lines 50-51) Please add citations here: "Literature on the effect of socio-economic status is mixed and comparisons are hindered by the multitude of different measures used." The authors have decided to not add a statement on the possibility of wealth and culture when discussing their findings for SES due to word count. I advise the authors to consider a brief, one sentence statement, which is doable with editorial revisions. I advise the same regarding the finding for British exposure to harms not being attirbutable to AUDIT score as both findings suggest potential socio-cultural factors may play a role in harm. This can be said quite briefly.

We have added that the study did not measure wealth per se so measures like employment and social grade are proxies for this. We added a paragraph about our findings in relation to socio-cultural differences.

Reviewer: 3

Reviewer Name: Philip Clare

Institution and Country: UNSW Sydney, Australia Competing Interests: None to declare

The revision present a marked improvement over the original submission. At this stage, my main concern lies with the requests made by reviewers for additional analyses which would add to the value of this paper, but which the authors have chosen not to perform because of the 'word limit'. Given the fact that results can be presented in supplementary materials, and discussed in text in very few words (if the results are consistent with the primary analysis), I do not feel this is a valid excuse. Where the analyses requested are outside the scope of the study, I accept the decision to not include them. In particular, the discussion of gender differences is complex enough to warrant its own investigation. But (for example), failing to examine the complexity of harm severity because this study is exploratory leaves open the possibility that the findings do not generalise beyond common but low-impact 'harms'. Similarly, discussion of SES being curtailed because the study lacks a 'good' measure, is somewhat weak - there is a wealth of studies using education in particular as a proxy for SES, so a nuanced consideration of SES is certainly possible at not, as the authors state, "impossible".

We believe that we have examined the data according to harm severity by focusing on aggressive harms specifically in addition to all harms. We took the pragmatic decision to focus on aggressive harms as we felt the three aggressive harms were most severe and could be meaningfully grouped. We have added text to the second paragraph of the discussion to acknowledge your concern.

The discussion includes a section on education, employment status and type of accommodation (which we state are proxy measures for SES). We have added an additional study to the section on education.

Finally, regarding the statistical methodology: I disagree with the strong claim that exploratory studies need not concern themselves with Type I error, but am willing to be forgiving given the otherwise important nature of the work.

However, with regards to the sample size: R1 was correct to point out that a rationale for the sample size should be presented. That no formal power analysis was conducted is unfortunate, but common in this kind of research (although I question the authors statement that they could not at all estimate effect sizes - if nothing else, power could have been calculated based on a threshold of minimum relevant difference). But the article should at least state that the sample size is in line with other similar studies.

We have added to the methods that the sample size was larger than other studies on AHTO conducted in the UK.

I also question the authors' rebuttal of my suggestion of log-binomial regression for relative risk ratios instead of odds ratios - for one, only 1 of the other 3 authors commented on the validity of the approach, and they confessed to not be a statistician. Relying on the literature only works if the literature is clearly correct - but there are a number of studies discussing the limitations of using logistic regression when what is really being discussed is incidence not odds (which you do by presenting the prevalence as a proportion of the group). See:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3399983/

https://academic.oup.com/aje/article/157/10/940/290159

We believe that logistic regression is a valid approach to the multivariate analyses (similar papers published in BMJ Open use the same approach). When we discuss the multivariate analysis, we do it from the perspective of comparing subgroups to identify potential associations, not estimating prevalence of harm for these subgroups specifically. We do not feel that presenting both prevalence findings and comparative odds in one paper is problematic; both have a clear purpose.

The key aspect of the papers you sent was the incorrect presentation and interpretation of odds ratios (i.e. using them to estimate of relative risk). We agree that the reporting and interpretation of odds ratios need careful consideration to avoid any confusion. Nowhere in our paper do we use the odds ratios to estimate relative risk. We have revised the paper to ensure that nowhere can the reader be inaccurately assume we are talking about risk and not odds, including revising the generic term 'risk factor'.

As a minor point - I'm still not clear how precisely social grade was evaluated as non-significant. In table 2, the 95% CIs suggest C1 and E have sig higher harm than AB, C2 and D.

The bivariate tests were used to assess whether each independent variable was associated with harm (and therefore inclusion/exclusion from the multivariate model), and not stratum level differences with the reference category. The bivariate test resulted in a p value of 0.142. The confidence intervals overlap for all categories (though only just for C1 and E in the 'any harm' as you identified).

Reviewer: 4

Reviewer Name: Grace Chan

Institution and Country: University of Connecticut School of Medicine, USA Competing Interests: None

This revision had cleared up many things. There are at least few more minor modifications that may enhance the manuscript further.

1. Weighted or non-weighted (i.e., raw) data? Please add this information to the title of Figures 1 – 2, and Supplementary Tables 3 – 4.

We have made these changes.

2. Please add n and/or % to Figure 2.

We have added a footnote to Figure 2 to indicate the n value.

3. Order of harm type in Supplementary Tables 3 and 4: Please consider matching the row orders with those in Table 1 and Supplementary Table 2.

We have made this change.

4. Order of perpetrator type in Supplementary Table 4: Please consider starting with the most frequent, i.e., "a friend", to the least frequent, i.e., "A work colleague" according to Figure 1.

We have made this change.

VERSION 3 – REVIEW

REVIEWER	Lauren Kaplan Division of General Internal Medicine, University of California, San Francisco
REVIEW RETURNED	19-Jul-2018

GENERAL COMMENTS	I commend the authors for their substantial revisions and integrating the reviewers' suggestions. I have one minor remaining point, which I am confident the authors can address. Discussion (page 15; lines 3-8): "Measures such as educational attainment, type of accommodation, social grade and employment status are proxy measures for socio-economic status. Literature on the effect of socio-economic status is mixed and comparisons are hindered by the multitude of different measures used in different studies. In this study social grade was not significantly associated with harm or aggressive harm in the bivariate analyses but other socio-economic variables were." The authors added this in response to prior reviews but the text does not contextualize or interpret their results. The purpose of a Discussion section is to do so. The text the authors have added is vague. If the authors can revise this text to specify the measures of SES that were significant and write a few brief statements as to why they think these variables were significant, then this manuscript would be acceptable for publication. I recommend the authors work with an editor to improve the writing of the manuscript. There are areas throughout the manuscript where the writing could be made more clear and concise.

REVIEWER	Philip Clare National Drug & Alcohol Research Centre, UNSW Sydney, Australia
REVIEW RETURNED	30-Jul-2018

GENERAL COMMENTS	The authors have if not responded, then at least provided adequate justification for the lack of response to previous comments from reviewers. While there remain a small number of typographical errors (for example, P3/L16 - there should be a word after "England") and similar matters, these are quite minor and may be corrected in editing.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Lauren Kaplan
Institution and Country: Division of General Internal Medicine, University of California, San Francisco
Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below
I commend the authors for their substantial revisions and integrating the reviewers' suggestions. I have one minor remaining point, which I am confident the authors can address.

Discussion (page 15; lines 3-8):

"Measures such as educational attainment, type of accommodation, social grade and employment status are proxy measures for socio-economic status. Literature on the effect of socio-economic status is mixed and comparisons are hindered by the multitude of different measures used in different studies. In this study social grade was not significantly associated with harm or aggressive harm in the bivariate analyses but other socio-economic variables were."

The authors added this in response to prior reviews but the text does not contextualize or interpret their results. The purpose of a Discussion section is to do so. The text the authors have added is vague. If the authors can revise this text to specify the measures of SES that were significant and write a few brief statements as to why they think these variables were significant, then this manuscript would be acceptable for publication.

Paragraphs 10 to 13 provide details of the findings and literature on social status, qualifications, employment status and housing tenure respectively. We have added the findings of more recent studies into these paragraphs. We have added further detail to paragraph 14 which provides the interpretation of these findings. No clear picture emerges in relation to the influence of socio-economic status on harm either from our findings or from the literature and we have made this point clearer.

I recommend the authors work with an editor to improve the writing of the manuscript. There are areas throughout the manuscript where the writing could be made more clear and concise.

The manuscript has been reviewed again and we have tried to make the text clearer and more concise. If the editor would like to indicate sections where the clarity could be improved, we would happily make these changes.

Reviewer: 3
Reviewer Name: Philip Clare
Institution and Country: National Drug & Alcohol Research Centre, UNSW Sydney, Australia Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
The authors have if not responded, then at least provided adequate justification for the lack of response to previous comments from reviewers. While there remain a small number of typographical errors (for example, P3/L16 - there should be a word after "England") and similar matters, these are quite minor and may be corrected in editing.

Thank you for highlight this error (P3/L16) which we have corrected. We have asked a person who has not previously read the manuscript to do so. We would also welcome any suggestions from the editor on areas where the text could be improved.