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)	How Australian General Practitioners engage in discussions about alcohol with their patients: a cross-sectional study
AUTHORS	Miller, Emma; Ramsey, Imogen; TRan, Ly; Tsourtos, George; Baratiny, Genevieve; Manocha, Ramesh; Olver, Ian

VERSION 1 - REVIEW

REVIEWER	Michael Tam General Practice Unit, South Western Sydney Local Health District, and Ingham Institute of Applied Medical Research, NSW Australia; School of Public Health and Community Medicine, UNSW Australia, NSW Australia A number of my publications are cited in this paper
REVIEW RETURNED	30-Aug-2016

GENERAL COMMENTS	Review Summary Thank you for this opportunity to review this paper. This study examined potential barriers and facilitators to discussions between patients and GPs using questionnaires. The topic is one of some interest to clinical medicine and public health, especially to Australian general practice. I need to declare that I was serendipitously one of the peer reviewers on a previous version of this paper that was submitted to the Medical Journal of Australia. I am pleased to see that the authors have made use of the criticisms and suggestions given by all the peer reviewers previously and have substantially improved their paper. My overall conclusion is that the data and results of this study contributes to the field and I would be supportive for this study to be published. All critical comments following needs to be taken from this point of view. However, there will need to be some important revisions, especially with regards to how the results are interpreted, and the limitations acknowledged. Major issues Bivariate and multivariate analyses and interpretations The interpretation of these quantitative analyses are entirely dependent on whether the qualitative labels of the factors (which do not appear to have been constructed in a planned or peer reviewed manner) accurately represent the actual phenomenon that was measured from the participant responses. I have some doubts related to the construct validity of the questionnaires, and some major doubts to the qualitative labels used to group responses.

Importantly, I note that the labels of the factors in this version of the manuscript are substantially different to the labels used in the prior version. This further reduces my confidence of the validity of the labels. Moreover, this suggests to me that at least part of the model development was EXPLORATORY and was not designed a priori. For instance, table 4 has three outcomes in the model labelled "Likely to routinely ask about alcohol consumption" and "Confidence in ability to assess and manage alcohol issues", and "Usually not enough time to enquire about each patient's alcohol intake", with 4, 5, and 3 labelled factors respectively. Your previous paper with the equivalent model had three outcomes labelled "Does not routinely ask", "Low confidence" and "Insufficient skills", with 2, 3, and 2 factors respectively.

Between these two drafts, the framing of the model outcomes is quite different, and, there are different numbers of factors in the analyses.

The authors must explain their data analysis process and more clearly note that the analyses were exploratory. They will need to better explain how they came to the qualitative labels for their factors – some of this will be aided by including the actual questionnaire(s) used.

It is my view that the results of this study as they have been presented, though interesting and perhaps indicative of certain phenomena, contain much more uncertainty than acknowledged by the authors. These GPs are not representative of GPs in general. The construction of the questionnaire, the terminology used, and the labelling and focus of subsequent analyses are likely to have been influenced and constrained by pre-existing conceptual biases held by the investigators on the topic.

The authors place substantial confidence in their analytic approach and model development, and this is not a confidence that I share having read through two versions of this paper.

Lastly, even assuming that the analysis and model developed is largely valid within the context of the limited range of participants, the "adjusted agreement ratios" in the final model (table 4), for "Likely to routinely ask about alcohol consumption", range from the most positive at 1.12, to least positive at 0.80. To me, this appears to mean that all the factors described have a relatively small impact on the likelihood of "routinely ask about alcohol". It is possible, if not probable, that this study did not capture the range of explanations and factors that contribute to the majority of the variation in this phenomenon.

Minor issues
Page 4, Line 38: "Little is known about how GPs initiate discussions about alcohol with their patients in Australian general practice settings…"
This statement is referenced to citation 15, (Johnson, Jackson, Guillaume, Meier, & Goyder, 2011), a systematic review of qualitative evidence of barriers and facilitators of implementing alcohol screening and brief interventions, and does not seem entirely appropriate. I wonder whether this is an error?

Page 4, Line 45: "Pennay et al. propose confidence in managing alcohol issues as a potential barrier to screening and brief intervening in alcohol issues."
This is a somewhat narrow point to take from the "treatment: strategies" section of the (Pennay, Lubman, & Frei, 2014) paper in

Australian Family Physician. This was from a single statement in that paper. This statement in (Pennay et al., 2014) in turn referenced the phase 3 of the WHO Collaborative Project from two decades ago – it was published in 1998. I generally agree with the concept that confidence in managing alcohol issues is something that has an effect on GP alcohol screening behaviours, but a better reference should be found.

A theoretical framework that is used in the field are the concepts of "role security" and "therapeutic commitment". Confidence would be subsumed within these concepts.

Although there has been enthusiasm from various researchers that influencing one or both of these factors may improve alcohol screening and brief intervention rates (which is an intuitive hypothesis), there is some empirical evidence that it does not. For instance, this recent paper by (Bendtsen et al., 2015) found that professional attitudes did not influence SBI rates. One observation to reflect on is the conceptual bias that many in the field have in seeking clinician-related factors/deficiencies as the explanation for the phenomena (e.g., not enough knowledge, not enough resources, wrong attitude) and plan for intervention (e.g., education, support, and mentorship respectively). It might be worth considering that this study similarly focussed almost exclusively on clinician-related factors and thus, any description of the phenomena, as well as recommendations for intervention are likely to be incomplete.

Existing literature on barriers and facilitators
Although I respect the authors' prerogative into placing weight on specific ideas and concepts in prior literature in terms of barriers and facilitators, I feel that they need to explain why so much weight was given to the review article by (Pennay et al., 2014) in the paragraph on page 4, lines 23 to 52.

With all due respect to (Pennay et al., 2014), I do not believe that they discussed the causes of poor uptake of alcohol screening and brief interventions by GPs in much depth in their paper. In defence of their paper, this was not their intention in the section where they describe it. I would note further, that none of the authors in the (Pennay et al., 2014) paper were general practitioners themselves. As a GP, I found the strategies that were provided were broad based and generic, rather than focussed and allowing specific clinical behaviour change (e.g., "Finding a way to connect with patients as part of providing general lifestyle, dietary or mental health advice might be one way of overcoming barriers to addressing heavy alcohol consumption in primary care").

The authors cited (Johnson et al., 2011) earlier (citation 15). As this is a systematic review of the qualitative evidence, I would have expected the authors to have discussed its results in more depth.

Page 7, line 15: "Overall, females were significantly younger than males in this group – d t(857) – 5.64, p < 0.001 (data not shown)". This is a relatively minor point, but the t-test statistics are mostly uninterpretable when presented in text, and is arguably not the important information for the reader in any case. Simply describing the mean and SD of the ages of female and male GP participants would be more helpful.

Page 8, lines 11-13: "Participants were far less likely to agree they had ability to identify at-risk drinking; particularly female participants"

It needs to be noted that this is an interpretation by the authors, and it is unclear whether the survey respondents would have interpreted this statement in the same way. Assuming the text description in Table 2 is verbatim from the questionnaire itself, the specific statement is, "able to tell if patients have alcohol issues". Specifically, "ability to identify" is not conceptually identical to "able to tell if". As a GP, I interpret the first statement as having the skills to assess (which is already covered by the second question in table 2). I interpret "able to tell if" more as my guess or intuition of how often I am misled by patients.

Also, "alcohol issues" is a vague and non-clinical term. It is unclear how the GP respondents interpreted those words. However, I suspect that many GPs will not have interpreted "alcohol issues" as "at-risk" or "risky" drinking. Rather, they may have interpreted this as an alcohol use disorder, or WHO "harmful drinking".

Reflecting on how I would have interpreted that question in a questionnaire, I would have probably identified as a construct, whether I more often than not correctly identify someone as having an alcohol use disorder who is trying to avoid detection in the consultation.

The point that I'm making here, which I also made in the previous review and was similarly commented on by other peer reviewers, is that the language used in the questionnaire has conceptual ambiguities. It is probable, in my view, that some (? many) of the respondents may have interpreted the questions in a manner that is different to the intended concept.

Page 8, line 56-57: "Approximately 30% of participants agreed that they ask about alcohol consumption depending on patients' demographic and socioeconomic characteristics."

The authors will need to provide the actual question/stem here. In figure 1, it states, "Depends on patient's SES/occupation" and "Depends on patient's age/sex/ethnicity". It is unclear what was actually asked so the GP responses are uninterpretable to me. What exactly did the GP respondents agree or not agree to? In this section, the authors state "they ask about alcohol consumption". In a later section, they write "they raised alcohol issues depending on…" There are multiple interpretations. Some GPs might perceive the question as asking whether they ask in addition to usual care, or are prompted to ask by these factors. A negative response may be related to (i) they don't believe that they are relevant, or (ii) they believe they are relevant, but they ask "routinely" anyway, so it does not "depend" on these factors.

Overall, it would be beneficial if a copy of the full questionnaire were available.

Page 9, text and figure 1

To reduce the statistics in text, consider noting that the comparisons are by Chi-squared analysis in the details of figure 1, and place the p-values next to the female-male comparison horizontal bars. This may make the text sections easier to read.

I leave it up to the authors, but some of the actual differences between female and male responses don't appear to be particularly important, even if they were statistically significant. Given the probable significant differences between female and male GPs (these were educational events where the overwhelming majority of attendees were women – it seems quite probable that the male GP participants in particular are not representative of male GPs in

general), it is questionable whether any conclusions that are generalisable can be made about differences in responses between sexes. I raise this from the perspective of whether some of these differences are worthy of noting in text, or whether figure 1 could potentially simply be a table.

Page 10, line 5
"Figure 1" – this is a typo – the authors are referring to figure 2.

Page 15, lines 25 to 28: "In contrast to the reported perspectives of patients,[17] our GP participants overwhelmingly supported the statement that, in a usual month, they were likely to routinely ask patients about their alcohol status."
Although I agree that GPs likely overestimate their asking and recording of alcohol status of their patients, it does need to be noted that citation 17 (Aalto, Pekuri, & Seppä, 2002) was a study conducted in Finland, 18 years ago. The drinking culture in Finland, though having similarities with Australia, is a little unusual in that though binge drinking is especially common, it is usually at home. This has to do with where and how alcohol is available for sale. Northern Europe and Australia are described as having "temperance" or "dry" or drinking culture, where drinking is perceived in a morally ambiguous way socially (The Social Issues Research Centre, 1998). This is likely exaggerated in Finland as compared to Australia – that is, both the normalisation of heavy alcohol consumption, and the stigma of being seen as someone with an alcohol problem.
If the findings that GPs alcohol assessment behaviours are influenced by the society and culture in which they sited (citation 7) (Tam, Zwar, & Markham, 2013), it might be that there are important differences in the phenomena between Finland at the end of the 20th century, and contempory Australia.

Page 15, lines 42-47: "The presentation rankings (see Figure 1) support evidence that GPs tend to rely more on clinical judgements and medical conditions at consultation when deciding to initiate conversations about alcohol with patients, rather than enquiring as routine practice."
I assume the authors are referring to Figure 2.
This statement is likely an over-interpretation. Assuming that the title in figure 2 was the verbatim stem in the questionnaire, "presentations most likely to PROMPT alcohol discussions", then that question asked GPs which of these scenarios effectively are most likely to be directly "caused" by alcohol, and asked them to rank only the top three.
Furthermore, I feel that the authors have not followed through with the significance of their acknowledgement that, "it is possible that 'routinely ask' was not interpreted to mean that GPs universally asked patients at each visit but that they routinely did so in certain circumstances."
An "episode of care" by a GP may take place over several consultations visits. As an illustration, the patient may present with a new issue. The first consultation may be focussed on the immediate clinical assessment to minimise danger, along with requests for further investigation. The second consultation may be to complete the assessment, which has now had the benefit of time to see how it has progressed, along with investigation results. Early management

may be suggested and recommended at this consultation. A third follow up consultation may take place to assess how the new issue has progressed/resolved.

It would seem meaningless for a GP to repeatedly assess alcohol consumption, especially with a formal tool like the AUDIT or AUDIT-C, at each of these near-spaced consultations. The authors should consider that the adults in Australia presents to GPs 5 or 6 times a year (median).

Pages 15-16, lines 55 to 7: "It is of concern, however, that presentations such as 'suspicious or frequent injuries' and 'frequent requests for sickness certificates' were ranked in the top three presentations by 20% or fewer of GPs, given that injuries and work absenteeism are very common outcomes of harmful drinking."

The authors need to better consider what their "concern" is, given that they asked the respondents to only rank 3 items. Are they suggesting that these two items are more PREVALENT than altered LFT results, or suspected clinical depression/anxiety? Abnormal LFT results are very common and in contemporary society, alcohol use is likely the most common single explanation, or an important contributor. The 12-month prevalence of mental illness (most frequently, anxiety and depressive disorders) is 20% in the adult population in Australia.

The use of the 7% of "problems" managed at general practices as an implication of its small number, is misleading without acknowledging the context that the MOST frequent problem managed occurs at only around 10% of consultations in the BEACH dataset. The casemix in general practice is very broad.

To be a little blunt, this statement appears not well informed of the clinical context of family medicine. It strikes me as a somewhat inauthentic piece of criticism of clinical reasoning of practitioners who are in fact, the experts in family medicine, by authors who do not appear to work as clinicians in that context.

Furthermore, this interpretation is limited by the categorical heterogeneity of the items provided in the list, which mixes objective and specific events such as "patient smells of alcohol", and longer term subjective scenarios like "frequent requests for sickness certificates". To be able to make this interpretation of "concern", the authors will need to have clarity as to WHY the respondents answered the way they did. The method of data collection precludes this understand. Strong statements based on conjectures of the clinical reasoning processes of the respondents, in my view, should be avoided.

Page 16, lines 33-35: "Fewer than of males and female GPs agreed they were sufficiently informed about alcohol misuse and related issues during their medical education."

There appears to be a missing value after "fewer than…"

Page 17, lines 26-30: "Routine primary screening with the shorter AUDIT-C might be seen as less time consuming and potentially less intrusive than the full AUDIT has been perceived to be.[42]"

(Beich, Gannik, & Malterud, 2002) is probably the better reference if only one is considered.

Page 17, lines 30-33: "Further research might evaluate the

acceptability of the AUDIT-C for use in routine practice and its impact on GP confidence in identifying at-risk patients."
Citation 7, (Tam et al., 2013) also examined Australian GP perspectives to the AUDIT-C. It was not perceived positively.

Page 17, lines 37-40: "Brief screening tools, such as the AUDIT-C, may also help to address perceptions of 'lack of time', which emerged as an important barrier to routinely enquiring about alcohol consumption"
Although this has been a common assumption, this statement is improbable when considered from the perspective of how consultations actually run, and also from the perspective of history. Screening instruments like the original CAGE and then MAST have been around since the 1970s, and the CAGE is simpler than the AUDIT-C.
If we think about workflow, let's say that an AUDIT-C can be completed, scored, and discussed on average in 30 seconds (which is improbably quick as someone who has actually used the AUDIT and AUDIT-C in actual practice). Implementing this would imply that over a standard work day (30 patients), the time devoted to this single activity is 15 minutes. The opportunity cost is at least a patient consultation per consulting day. Over the course of a month, it will average out to the equivalent of an entire working day per month. What is the benefit as seen at the level of an individual clinician?
This was first clearly described by (Beich, Thorsen, & Rollnick, 2003) in a systematic review and meta-analysis that created a huge controversy in the SBI field. To quote:
"Overall, in 1000 screened patients, 90 screened positive and required further assessment, after which 25 qualified for brief intervention; after one year 2.6 (95% confidence interval 1.7 to 3.4) reported they drank less than the maximum recommended level."
There is further, a disparity between the academically perceived "efficacy" of screening and brief interventions, versus the on-the-ground pragmatically experienced "effectiveness" of SBIs. Much of the evidence described as supportive of SBI in this paper are rather old. Many of those individual papers are subsumed into the Cochrane systematic review by (Kaner et al., 2007) (your citation 11).
There have been a number of pragmatic trials of SBI in general practice since the Kaner meta-analysis to more clearly determine the effectiveness of SBI, and ALL of these have been NEGATIVE (Beich, Gannik, Saelan, & Thorsen, 2007; Butler et al., 2013; Hilbink, Voerman, van Beurden, Penninx, & Laurant, 2012; Kaner et al., 2013; Williams et al., 2014) (see this non-peer reviewed article published in a medical periodical, "The Medical Republic": https://gpunit.org/2015/11/18/does-alcohol-screening-work-in-general-practice/).
The alcohol screening and brief intervention field is somewhat split on the question of how to interpret these new data. However, there is at least a group that believes that SBI, as they have been typically conceived for implementation in family medicine (i.e., for "routine" implementation in standard consulting) might not be effective (Clossick & Woodward, 2014; Nilsen, 2010; Saitz, 2014).
Although I don't necessarily expect the authors to take that position, it should probably at least be acknowledged that the evidence for the routine implementation of SBI in general practice is not as categorically supportive as it is typically made out to be (for instance, in the narrative of this paper).

Page 19, lines 20-22: "Routine use of brief alcohol risk assessment instruments, such as the AUDIT-C, might improve confidence in assessment alcohol issues…"

The authors might want to consider introducing the perspective of "screening as intervention". One of the interpretations of the negative effects of the modern SBI effectiveness trials is that the control groups (who usually received alcohol screening in the form of something like the AUDIT-C, and then a pamphlet) improved just as much (and in several of the trials, MORE) than the intervention groups who also received brief interventions.

The missing link then is one of translation – how do we actually implement something like the AUDIT-C into regular practice? It is improbable that separate questionnaires, especially those that do not integrate with modern electronic health records and the system workflows of real practice will be seen as pragmatic. GPs have been exhorted over decades to use various alcohol screening questionnaires and in my view, this advice in and of itself in unlikely to be effective. We have previously argued this in an opinion piece (Tam, Leong, & Zwar, 2015):

"Newer or updated alcohol-screening questionnaires are unlikely to be the answer on their own. Rather, we [general practitioners] need to be equipped with strategies that can be practically implemented in our local contexts."

A number of modalities have been trialled in recent years such as waiting room electronic surveys on an iPad like device, and GP facilitated online alcohol interventions, may be promising.

References

Aalto, M., Pekuri, P., & Seppä, K. (2002). Primary health care professionals' activity in intervening in patients' alcohol drinking: a patient perspective. Drug Alcohol Depend, 66(1), 39-43.

Beich, A., Gannik, D., & Malterud, K. (2002). Screening and brief intervention for excessive alcohol use: qualitative interview study of the experiences of general practitioners. BMJ, 325(7369), 870.

Beich, A., Gannik, D., Saelan, H., & Thorsen, T. (2007). Screening and brief intervention targeting risky drinkers in Danish general practice--a pragmatic controlled trial. Alcohol Alcohol, 42(6), 593-603. doi:10.1093/alcalc/agm063

Beich, A., Thorsen, T., & Rollnick, S. (2003). Screening in brief intervention trials targeting excessive drinkers in general practice: systematic review and meta-analysis. BMJ, 327(7414), 536-542. doi:10.1136/bmj.327.7414.536

Bendtsen, P., Anderson, P., Wojnar, M., Newbury-Birch, D., Mussener, U., Colom, J., . . . Gual, A. (2015). Professional's Attitudes Do Not Influence Screening and Brief Interventions Rates for Hazardous and Harmful Drinkers: Results from ODHIN Study. Alcohol Alcohol, 50(4), 430-437. doi:10.1093/alcalc/agv020

Butler, C. C., Simpson, S. A., Hood, K., Cohen, D., Pickles, T., Spanou, C., . . . Rollnick, S. (2013). Training practitioners to deliver opportunistic multiple behaviour change counselling in primary care: a cluster randomised trial. BMJ, 346, f1191. doi:10.1136/bmj.f1191

Clossick, E., & Woodward, S. (2014). Alcohol brief interventions in general practice. British Journal of Healthcare Management, 20(10), 468-477. doi:doi:10.12968/bjhc.2014.20.10.468

Hilbink, M., Voerman, G., van Beurden, I., Penninx, B., & Laurant, M. (2012). A randomized controlled trial of a tailored primary care program to reverse excessive alcohol consumption. J Am Board Fam Med, 25(5), 712-722. doi:10.3122/jabfm.2012.05.120070

Johnson, M., Jackson, R., Guillaume, L., Meier, P., & Goyder, E. (2011). Barriers and facilitators to implementing screening and brief intervention for alcohol misuse: a systematic review of qualitative evidence. J Public Health (Oxf), 33(3), 412-421. doi:10.1093/pubmed/fdq095

Kaner, E. F., Beyer, F., Dickinson, H. O., Pienaar, E., Campbell, F., Schlesinger, C., . . . Burnand, B. (2007). Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev(2), CD004148. doi:10.1002/14651858.CD004148.pub3

Kaner, E. F., Bland, M., Cassidy, P., Coulton, S., Dale, V., Deluca, P., . . . Drummond, C. (2013). Effectiveness of screening and brief alcohol intervention in primary care (SIPS trial): pragmatic cluster randomised controlled trial. BMJ, 346, e8501. doi:10.1136/bmj.e8501

Nilsen, P. (2010). Brief alcohol intervention--where to from here? Challenges remain for research and practice. Addiction, 105(6), 954-959. doi:10.1111/j.1360-0443.2009.02779.x

Pennay, A., Lubman, D., & Frei, M. (2014). Alcohol: prevention, policy and primary care responses. Aust Fam Physician, 43, 356-361.

Saitz, R. (2014). The best evidence for alcohol screening and brief intervention in primary care supports efficacy, at best, not effectiveness: you say tomato, I say tomato? That's not all it's about. Addict Sci Clin Pract, 9, 14. doi:10.1186/1940-0640-9-14

Tam, C. W. M., Leong, L. H., & Zwar, N. (2015). Let's listen to patients' and GPs' perspectives on alcohol-screening research. Aust Fam Physician, 44(6), 427-428.

Tam, C. W. M., Zwar, N., & Markham, R. (2013). Australian general practitioner perceptions of the detection and screening of at-risk drinking, and the role of the AUDIT-C: a qualitative study. BMC Fam Pract, 14, 121. doi:10.1186/1471-2296-14-121

The Social Issues Research Centre. (1998). Social and cultural aspects of drinking: a report to the European commission Retrieved from http://www.sirc.org/publik/social_drinking.pdf

Williams, E. C., Rubinsky, A. D., Chavez, L. J., Lapham, G. T., Rittmueller, S. E., Achtmeyer, C. E., & Bradley, K. A. (2014). An early evaluation of implementation of brief intervention for unhealthy alcohol use in the US Veterans Health Administration. Addiction, 109(9), 1472-1481. doi:10.1111/add.12600

REVIEWER	Pablo Barrio
	Addictive Behaviors Unit, Clinic Hospital, Barcelona, Spain
REVIEW RETURNED	06-Sep-2016

| GENERAL COMMENTS | Miller and colleagues tackle a very interesting and important topic: how GPs deal with alcohol in primary care. While this is not a new topic, I think the paper will add relevant information to the literature. However, some points, mostly minor, need to be addressed.
-In the introduction section, I think there is a need to better express why is this study needed. The jump from the second to the third paragraph is a little artificial. I think trying to connect the "dots" and making the case for the study will substantially improve the flow of the manuscript.
-The description of the survey is a little bit confusing and not easy to follow. Maybe authors could state something like: the survey consisted mainly of 4 parts. The first was desgined to gather sociodemographic information. The second…. |

-The only major concern I would state for this paper is in the Data analysis section. It looks like the authors prepared some questions with multiple options, in an ordinal scale, but then they reconverted this into a binary variable ("likert scale agreement categories were dichotomized…."). While I am aware that it would not be the first investigation doing this, it implies a loss of information, and I think authors should at least give an explanation about why they have done so. Alternatively, ordinal regression models could also be constructed.

-The second paragraph in the Data analysis section, second sentence, I think a parenthesis would work better than a coma for the section (likelihood to routinely enquire…..at-risk drinking).

-In the results section the authors talk about managing and identification alcohol problems. I think these are two separate concepts, and maybe in the introduction they should be better delineated, stressing the importance of both identificating and managing.

-In the description of the multivariate analysis, I would recommend authors to rewrite results in order to make it much more easily readable and understandable. Lots of reasons (which usually imply long sentences) are mixed and followed one by another, therefore I think readers will end up not really remembering which are the significant facilitators and barriers. Maybe authors could only describe the significant ones.

-while I think the discussion is well written and the relevant findings are properly analyzed, what I miss is more related literature. Maybe it would be worth to look at other non-Australia studies tackling the same topic. For example there are many of this kind in Europe (for example :

Ann Fam Med. 2015 Jan-Feb;13(1):28-32. doi: 10.1370/afm.1742.General practitioners recognizing alcohol dependence: a large cross-sectional study in 6 European countries. Rehm J1, Allamani A1, Della Vedova R1, Elekes Z1, Jakubczyk A1, Landsmane I1, Manthey J2, Moreno-España J1, Pieper L1, Probst C1, Snikere S1, Struzzo P1, Voller F1, Wittchen HU1, Gual A1, Wojnar M)

OR

Professional's Attitudes Do Not Influence Screening and Brief Interventions Rates for Hazardous and Harmful Drinkers: Results from ODHIN Study

Preben Bendtsen, Peter Anderson, Marcin Wojnar, Dorothy Newbury-Birch, Ulrika Müssener, Joan Colom, Nadine Karlsson, KrzysztofBrzózka, Fredrik Spak, Paolo Deluca, Colin Drummond, Eileen Kaner, Karolina Kłoda, Artur Mierzecki, Katarzyna Okulicz-Kozaryn, KathrynParkinson, Jillian Reynolds, Gaby Ronda, Lidia Segura, Jorge Palacio, Begoña Baena, Luiza Slodownik, Ben van Steenkiste, AmyWolstenholme, Paul Wallace, Myrna N. Keurhorst, Miranda G.H. Laurant, Antoni Gual

-Finally, there are some minor typos that should be corrected (figure 1 when it should say figure 2, Fewer than___of males and females (percentage missing), etc)

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Michael Tam

Review
Summary
Thank you for this opportunity to review this paper. This study examined potential barriers and facilitators to discussions between patients and GPs using questionnaires. The topic is one of some interest to clinical medicine and public health, especially to Australian general practice.

I need to declare that I was serendipitously one of the peer reviewers on a previous version of this paper that was submitted to the Medical Journal of Australia. I am pleased to see that the authors have made use of the criticisms and suggestions given by all the peer reviewers previously and have substantially improved their paper.

My overall conclusion is that the data and results of this study contributes to the field and I would be supportive for this study to be published. All critical comments following needs to be taken from this point of view. However, there will need to be some important revisions, especially with regards to how the results are interpreted, and the limitations acknowledged.

Major issues

1. Bivariate and multivariate analyses and interpretations The interpretation of these quantitative analyses are entirely dependent on whether the qualitative labels of the factors (which do not appear to have been constructed in a planned or peer reviewed manner) accurately represent the actual phenomenon that was measured from the participant responses. I have some doubts related to the construct validity of the questionnaires, and some major doubts to the qualitative labels used to group responses.

Response:
While the survey was not formally evaluated in a validation study, its items were developed using the peer reviewed literature with the input of experts in the field of alcohol and general practice. Further, the survey was piloted among the research team and its network (including affiliated GPs) and was further refined after feedback from the pilot participants. In developing the survey, the priority was to keep the survey as brief as possible to increase participation and completeness of data. This meant that that it was not possible to include previously validated, often lengthy, instruments while retaining focus on our aims. However, our responses to the piloting suggested consistency between the intended meaning and its comprehension. We do believe the questionnaire was constructed to measure the phenomenon of interest – i.e. the factors that are enhance or impede GP-patient alcohol discussions.

We have added the following statement to the Limitations and Recommendations' section in the paper (page 20):

"Although it was informed by the literature, and developed in consultation with experts in the field, the survey did not include previously validated measures. Feedback from the pilot and responses to the survey suggested consistency between the intended meaning of questions and their comprehension, however confirming our conclusions with further study is recommended."

2. Importantly, I note that the labels of the factors in this version of the manuscript are substantially different to the labels used in the prior version. This further reduces my confidence of the validity of

the labels. Moreover, this suggests to me that at least part of the model development was EXPLORATORY and was not designed a priori.

For instance, table 4 has three outcomes in the model labelled "Likely to routinely ask about alcohol consumption" and "Confidence in ability to assess and manage alcohol issues", and "Usually not enough time to enquire about each patient's alcohol intake", with 4, 5, and 3 labelled factors respectively. Your previous paper with the equivalent model had three outcomes labelled "Does not routinely ask", "Low confidence" and "Insufficient skills", with 2, 3, and 2 factors respectively. Between these two drafts, the framing of the model outcomes is quite different, and, there are different numbers of factors in the analyses.

The authors must explain their data analysis process and more clearly note that the analyses were exploratory. They will need to better explain how they came to the qualitative labels for their factors – some of this will be aided by including the actual questionnaire(s) used.

Response:
Following the previous submission, the paper was substantially amended including a complete revisiting of the analysis. The broad approach to the analysis was unchanged between papers – specifically aiming to identify factors bivariately associated with the main outcomes (likelihood to routinely inquire about alcohol and confidence in managing alcohol issues) and then to test the independence of these factors using log binomial models. The over-representation of female participants was anticipated so controlling for this was assumed in planning for our multivariate analyses. However, specific aspects changed in line with the comments from the previous review, as follows:

• The main problem for the previous review was that the presented data labels appeared to be ambiguous raising concerns the information could be invalid. Thus, on the re-analysis, the data labels are changed to more closely reflect the actual questions asked, rather than the short-hand versions which previously appeared.

• The previous review pointed out the loss of data in dichotomising the Likert responses in a manner that excluded the 'don't know' or 'uncommitted' categories. In the new analyses, the dichotomised variables included all data points – e.g. agreement categories were dichotomised to create measures of agreement – combining 'strongly agree' with 'agree' and combining 'neither', 'disagree' and 'strongly disagree'. Likert scales for likelihood ('very likely' to 'very unlikely') were dichotomised in a similar way. This approach increased the numbers in the analyses, which resulted in increased power to identify statistically significant bivariate differences and, ultimately, an increased number of factors that could be tested in the multivariate models.

• Concentrating on 'agreement' and 'likelihood' has increased the consistency of the new analyses as well as reframing the models more positively. Lack of clarity was identified as an issue in our previous analysis, which also focussed on negative (rather than positive) valences and conveyed a sense of 'risk' and deficit that was not intended.

Thanking the reviewer for the suggestion we have submitted the questionnaire with this review.

3. It is my view that the results of this study as they have been presented, though interesting and perhaps indicative of certain phenomena, contain much more uncertainty than acknowledged by the authors. These GPs are not representative of GPs in general. The construction of the questionnaire, the terminology used, and the labelling and focus of subsequent analyses are likely to have been influenced and constrained by pre-existing conceptual biases held by the investigators on the topic. The authors place substantial confidence in their analytic approach and model development, and this is not a confidence that I share having read through two versions of this paper.

Response:
The representativeness of the sample is discussed in the 'Limitations' section. We trust that our responses to the reviewers above concerns have decreased some of the uncertainties about the

questionnaire construction and labelling of outcomes. As can be seen from the submitted survey, the revisited analysis now contains labels that more closely reflect the questions posed to the GPs. We do agree that the previous analysis seen by the reviewer could well have been interpreted as negatively biased due to its positioning of associated factors from a deficit model.

4. Lastly, even assuming that the analysis and model developed is largely valid within the context of the limited range of participants, the "adjusted agreement ratios" in the final model (table 4), for "Likely to routinely ask about alcohol consumption", range from the most positive at 1.12, to least positive at 0.80. To me, this appears to mean that all the factors described have a relatively small impact on the likelihood of "routinely ask about alcohol". It is possible, if not probable, that this study did not capture the range of explanations and factors that contribute to the majority of the variation in this phenomenon.

Response:
We agree with the reviewer that the identified differences were not large. The ability for a cross-sectional survey to fully explore all related factors exhaustively demand a pre-existing level of certainty about an issue. While the questions were based on information from the relatively few published studies on this issue, together with the knowledge of our networked experts, the level of pre-existing knowledge did not allow for an exhaustive exploration in our brief survey. Complex phenomena, such as GP practice, are likely to have a number of contributing factors. Nonetheless, we have been able to identify some independent predictors of likelihood to routinely ask about alcohol, some of which may be amenable to modification. We would also expect our results to be expanded upon in future studies to obtain the full story.

For greater clarity, we have added the following statement to the end of the 'Limitations and Recommendations' section (page 21):

"Our analyses identified independent predictors of likelihood to routinely ask about alcohol and confidence in managing alcohol issues, however the absolute differences were not large – possibly indicating that not all contributing factors were uncovered. Importantly, however, some of the contributory factors identified in our study are potentially modifiable. More research on this complex issue is required, however our findings could help to inform the development of actions aimed at enhancing preventive practice."

Minor issues

5. Page 4, Line 38: "Little is known about how GPs initiate discussions about alcohol with their patients in Australian general practice settings…" This statement is referenced to citation 15, (Johnson, Jackson, Guillaume, Meier, & Goyder, 2011), a systematic review of qualitative evidence of barriers and facilitators of implementing alcohol screening and brief interventions, and does not seem entirely appropriate. I wonder whether this is an error?

Response:
With our thanks for noting this error; we have now updated the reference to:

New reference:
[15] Wellard L, Corsini N, Hughes C: Discussing alcohol and cancer with patients: Knowledge and practices of general practitioners in New South Wales and South Australia. Australian Family Physician 2016, 45(8):588-593.

6. Page 4, Line 45: "Pennay et al. propose confidence in managing alcohol issues as a potential

barrier to screening and brief intervening in alcohol issues."

This is a somewhat narrow point to take from the "treatment: strategies" section of the (Pennay, Lubman, & Frei, 2014) paper in Australian Family Physician. This was from a single statement in that paper. This statement in (Pennay et al., 2014) in turn referenced the phase 3 of the WHO Collaborative Project from two decades ago – it was published in 1998. I generally agree with the concept that confidence in managing alcohol issues is something that has an effect on GP alcohol screening behaviours, but a better reference should be found.

A theoretical framework that is used in the field are the concepts of "role security" and "therapeutic commitment". Confidence would be subsumed within these concepts.

Although there has been enthusiasm from various researchers that influencing one or both of these factors may improve alcohol screening and brief intervention rates (which is an intuitive hypothesis), there is some empirical evidence that it does not. For instance, this recent paper by (Bendtsen et al., 2015) found that professional attitudes did not influence SBI rates. One observation to reflect on is the conceptual bias that many in the field have in seeking clinician-related factors/deficiencies as the explanation for the phenomena (e.g., not enough knowledge, not enough resources, wrong attitude) and plan for intervention (e.g., education, support, and mentorship respectively). It might be worth considering that this study similarly focussed almost exclusively on clinician-related factors and thus, any description of the phenomena, as well as recommendations for intervention are likely to be incomplete.

Response:
We have included another reference and changed the text as follows (page 4):

"In a review of the world wide literature, Yoast et al [17] found that lack of self-efficacy was the chief reason proposed by physicians for not providing care around substance use. They state (page 83): "They do not have familiarity or experience with screening and intervention techniques, lack confidence in their skills to intervene, and doubt the effectiveness of the help they had provide to patients.""

New reference:
[17] Yoast RA, Wilford BB, Hayashi SW: Encouraging Physicians to Screen for and Intervene in Substance Use Disorders: Obstacles and Strategies for Change. Journal of Addictive Diseases 2008, 27(3):77-97.

We agree that research focussing purely on the 'GP deficit' model will not be able to provide a complete picture of the issue and how to address it. We have added the following statement to the end of the 'Limitations and Recommendations' section in the paper (page 21):

"Further work focusing on the contribution of factors external to the GP, such as patient characteristics and expectations, would also allow for a more complete picture of this issue."

7. Existing literature on barriers and facilitators Although I respect the authors' prerogative into placing weight on specific ideas and concepts in prior literature in terms of barriers and facilitators, I feel that they need to explain why so much weight was given to the review article by (Pennay et al., 2014) in the paragraph on page 4, lines 23 to 52.

With all due respect to (Pennay et al., 2014), I do not believe that they discussed the causes of poor uptake of alcohol screening and brief interventions by GPs in much depth in their paper. In defence of their paper, this was not their intention in the section where they describe it. I would note further, that none of the authors in the (Pennay et al., 2014) paper were general practitioners themselves. As a GP, I found the strategies that were provided were broad based and generic, rather than focussed and allowing specific clinical behaviour change (e.g., "Finding a way to connect with patients as part

of providing general lifestyle, dietary or mental health advice might be one way of overcoming barriers to addressing heavy alcohol consumption in primary care").

Response:
We have removed the Pennay reference in favour of an international review (Yoast et al 2008) and a European survey of more 2345 GPs. Consequently, we have changed the text in the 'Introduction' as follows (page 5);

In their survey of more than 2300 GPs in eight European countries, Anderson et al [18] found that GPs with higher education about alcohol and GPs with confidence in managing patients with alcohol issues were more likely to manage such patients, while those who believed alcohol was a 'disease' or those who viewed drinking as a personal, rather than medical, responsibility tended to manage fewer patients with alcohol issues."

New reference:
[18] Anderson P, Wojnar M, Jakubczyk A, Gual A, Segura L, Sovinova H, Csemy L, Kaner E, Newbury-Birch D, Fornasin A et al: Managing Alcohol Problems in General Practice in Europe: Results from the European ODHIN Survey of General Practitioners. Alcohol and Alcoholism 2016, 51(5):630-630.

8. The authors cited (Johnson et al., 2011) earlier (citation 15). As this is a systematic review of the qualitative evidence, I would have expected the authors to have discussed its results in more depth.

Response:
This reference was replaced by the Wellard et al (2016) reference in response to the reviewer's comment no. 5 (above).

9. Page 7, line 15: "Overall, females were significantly younger than males in this group – d t(857) – 5.64, p < 0.001 (data not shown)".
This is a relatively minor point, but the t-test statistics are mostly uninterpretable when presented in text, and is arguably not the important information for the reader in any case. Simply describing the mean and SD of the ages of female and male GP participants would be more helpful.

Response:
Since the analysis was done separately to what is provided in the table, we believe it is important to provide the t-test statistics to demonstrate what process was undertaken. We do agree, however, that highlighting the respective ages of the participants is valuable here. We have amended the text in the 'Results' section as follows (page 7):

"Overall, females were significantly younger than males in this group – mean 46.5 versus 52.2 years – d t(857) -5.64, p<0.001 (data not shown)."

10. Page 8, lines 11-13: "Participants were far less likely to agree they had ability to identify at-risk drinking; particularly female participants"
It needs to be noted that this is an interpretation by the authors, and it is unclear whether the survey respondents would have interpreted this statement in the same way. Assuming the text description in Table 2 is verbatim from the questionnaire itself, the specific statement is, "able to tell if patients have alcohol issues".
Specifically, "ability to identify" is not conceptually identical to "able to tell if". As a GP, I interpret the first statement as having the skills to assess (which is already covered by the second question in table

2). I interpret "able to tell if" more as my guess or intuition of how often I am misled by patients. Also, "alcohol issues" is a vague and non-clinical term. It is unclear how the GP respondents interpreted those words. However, I suspect that many GPs will not have interpreted "alcohol issues" as "at-risk" or "risky" drinking. Rather, they may have interpreted this as an alcohol use disorder, or WHO "harmful drinking".

Reflecting on how I would have interpreted that question in a questionnaire, I would have probably identified as a construct, whether I more often than not correctly identify someone as having an alcohol use disorder who is trying to avoid detection in the consultation.

The point that I'm making here, which I also made in the previous review and was similarly commented on by other peer reviewers, is that the language used in the questionnaire has conceptual ambiguities. It is probable, in my view, that some (? many) of the respondents may have interpreted the questions in a manner that is different to the intended concept.

Response:
The wording in the survey, is "I am usually able to tell if patients are having alcohol-related problems." When the statement was posed, we did intend it as a statement of intuition or guess rather than a question about skills and capacity. We have now changed the wording of the statement in the 'Results' section (page 8) as follows:

"Participants were far less likely to agree they were able to 'tell' if their patients had alcohol issues; particularly female participants."

Similar to other surveys, there is potential for ambiguity. Although we do not necessarily believe that it can be argued in this instance that different respondent interpretations fundamentally challenge the analysis, we do believe that it is appropriate to comment on this and have added the following statement to the 'Discussion' section (page 17):

It is possible that GPs interpreted the statement about being able to tell if their patients had alcohol issues differently – e.g. either they are able to intuitively know (or guess), regardless of any information offered by the patient, or that they have confidence in their clinical ability to identify such patients. Indeed, agreeing they were 'usually able to tell' if their patients had alcohol issues independently predicted GPs confidence in their ability to manage alcohol issues and both of these factors independently predicted likelihood to routinely ask about alcohol consumption. Nonetheless, strategies aimed at enhancing assessment abilities might increase likelihood of routinely asking about alcohol.

11. Page 8, line 56-57: "Approximately 30% of participants agreed that they ask about alcohol consumption depending on patients' demographic and socioeconomic characteristics."
The authors will need to provide the actual question/stem here. In figure 1, it states, "Depends on patient's SES/occupation" and "Depends on patient's age/sex/ethnicity". It is unclear what was actually asked so the GP responses are uninterpretable to me. What exactly did the GP respondents agree or not agree to? In this section, the authors state "they ask about alcohol consumption". In a later section, they write "they raised alcohol issues depending on…"
There are multiple interpretations. Some GPs might perceive the question as asking whether they ask in addition to usual care, or are prompted to ask by these factors. A negative response may be related to (i) they don't believe that they are relevant, or (ii) they believe they are relevant, but they ask "routinely" anyway, so it does not "depend" on these factors.
Overall, it would be beneficial if a copy of the full questionnaire were available.

Response:
The exact wording for the two statement is: "I raise alcohol issues with patients dependent on the patient's age, sex or ethnicity" and "I am more likely to bring up alcohol issues with patients from

particular occupations or from particular socio-economic groups". Although we believe that these statement are quite unambiguous, in addition to supplying a copy of the survey (agreed in response to the reviewer's comment no. 2), for clarity we have amended the text in the 'Results' section (page 9):

"Approximately 30% of participants agreed that they 'raise alcohol issues with patients dependent on the patient's age, sex or ethnicity' and they were 'more likely to bring up alcohol issues with patients from particular occupations or from particular socio-economic groups."

12. Page 9, text and figure 1
To reduce the statistics in text, consider noting that the comparisons are by Chi-squared analysis in the details of figure 1, and place the p-values next to the female-male comparison horizontal bars. This may make the text sections easier to read.
I leave it up to the authors, but some of the actual differences between female and male responses don't appear to be particularly important, even if they were statistically significant. Given the probable significant differences between female and male GPs (these were educational events where the overwhelming majority of attendees were women – it seems quite probable that the male GP participants in particular are not representative of male GPs in general), it is questionable whether any conclusions that are generalisable can be made about differences in responses between sexes. I raise this from the perspective of whether some of these differences are worthy of noting in text, or whether figure 1 could potentially simply be a table.

Response:
On the whole, we would prefer to leave the male and female responses in place as they do provide further rationale for controlling for sex in the multivariate analyses. We did try to amend the chart but inserting the p-values as suggested by the reviewer did not improve the readability. We have, however, changed the font type and size in figures 1 and 2 and we trust they are both now more easy to read (see 'Results', pages 10-11)

13. Page 10, line 5
"Figure 1" – this is a typo – the authors are referring to figure 2.

Response:
We thank the reviewer for the observation – this has now been corrected.

14. Page 15, lines 25 to 28: "In contrast to the reported perspectives of patients,[17] our GP participants overwhelmingly supported the statement that, in a usual month, they were likely to routinely ask patients about their alcohol status."
Although I agree that GPs likely overestimate their asking and recording of alcohol status of their patients, it does need to be noted that citation 17 (Aalto, Pekuri, & Seppä, 2002) was a study conducted in Finland, 18 years ago. The drinking culture in Finland, though having similarities with Australia, is a little unusual in that though binge drinking is especially common, it is usually at home. This has to do with where and how alcohol is available for sale.
Northern Europe and Australia are described as having "temperance" or "dry" or drinking culture, where drinking is perceived in a morally ambiguous way socially (The Social Issues Research Centre, 1998). This is likely exaggerated in Finland as compared to Australia – that is, both the normalisation of heavy alcohol consumption, and the stigma of being seen as someone with an alcohol problem.
If the findings that GPs alcohol assessment behaviours are influenced by the society and culture in which they sited (citation 7) (Tam, Zwar, & Markham, 2013), it might be that there are important differences in the phenomena between Finland at the end of the 20th century, and contempory Australia.

Response:

We agree and thank the reviewer for pointing out the unsuitability of the cited study in the light of its age and context. We have amended the text in the 'Discussion' section (page 15) as follows:

"In contrast to the reported perspectives of GPs elsewhere,[20] our GP participants overwhelmingly supported the statement that, in a usual month, they were likely to routinely ask patients about their alcohol status."

New reference:
[20] Mules T, Taylor J, Price R, Walker L, Singh B, Newsam P, Palaniyappan T, Snook T, Ruselan M, Ryan J et al: Addressing patient alcohol use: a view from general practice. J Prim Health Care 2012, 4(3):217-222.

15. Page 15, lines 42-47: "The presentation rankings (see Figure 1) support evidence that GPs tend to rely more on clinical judgements and medical conditions at consultation when deciding to initiate conversations about alcohol with patients, rather than enquiring as routine practice."
I assume the authors are referring to Figure 2.
This statement is likely an over-interpretation. Assuming that the title in figure 2 was the verbatim stem in the questionnaire, "presentations most likely to PROMPT alcohol discussions", then that question asked GPs which of these scenarios effectively are most likely to be directly "caused" by alcohol, and asked them to rank only the top three.
Furthermore, I feel that the authors have not followed through with the significance of their acknowledgement that, "it is possible that 'routinely ask' was not interpreted to mean that GPs universally asked patients at each visit but that they routinely did so in certain circumstances."
An "episode of care" by a GP may take place over several consultations visits. As an illustration, the patient may present with a new issue. The first consultation may be focussed on the immediate clinical assessment to minimise danger, along with requests for further investigation. The second consultation may be to complete the assessment, which has now had the benefit of time to see how it has progressed, along with investigation results. Early management may be suggested and recommended at this consultation. A third follow up consultation may take place to assess how the new issue has progressed/resolved.
It would seem meaningless for a GP to repeatedly assess alcohol consumption, especially with a formal tool like the AUDIT or AUDIT-C, at each of these near-spaced consultations. The authors should consider that the adults in Australia presents to GPs 5 or 6 times a year (median).

Response:
As can be seen in the now submitted survey, the specific wording for the question was "In order of importance (1 being most important), what are the three most likely types of presentations that might prompt you to initiate conversations about alcohol with your patients? (please write the numbers 1, 2 and 3 in the boxes)." Given that the very first option provided was 'I usually ask about alcohol', we believe that misinterpreting this question might have been minimised. To clarify this, we have now amended the text in the 'Discussion' section (page 18) as follows:

"Rather than ranking the top three presentations that prompt alcohol discussions, it is possible that some GPs may have ranked presentations that can be directly attributed to alcohol. The positioning of 'I usually ask about alcohol consumption' as the first option, however, was likely to have minimised the possibility of misunderstanding."

We agree that the statement 'i.e. at each consultation' is misleading as it was more important to capture the situation of a patient presenting with a new set of problems. We have amended the clarifying statement in the 'Discussion' section (page 18) as follows:

"…(i.e. at each new presentation)."

16. Pages 15-16, lines 55 to 7: "It is of concern, however, that presentations such as 'suspicious or frequent injuries' and 'frequent requests for sickness certificates' were ranked in the top three presentations by 20% or fewer of GPs, given that injuries and work absenteeism are very common outcomes of harmful drinking."

The authors need to better consider what their "concern" is, given that they asked the respondents to only rank 3 items. Are they suggesting that these two items are more PREVALENT than altered LFT results, or suspected clinical depression/anxiety? Abnormal LFT results are very common and in contemporary society, alcohol use is likely the most common single explanation, or an important contributor. The 12-month prevalence of mental illness (most frequently, anxiety and depressive disorders) is 20% in the adult population in Australia.

The use of the 7% of "problems" managed at general practices as an implication of its small number, is misleading without acknowledging the context that the MOST frequent problem managed occurs at only around 10% of consultations in the BEACH dataset. The casemix in general practice is very broad.

To be a little blunt, this statement appears not well informed of the clinical context of family medicine. It strikes me as a somewhat inauthentic piece of criticism of clinical reasoning of practitioners who are in fact, the experts in family medicine, by authors who do not appear to work as clinicians in that context.

Furthermore, this interpretation is limited by the categorical heterogeneity of the items provided in the list, which mixes objective and specific events such as "patient smells of alcohol", and longer term subjective scenarios like "frequent requests for sickness certificates". To be able to make this interpretation of "concern", the authors will need to have clarity as to WHY the respondents answered the way they did. The method of data collection precludes this understand. Strong statements based on conjectures of the clinical reasoning processes of the respondents, in my view, should be avoided.

Response:

According to the two sources that were referenced here, 10-18% of injuries in emergency departments are alcohol related (where presumably only the more serious injuries are seen), and that alcohol misuse is also strongly linked to absenteeism at work. The concern here stems from the potential for clinical judgements that do not always include consideration of important indicators of problematic drinking. We agree that the statement may over-state this concern and have amended the text in the 'Discussion' section as follows (page 16-17):

"The presentation rankings (see Figure 2) support evidence that GPs tend to rely more on clinical judgements and medical conditions at consultation when deciding to initiate conversations about alcohol with patients, rather than enquiring as routine practice.[22] Given that from 2013 to 2014 LFTs were only ordered at 2.4% of GP encounters, and anxiety and depression represented 7% of problems managed at general practices,[4] it is likely that a large number of patients with problematic alcohol behaviours could remain undetected should these presentations be the primary prompt for enquiry. Injuries and work absenteeism are very common outcomes of harmful drinking,[23, 24] yet 'suspicious or frequent injuries' and 'frequent requests for sickness certificates' were ranked in the top three presentations by 20% or fewer of GPs in our survey. Mitchell et al[25] concluded that healthcare professionals struggle to identify problem drinking in clinical practice, using clinical judgements to identify half of those with alcohol use disorder based on clinical judgement and accurately noting alcohol use disorder in only a third of actual cases. It is important to note that our participants were asked to rank the top three most important presentations, rather than rate an exhaustive list of potential presentations, and non-selection of an option should not be interpreted as meaning GPs ignored that presentations. Further, among the broad range of presenting problems in general practice, presentations most likely to prompt alcohol discussions are likely to be highly contextual. Yet

finding ways to enhance clinical decision making in regard to problematic alcohol use is likely to be beneficial given our results."

New reference:
[25] Mitchell AJ, Meader N, Bird V, Rizzo M: Clinical recognition and recording of alcohol disorders by clinicians in primary and secondary care: meta-analysis. The British Journal of Psychiatry 2012, 201(2):93.

17. Page 16, lines 33-35: "Fewer than of males and female GPs agreed they were sufficiently informed about alcohol misuse and related issues during their medical education."
There appears to be a missing value after "fewer than…"

Response:
With thanks, this has now been amended (page 17):

"Fewer than half of male and female GPs…"

18. Page 17, lines 26-30: "Routine primary screening with the shorter AUDIT-C might be seen as less time consuming and potentially less intrusive than the full AUDIT has been perceived to be.[42]"
(Beich, Gannik, & Malterud, 2002) is probably the better reference if only one is considered.

Response:
With thanks for the suggestion, we have included it along with the Brady et al (2002) reference to include the Australian and International experience.

19. Page 17, lines 30-33: "Further research might evaluate the acceptability of the AUDIT-C for use in routine practice and its impact on GP confidence in identifying at-risk patients."
Citation 7, (Tam et al., 2013) also examined Australian GP perspectives to the AUDIT-C. It was not perceived positively.

Response:
We thank the reviewer, whose paper was cited earlier in the manuscript, but does seem relevant in this section of the discussion. We have amended this in the 'Discussion' section as follows (page 19):

"In their qualitative study, Tam et al[7] found that GP responses to even the shorter version were largely negative. Future larger studies across a range of primary health care settings may find ways to increase the acceptability and usefulness of such tools for routine practice as well as assess their impact on GP confidence in identifying at-risk patients."

20. Page 17, lines 37-40: "Brief screening tools, such as the AUDIT-C, may also help to address perceptions of 'lack of time', which emerged as an important barrier to routinely enquiring about alcohol consumption"
Although this has been a common assumption, this statement is improbable when considered from the perspective of how consultations actually run, and also from the perspective of history. Screening instruments like the original CAGE and then MAST have been around since the 1970s, and the CAGE is simpler than the AUDIT-C.
If we think about workflow, let's say that an AUDIT-C can be completed, scored, and discussed on average in 30 seconds (which is improbably quick as someone who has actually used the AUDIT and

AUDIT-C in actual practice). Implementing this would imply that over a standard work day (30 patients), the time devoted to this single activity is 15 minutes. The opportunity cost is at least a patient consultation per consulting day. Over the course of a month, it will average out to the equivalent of an entire working day per month. What is the benefit as seen at the level of an individual clinician?

This was first clearly described by (Beich, Thorsen, & Rollnick, 2003) in a systematic review and meta-analysis that created a huge controversy in the SBI field. To quote:

"Overall, in 1000 screened patients, 90 screened positive and required further assessment, after which 25 qualified for brief intervention; after one year 2.6 (95% confidence interval 1.7 to 3.4) reported they drank less than the maximum recommended level."

There is further, a disparity between the academically perceived "efficacy" of screening and brief interventions, versus the on-the-ground pragmatically experienced "effectiveness" of SBIs. Much of the evidence described as supportive of SBI in this paper are rather old. Many of those individual papers are subsumed into the Cochrane systematic review by (Kaner et al., 2007) (your citation 11). There have been a number of pragmatic trials of SBI in general practice since the Kaner meta-analysis to more clearly determine the effectiveness of SBI, and ALL of these have been NEGATIVE (Beich, Gannik, Saelan, & Thorsen, 2007; Butler et al., 2013; Hilbink, Voerman, van Beurden, Penninx, & Laurant, 2012; Kaner et al., 2013; Williams et al., 2014) (see this non-peer reviewed article published in a medical periodical, "The Medical Republic": https://gpunit.org/2015/11/18/does-alcohol-screening-work-in-general-practice/).

The alcohol screening and brief intervention field is somewhat split on the question of how to interpret these new data. However, there is at least a group that believes that SBI, as they have been typically conceived for implementation in family medicine (i.e., for "routine" implementation in standard consulting) might not be effective (Clossick & Woodward, 2014; Nilsen, 2010; Saitz, 2014).

Although I don't necessarily expect the authors to take that position, it should probably at least be acknowledged that the evidence for the routine implementation of SBI in general practice is not as categorically supportive as it is typically made out to be (for instance, in the narrative of this paper).

Response:
In our response to the reviewer's previous comment (no. 19) and to the reviewer's next comment (no.20) we have made a few changes that may help to address this issue, including referring to negative and neutral assessments of the effectiveness of SBI. In our next response (to comment no. 21) we also point out that we do not advocate routine SBI but suggest that screening tools may assist GPs to initiate discussions about alcohol consumption and identifying patients with problems in relation to alcohol.

21. Page 19, lines 20-22: "Routine use of brief alcohol risk assessment instruments, such as the AUDIT-C, might improve confidence in assessment alcohol issues…"

The authors might want to consider introducing the perspective of "screening as intervention". One of the interpretations of the negative effects of the modern SBI effectiveness trials is that the control groups (who usually received alcohol screening in the form of something like the AUDIT-C, and then a pamphlet) improved just as much (and in several of the trials, MORE) than the intervention groups who also received brief interventions.

The missing link then is one of translation – how do we actually implement something like the AUDIT-C into regular practice? It is improbable that separate questionnaires, especially those that do not integrate with modern electronic health records and the system workflows of real practice will be seen as pragmatic. GPs have been exhorted over decades to use various alcohol screening questionnaires and in my view, this advice in and of itself in unlikely to be effective. We have previously argued this in an opinion piece (Tam, Leong, & Zwar, 2015):

"Newer or updated alcohol-screening questionnaires are unlikely to be the answer on their own. Rather, we [general practitioners] need to be equipped with strategies that can be practically

implemented in our local contexts."
A number of modalities have been trialled in recent years such as waiting room electronic surveys on an iPad like device, and GP facilitated online alcohol interventions, may be promising.

Response:
In our paper we do not necessarily advocate for universal introduction of SBI but propose brief screening of alcohol as a means to assist with initiating discussions about alcohol consumption and identifying patients with problems in relation to alcohol. We have amended some of the text to better reflect this position in the 'Discussion' section as follows (page 19):

"There is also evidence that the screening process itself may be the more effective component, with positive outcomes not varying according to intensity of the intervention in Emergency Department patients.[47] Assessing the usefulness of using various formats of brief screening in primary health care (for example, whether completed by the patient, GP or practice nurse) for the purposed of facilitating discussion could form the basis for future study.

Brief alcohol screening tools may assist GPs to initiate discussions about alcohol in a non-personalised way."

New reference:
[47] Drummond C, Deluca P, Coulton S, Bland M, Cassidy P, Crawford M, Dale V, Gilvarry E, Godfrey C, Heather N et al: The Effectiveness of Alcohol Screening and Brief Intervention in Emergency Departments: A Multicentre Pragmatic Cluster Randomized Controlled Trial. PLoSone 214, 9 (6):e99463.

Reviewer: 2
Reviewer Name: Pablo Barrio
Institution and Country: Addictive Behaviors Unit, Clinic Hospital, Barcelona, Spain
Competing Interests: None declared

Miller and colleagues tackle a very interesting and important topic: how GPs deal with alcohol in primary care. While this is not a new topic, I think the paper will add relevant information to the literature. However, some points, mostly minor, need to be addressed.

1. In the introduction section, I think there is a need to better express why is this study needed. The jump from the second to the third paragraph is a little artificial. I think trying to connect the "dots" and making the case for the study will substantially improve the flow of the manuscript.

Response:
We agree that summarising the gaps and significance of the study is required and have added the following paragraph (and new reference) as suggested in the 'Introduction' section as follows (page 5):

According to Knox et al[19], almost 90% of Australians visit a GP on at least one occasion each year with an average of six visits per person annually, providing GPs with substantial opportunity to discuss alcohol behaviour with their patients. Given their substantiated efficacy in prevention, more information is required from GPs regarding their perceptions around discussing alcohol consumption. This may inform the development of strategies assisting GP's to discuss alcohol intake with their patients more frequently, effectively manage any alcohol issues identified and ultimately reduce the impact of alcohol-related harm in the population.

New reference:

[19] Knox SA, Harrison CM, Britt HC, Henderson JV: Estimating prevalence of common chronic morbidities in Australia. Medical Journal of Australia 2008, 189(2):66-70.

2. The description of the survey is a little bit confusing and not easy to follow. Maybe authors could state something like: the survey consisted mainly of 4 parts. The first was desgined to gather sociodemographic information. The second….

Response:
Reflecting the suggestion of the reviewer, we have amended the description of the survey in the 'Methods' section as follows (page 6):

"To reduce the burden on participants, the survey was presented on one page containing four types of questions. The first part of the survey included demographic information such as gender, age, years worked in general practice, and employment fraction (full-time or part-time). Forming the two main outcome variables, participants were asked to indicate how likely they were, in a usual month of general practice, to routinely ask about their patients' alcohol consumption, as well as their confidence and ability to assess and manage at-risk drinking. In the third grouping of questions, participants rated their level of agreement with 14 statements regarding potential barriers and facilitators to asking about alcohol on a five-point Likert scale (1 = 'strongly disagree' to 5 = 'strongly agree'). Participants also ranked the top three patient presentations that would prompt them to ask about a patient's alcohol consumption, from 12 possible options (an 'other' option was also offered). Other questions relating to cancer and alcohol will form part of a separate report."

3. The only major concern I would state for this paper is in the Data analysis section. It looks like the authors prepared some questions with multiple options, in an ordinal scale, but then they reconverted this into a binary variable ("likert scale agreement categories were dichotomized…."). While I am aware that it would not be the first investigation doing this, it implies a loss of information, and I think authors should at least give an explanation about why they have done so. Alternatively, ordinal regression models could also be constructed.

Response:
We have used the binary logistic method because the cell numbers at the extremes tended to be too small for meaningful analysis. For instance, our main outcome variable 'likely to routinely inquire about alcohol' had the following counts for the Likert scale: 'very likely'=387; 'likely' =375; 'neither likely nor unlikely'=73; 'unlikely'=34; and 'very unlikely'=5. A similar pattern was seen on the other outcome variable ('confidence in ability to manage and assess alcohol issues') as well as a number of the 14 barrier and facilitator questions. The analytical decision was to dichotomise by modelling on agreement as this provided a positive valence and meant that the 'neither' categories were not excluded. To clarify this, we have amended the description in the 'Methods' section as follows (page 7):

To compensate for small counts in the extreme categories of the independent variables (and many of the independent variables), Likert scale agreement categories were dichotomised to create measures of agreement in subsequent analyses – combining 'strongly agree' with 'agree' and combining 'neither', 'disagree' and 'strongly disagree'. Likert scales for likelihood ('very likely' to 'very unlikely') were dichotomised in a similar way.

4. The second paragraph in the Data analysis section, second sentence, I think a parenthesis would work better than a coma for the section (likelihood to routinely enquire…..at-risk drinking).

Response:
We thank the reviewer for this suggestion and have amended the sentence as advised.

5. In the results section the authors talk about managing and identification alcohol problems. I think these are two separate concepts, and maybe in the introduction they should be better delineated, stressing the importance of both identificating and managing.

Response:
We agree that these two concepts should be more delineated. We believe that the concepts are currently discussed more separately in the 'Discussion' but we have clarified the distinction by adding a statement to the 'Introduction' section as follows (page 4):

"Identifying potentially harmful alcohol consumption is a necessary first step in managing those issues once they are detected and there is some evidence to suggest that the two concepts, detection and management, may be linked via reduced GP confidence in their ability to undertake follow up interventions.[16] In a review of the world wide literature, Yoast et al [17] found…"

6. In the description of the multivariate analysis, I would recommend authors to rewrite results in order to make it much more easily readable and understandable. Lots of reasons (which usually imply long sentences) are mixed and followed one by another, therefore I think readers will end up not really remembering which are the significant facilitators and barriers. Maybe authors could only describe the significant ones.

Response:
We agree that the first paragraph in particular is not clear and have amended it to increase the readability, while retaining sufficient information for readers to understand how the final models were derived. The amended paragraph in the 'Results' section is as follows (page 13):

"For the outcome 'routinely asking about alcohol' the following factors were no longer significant, after adjusting for other factors, and fell out of all the models: agreeing they had sufficient skills; agreeing they were sufficiently educated to manage alcohol issues; asking depending on the patients age, sex and ethnicity; fear of negative responses from patients; and anticipating communication difficulties. In our final model (see Table 4), confidence in their ability to manage and assess alcohol was the most important factor (agreement ratio of 1.12, 10% absolute difference). Agreeing they were 'usually able to tell' if patients have alcohol issues, and being female also independently predicted likelihood to routinely ask about alcohol consumption. Lack of time was a significant barrier to routinely asking, with those agreeing with the statement about lack of time being 20% less likely to 'routinely ask' (agreement ratio 0.80), with an absolute agreement difference of 18%.

7. while I think the discussion is well written and the relevant findings are properly analyzed, what I miss is more related literature. Maybe it would be worth to look at other non-Australia studies tackling the same topic. For example there are many of this kind in Europe
(for example :
Ann Fam Med. 2015 Jan-Feb;13(1):28-32. doi: 10.1370/afm.1742.General practitioners recognizing alcohol dependence: a large cross-sectional study in 6 European countries. Rehm J1, Allamani A1, Della Vedova R1, Elekes Z1, Jakubczyk A1, Landsmane I1, Manthey J2, Moreno-España J1, Pieper L1, Probst C1, Snikere S1, Struzzo P1, Voller F1, Wittchen HU1, Gual A1, Wojnar M)
OR
Professional's Attitudes Do Not Influence Screening and Brief Interventions Rates for Hazardous and Harmful Drinkers: Results from ODHIN Study
Preben Bendtsen, Peter Anderson, Marcin Wojnar, Dorothy Newbury-Birch, Ulrika Müssener, Joan

Colom, Nadine Karlsson, KrzysztofBrzózka, Fredrik Spak, Paolo Deluca, Colin Drummond, Eileen Kaner, Karolina Kłoda, Artur Mierzecki, Katarzyna Okulicz-Kozaryn, KathrynParkinson, Jillian Reynolds, Gaby Ronda, Lidia Segura, Jorge Palacio, Begoña Baena, Luiza Slodownik, Ben van Steenkiste, AmyWolstenholme, Paul Wallace, Myrna N. Keurhorst, Miranda G.H. Laurant, Antoni Gual

Response:
We thank the reviewer for the suggestion and trust that the addition of six international papers as part of our response to other comments has adequately addressed their concern.

8. Finally, there are some minor typos that should be corrected (figure 1 when it should say figure 2, Fewer than___of males and females (percentage missing), etc)

Response:
With thanks again, we have amended these errors as suggested.

VERSION 2 – REVIEW

REVIEWER	Michael Tam General Practice Unit, South Western Sydney Local Health District and Ingham Institute, NSW Australia School of Public Health and Community Medicine, UNSW Australia Some of my papers are referenced in this paper.
REVIEW RETURNED	06-Oct-2016

GENERAL COMMENTS	Thank you for the opportunity to review this revision. I congratulate the authors for making use of the feedback and improving their manuscript. I am largely satisfied with the changes and I feel that the majority of the issues have been addressed well. I have only minor issues that I would like the authors to address in their final version of the paper. Well done. Minor issues Page 4, Line 28 – "Red Book" Since the review, the 9th edition of the Red Book has been released. The authors may want to update the reference. Page 15, Lines 53-57 – Injuries and absenteeism I have commented on this in the previous review, and I note the author's response. I would counter that injuries as seen in emergency departments do not translate to injuries seen in general practice. The majority of injuries seen in general practice would not be due to alcohol misuse. The context is very different – ED is not general practice. Also, although individuals with alcohol use disorders may have frequent work absenteeism, the reverse logic does not hold. The majority of individuals coming to see a GP for an illness certificate (repeat or otherwise) do so not because of alcohol use problems, but because of chronic illness (both physical and mental). The questionnaire did not ask GPs whether alcohol use disorders are associated with work absenteeism. Rather, it asked whether requests for frequent illness certificates would prompt

questioning about alcohol, amongst a list of other scenarios.
The updated language in this section of the paper has been improved, but in my view, the rationale underlying the statement "Injuries and work absenteeism are very common outcomes of harmful drinking,[23, 24] yet 'suspicious or frequent injuries' and 'frequent requests for sickness certificates' were ranked in the top three presentations by 20% or fewer of GPs in our survey" is weak. I agree with the overarching thesis of this section of the paper, that GPs use clinical judgements of individuals to identity risky drinking (and so they should!) and tend not to use more systematic or health screening approaches. The use of the results on injuries and work absenteeism, in my view, do not especially strengthen or support this idea.
I do not consider this to be a critical issue, and the existing text is publishable. However, I feel this section would be improved and streamlined by simply removing the interpretation of results on injuries and work absenteeism as a noteworthy point.

Page 17, Lines 41-43 – AUDIT-C and WHO
I'm not sure it would be accurate to describe the AUDIT-C as a WHO developed modification of the AUDIT. My understanding is that the consumption items of the AUDIT was first tested as a brief screening tool by Bush and colleagues (who labelled it the "AUDIT-C") in a Veteran Affairs Population: http://archinte.jamanetwork.com/article.aspx?articleid=208954 Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA, for the Ambulatory Care Quality Improvement Project (ACQUIP). The AUDIT Alcohol Consumption Questions (AUDIT-C): An Effective Brief Screening Test for Problem Drinking. Arch Intern Med. 1998;158(16):1789-1795. doi:10.1001/archinte.158.16.1789.

REVIEWER	Pablo Barrio
	Addictive Behaviors Unit, Clinic Hospital, Barcelona, Spain
	I have received consulting fees from Lundbeck and travel grants from Pfizer.
REVIEW RETURNED	04-Oct-2016

GENERAL COMMENTS	I believe authors made a great effort addressing all reviewers' comments in a very expeditive manner.
	I think both my minor and major concerns have been satisfactorily addressed, and therefore I believe this paper is ready for publication.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Michael Tam

Thank you for the opportunity to review this revision. I congratulate the authors for making use of the feedback and improving their manuscript. I am largely satisfied with the changes and I feel that the majority of the issues have been addressed well.

I have only minor issues that I would like the authors to address in their final version of the paper. Well done.

Minor issues
1. Page 4, Line 28 – "Red Book"
Since the review, the 9th edition of the Red Book has been released. The authors may want to update the reference.

Response:
We thank the reviewer for the suggestion and have updated the reference and the page number citation for the quote.

2. Page 15, Lines 53-57 – Injuries and absenteeism I have commented on this in the previous review, and I note the author's response. I would counter that injuries as seen in emergency departments do not translate to injuries seen in general practice. The majority of injuries seen in general practice would not be due to alcohol misuse. The context is very different – ED is not general practice. Also, although individuals with alcohol use disorders may have frequent work absenteeism, the reverse logic does not hold. The majority of individuals coming to see a GP for an illness certificate (repeat or otherwise) do so not because of alcohol use problems, but because of chronic illness (both physical and mental). The questionnaire did not ask GPs whether alcohol use disorders are associated with work absenteeism. Rather, it asked whether requests for frequent illness certificates would prompt questioning about alcohol, amongst a list of other scenarios.
The updated language in this section of the paper has been improved, but in my view, the rationale underlying the statement "Injuries and work absenteeism are very common outcomes of harmful drinking,[23, 24] yet 'suspicious or frequent injuries' and 'frequent requests for sickness certificates' were ranked in the top three presentations by 20% or fewer of GPs in our survey" is weak.
I agree with the overarching thesis of this section of the paper, that GPs use clinical judgements of individuals to identity risky drinking (and so they should!) and tend not to use more systematic or health screening approaches. The use of the results on injuries and work absenteeism, in my view, do not especially strengthen or support this idea.
I do not consider this to be a critical issue, and the existing text is publishable. However, I feel this section would be improved and streamlined by simply removing the interpretation of results on injuries and work absenteeism as a noteworthy point.

Response:
We agree that the reference to the finding does not add to the matter under discussion and have removed the statement as recommended by the reviewer.

3. Page 17, Lines 41-43 – AUDIT-C and WHO
I'm not sure it would be accurate to describe the AUDIT-C as a WHO developed modification of the AUDIT. My understanding is that the consumption items of the AUDIT was first tested as a brief screening tool by Bush and colleagues (who labelled it the "AUDIT-C") in a Veteran Affairs Population: http://archinte.jamanetwork.com/article.aspx?articleid=208954
Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA, for the Ambulatory Care Quality Improvement Project (ACQUIP). The AUDIT Alcohol Consumption Questions (AUDIT-C): An Effective Brief Screening Test for Problem Drinking. Arch Intern Med. 1998;158(16):1789-1795. doi:10.1001/archinte.158.16.1789.

Response:
We thank the reviewer for this advice and have updated the text as follows:

"Bush et al[44] subsequently developed a modification of the AUDIT instrument, the three-item

AUDIT-C"

[44]Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA, for the Ambulatory Care Quality Improvement Project (ACQUIP): The AUDIT Alcohol Consumption Questions (AUDIT-C) - An Effective Brief Screening Test for Problem Drinking. Archives of Internal Medicine 1998, 158:1789-1795.

Reviewer: 2
Reviewer Name: Pablo Barrio
Institution and Country: Addictive Behaviors Unit, Clinic Hospital, Barcelona, Spain
Competing Interests: I have received consulting fees from Lundbeck and travel grants from Pfizer.

I believe authors made a great effort addressing all reviewers' comments in a very expeditive manner. I think both my minor and major concerns have been satisfactorily addressed, and therefore I believe this paper is ready for publication.

Response:
We thank the reviewer for this response and for their contribution to our manuscript.