

Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: *Screening, Behavioral Counseling, and Referral in Primary Care To Reduce Alcohol Misuse*

Draft review available for public comment from October 13, 2011 to November 10, 2011.

Research Review Citation: Jonas DE, Garbutt JC, Brown JM, Amick HR, Brownley KA, Council CL, Viera AJ, Wilkins TM, Schwartz CJ, Richmond ER, Yeatts J, Swinson Evans T, Wood SD, Harris RP. Screening, Behavioral Counseling, and Referral in Primary Care To Reduce Alcohol Misuse. Comparative Effectiveness Review No. 64. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 12-EHC055-EF. Rockville, MD: Agency for Healthcare Research and Quality. July 2012. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Structured Abstract	<p>Because "Referral" is mentioned in the title of the report, and because SAMHSA funds a very large program of screening, brief intervention, REFERRAL, and treatment, it would be useful I think to mention the evidence or lack there-of for referral in the abstract, briefly. This finding (and the lack of evidence identified for those with dependence identified by screening) will be useful to highlight because it will help guide the need for research of great practical importance. Referral would apply only to those with dependence.</p> <p>Also for the abstract, I think it is very important to mention KQ1 and the absence of a single study that addressed it. Something like, 'We identified no evidence regarding the effectiveness of screening and intervention versus no screening,' or something like that.</p>	<p>We have added the point about referral (and insufficient evidence) as well as the point about KQ1 to the abstract.</p> <p>"No studies randomized subjects, practices, or providers to screening and a comparator and none of the included studies reported follow-up with referrals as an outcome."</p>
Peer Reviewer #4	Structured Abstract	<p>p. vii, ll. 32-35: Since mortality and alcohol-related liver problems are not likely to be factors in most of the patient populations studied, it seems odd (at least) that they are cited as health outcomes considered. It would make as much sense to cite colds, another factor not likely to be affected by the intervention. If the data allow, it would be far more likely to find evidence of injuries and other acute conditions, for example. I would suggest dropping the sentence or revising it to indicate that the studies generally did not provide information upon which traditional, disease related health outcomes could be measured.</p>	<p>We disagree and this seems to be a minority opinion voiced by a single reviewer. Several TEP members and other peer reviewers, as well as some members at the USPSTF meeting, indicated the importance of reporting whether there is any evidence on these outcomes (mortality and alcohol-related liver problems).</p> <p>We agree that evidence of injuries and other acute conditions is also very important and we have added that to the abstract (insufficient evidence)</p>
Peer Reviewer #4	ES / General	<p>Overall, the Executive Summary contains many issues of fact, interpretation, and conceptualization that, in my judgment, require significant revision. Additionally, some of the issues make me wonder if the authors have sufficient experience in this field to conduct the study.</p>	<p>This is a general comment that sets up a number of other specific comments (see other comments from reviewer #4 and our responses), primarily focused on the background information rather than the methods, results, or discussion. Of note, none of the other reviewers had similar comments and they generally included comments about how well the introduction and literature review was written, highlighting that it is concise and builds a strong argument for examining screening and brief counseling services primary care settings. The comment is stated in a strongly negative manner, yet we were able to address each of the specific comments that followed with fairly simple edits.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Introduction	TABLE ES-1 is confusing but maybe it is just a typographical error. This comment also applies to Table 1 in the full report. The first row is bolded which implies that it includes everything below it, which I don't think is intentional (or correct). (It looks like hazardous use includes abuse, dependence etc). Also see below re the use of "recommended" amounts.	We fixed the typo that the first row is no longer bolded.
Peer Reviewer #1	ES / Introduction	Also in the definition of risky use in Table ES-1 I would suggest including the concept that it is 'consumption levels that increase the risk for health consequences.' The reason for this edit is that physicians and patients alike are often confused by recommendations in this area and pointing out it isn't an arbitrary limit, but rather an amount that risks health harm would make the terms clearer.	We have added the concept that risky use is based on consumption levels that increase the risk for health consequences, as suggested.
Peer Reviewer #1	ES / Introduction	You might also add "problem drinking" or "problem use" to the table because it often appears in the literature. There was a paper titled this in Alcohol Health and Research (NIAAA journal) some years back. Problem use is use with consequences that do not meet criteria for a disorder (you might also add that alcohol use disorder means harmful/abuse or dependence). The rationale for adding problem use to the list is that the Table does not include a category for these folks (people who do not meet the diagnostic criteria but who drink risky amounts and have had consequences).	<p>We have not added problem drinking or problem use because the definitions in published literature are very heterogeneous (it is used to mean different things across various articles). In fact, we think this term should be avoided if possible to minimize confusion. If it were generally defined as the reviewer defines it here, it could potentially be a useful term (but it generally is not defined that way).</p> <p>The article the reviewer references here actually defines it differently than he does. The article uses the term "problem drinking" to describe people who drink heavily or experience occasional problems from drinking but who do not have a history of severe physical dependence on alcohol. [Walitzer KS. Treating problem drinking. Alcohol Research & Health 1999;23(2): 138]. This actually would include harmful drinking and alcohol abuse, rather than identifying a separate, more discrete group.</p> <p>The IOM (1990) report, defined problem drinkers as those having moderate to substantial levels of alcohol consumption and possibly alcohol abuse but who are unlikely to have physical dependence on alcohol.</p>

Commentator & Affiliation	Section	Comment	Response
			<p>Many of the studies from Fleming et al define problem alcohol use as one or more positive responses to the CAGE questionnaire and evidence of alcohol-related medical or behavioral problems. But, this does not distinguish it from harmful use or from abuse (this definition encompasses harmful use and abuse).</p> <p>We would prefer to describe people who drink risky amounts and have had consequences, as just that---risky drinkers who have had consequences. We don't see the value in having additional categories, especially when an additional category doesn't have any clear utility.</p>
Peer Reviewer #1	ES / Introduction	<p>I believe there is an important error or typo in the paragraph under Table ES-1. My NEJM paper is referenced supporting a prevalence of misuse of 50%. But that is not correct. Table 1 in my paper indicates a prevalence of 30% of risky use. Problem use, abuse and dependence are not mutually exclusive from risky use (i.e. people with dependence are included among those who drink excessively) so one should not add the prevalences on the left hand side of the table). Instead of 50% prevalence my paper can be cited as a 30% estimate. This is similar to the 28% estimate published by NIAAA in their clinicians guide on the page found here http://pubs.niaaa.nih.gov/publications/Practitioner/CliniciansGuide2005/clinicians_guide17.htm (of note, since about 65% of the population drinks, that does mean that about 50% of those who drink report having had risky amounts but that is not the same as 50% of the population).</p>	<p>We have corrected the 50%, it now says "about 30%".</p>
Peer Reviewer #1	ES / Introduction	<p>Also for this paragraph it would be helpful to point out that 1 in 5 who screen positive in primary care will have dependence (4 in 5 will not). Sources might include Saitz et al 2009 J Gen Intern Med Single item screening, and 1998 Manwell, Fleming MF.</p>	<p>We have also added this point that about 1 in 5 who screen positive... (and thank you for the suggested references as both of those provide good support for this statement).</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Introduction	ES: Last para of Background in Exec Summary: You might include “family history of dependence” among the conditions/risk factors to whom standard recommended limits may not apply, as well as “mental health conditions” (along with the medical conditions).	The examples provided in this sentence are those for whom alcohol use is completely contraindicated. We do not agree that people with a family history of dependence should be included in such a list. Also, we consider mental health conditions to be included in medical conditions, and don’t wish to dichotomize the two.
Peer Reviewer #1	ES / Introduction	Table ES-2. Can brief interventions also include simply advice (i.e. not necessarily motivational) to cut down or quit? As written it seems like this would not count as a brief behavioral intervention but perhaps it should? Is there some (old or recent) literature on this from Nick Heather or Paul Wallace in the UK?	We have added to the information in Table ES-2 to further describe behavioral counseling interventions in primary care, with new references added. We edited the title as we realize it was missing the word “counseling” (now reads “What are brief behavioral counseling interventions...?”) as we believe this helps to clarify the interventions of interest. We have also added the following to the table, with references: “Behavioral counseling interventions include the range of personal counseling and related behavior-change interventions that are employed in primary care to help patients change health-related behaviors. “Counseling” here denotes a cooperative mode of work demanding active participation from both patient and clinician that aims to facilitate the patient’s independent initiative.”
Peer Reviewer #1	ES / Introduction	ES: Also last para of background: The definition of risky amounts should include the words “on average” as in “1 or fewer standard drinks per day on average” and should also include the maximum daily amounts (i.e. no more than 3 for women, 4 for men in a day). Those additions would more accurately reflect the limits, and would avoid confusion for patients and clinicians who may mistakenly read this and think the limit is no more than one drink in a day ever for women, for example.	We have revised this to clarify and avoid confusion: “Maximum recommended consumption is 3 or fewer standard drinks per day (7 per week) for adult women and for anyone older than 65 years of age, and 4 or fewer standard drinks per day (14 per week) for adult men.
Peer Reviewer #1	ES / Introduction	Introduction: Screening and Behavioral Counseling. I see here that a SAMHSA definition is used for “brief intervention.” This explains why the focus on “motivational discussion.” The research is much broader than this. I would recommend against using SAMHSA’s definition as authoritative and instead use what is in the peer reviewed studies of alcohol brief interventions.	See response to comment 30 above. We have added new text to define behavioral counseling interventions in primary care (which is more broad). We still also mention SAMHSA’s definition for brief intervention as we believe that is also an important one to include.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Introduction	In the full report Introduction I find this clause “problem drinkers who have risky or harmful alcohol use but do not meet the DSM-IV-TR criteria for abuse or dependence” which supports the need to include “problem drinking” in the Table.	We have edited this sentence to avoid use of the term problem drinkers (related to response to comment 27 above).
Peer Reviewer #1	ES / Introduction	This sentence is not accurate: “Patients receiving referrals to specialty care based on positive screening results appear more likely to accept appointments for alcohol-related counseling than those receiving usual care (i.e., those who were not screened).57” In that study ALL were screen-identified. The finding was that when nurses were able to refer directly, more patients accepted appointments than in standard physician care. BUT even more important is that even in that group 90% did not accept appointments. Furthermore this was just about appointment acceptance not attendance at an appointment.	Thank you for the clarification. We have reviewed the study again and have revised that sentence in the introduction to better reflect their findings and to focus on the more important point about not accepting appointments. “In addition, most patients with a positive screening result for a drinking problem are unlikely to accept referrals for alcohol-related counseling.”
Peer Reviewer #2	ES / Introduction	a minor comment is to unbold the text in the row of the table (e.g., ES-1) that defines risky drinking. When reading it in boldface it seemed like a heading rather than a definition.	We fixed this.
Peer Reviewer #2	ES / Introduction	The key questions are generally okay. I have one comment. KQ1 does not map well to the figure depicting the key questions (e.g., ES-4). The KQ says “screening followed by behavioral counseling”, but the figure shows a direct line from screening to outcomes. It is not possible to counsel without screening, so I believe the KQ is flawed or else the diagram needs to be modified. I am having a hard time figuring out how to modify the diagram. The KQ could be modified to “Does screening for alcohol misuse lead to reduced morbidity, mortality, etc.”	We believe the KQ is a good representation of what we were intending to answer and that this is a challenge with the analytic framework to draw something that reflects this. With KQ1, we were truly trying to look for evidence that would randomize subjects, providers, or practices to screening or a comparator (likely no screening, usual care), and that would also provide intervention for those people screening positive. To address the comment, we have modified the Analytic Framework to attempt to illustrate this more clearly by adding the word “Intervention” along the overarching arrow for KQ1.
Peer Reviewer #4	ES / Introduction	p. ES1, l. 12: Economic Costs of Excessive Alcohol Consumption in the U.S., 2006, Ellen E. Bouchery, MS, Henrick J. Harwood, Jeffrey J. Sacks, MD, MPH, Carol J. Simon, PhD, Robert D. Brewer, MD, MSPHAm J Prev Med 2011;41(5):516–524 is a more recent update than Harwood for footnote 6.	Thank you for the suggestion. We have updated the information in the report based on this article (and we now reference this source)

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Introduction	Ibid. I. 18: Injuries should not be limited by “trauma-related”, and specific mention should be made of fires, drowning, homicide, motor vehicle crashes, and pedestrian injuries. Acute alcohol-related harm exceeds that of chronic. See CDC cost fractions.	We agree and we have revised the text to remove “trauma-related” as a descriptor of injuries, and we now specifically mention all of the things listed in this comment, as well as others reported by the CDC to be related to alcohol.
Peer Reviewer #4	ES / Introduction	Ibid. II. 8-10: Citing ranges of various estimates based upon a 10+ year old paper is too easy a way out. There are recent, highly reliable studies based upon good surveys that could be used to provide an accurate picture, rather than this misleading combination of ranges	<p>We have added data from more recently-published surveys. These data fall within our originally presented ranges. We believe it is important to keep the other information as well, because it provides information to demonstrate that there is a range across various populations (regional variation),</p> <p>We have edited the text to read: “Older studies report a range of risky drinkers from 4 percent to 29 percent across primary care populations, with prevalence estimates of 0.3 to 10.0 percent for harmful drinkers and 2.0 to 9.0 percent for alcohol dependence.¹⁵ More recent data from the American Academy of Family Physicians National Research Network reveals that 21.3% of primary care patients reported hazardous drinking (based on the 3 quantity and frequency questions from the AUDIT-C).{Vinson, 2010} Alcohol dependence has lifetime prevalence rates on the order of 17 percent for men and 8 percent for women;¹⁸ prevalence of current dependence (within the last 12 months and as defined by the DSM-IV) is approximately 4% in the general adult population.{Hasin, 2007} Approximately one in five of those who screen positive for unhealthy alcohol use in primary care will have alcohol dependence (4 in 5 will not).^{16,17} Rates of alcohol-use disorders among medical outpatients are similar to those seen in the general population and are generally higher in males and younger people of all races/ethnicities.^{15,19}”</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Introduction	Ibid. II. 10-11: Lifetime prevalence is not a useful measure for dependence in this context, where the patient's present condition is the issue for medical identification and intervention. It would be far better to cite current condition data for dependence within the year, which is under 4% nationally and is likely to range up from less than 1% in primary care.	We agree and have added current condition prevalence data in the general population; see response to previous comment with the new wording of the text quoted above.
Peer Reviewer #4	ES / Introduction	Ibid. II. 17-22: These are NOT the NIAAA recommended guidelines. They are the old USDA guidelines for average daily consumption. This is a mistake of monumental proportions indicating either a gross lack of care or significant ignorance in this field.	These were the correct NIAAA recommendations, from averaging the recommended weekly amount to convert it to a daily average. We see how/why this was unclear based on the way we wrote it (failing to add "on average"). We have revised this to more directly match the wording of the NIAAA recommended guidelines: "Maximum recommended consumption is 3 or fewer standard drinks per day (7 per week) for adult women and for anyone older than 65 years of age, and 4 or fewer standard drinks per day (14 per week) for adult men."
Peer Reviewer #4	ES / Introduction	p. ES2, II. 32-33: If the CAGE is going to be cited here, some explanation of how it might possibly be used to identify the target population. You will need specifically to explain how an instrument that contains only questions about symptoms of dependence can be used to identify a target population the vast majority of which is not likely to have any symptoms of dependence. Please note that the fact that the CAGE has been used by providers to identify abuse and dependence is NOT evidence that it ought to be used to identify the population this paper designates as its target. So why is it mentioned here?	We have explained this in the report, in the ES results and in the results. We make it very clear that CAGE is not a good screening test for risky/hazardous drinking. We still mention the CAGE in the place related to this comment b/c this is in the intro where we're listing all of the instruments that have been studied. The sentence says: "The most commonly studied instruments include the..." and we go on to list them.
Peer Reviewer #4	ES / Introduction	Ibid. I. 33: Why "versions of" a very specific reference to something? There are no versions of what NIAAA recommends. There are, however, different versions of the single question screen.	Good point, we have clarified this by revising the text. It now says "versions of the single question screen."
Peer Reviewer #4	ES / Introduction	Ibid. II. 54-55: The assumption stated about brief interventions is not dependent on "if they are effective". They are not effective with many patients, but the assumption stands.	Good point, we have deleted "if they are effective" from the sentence.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Introduction	<p>p. ES3, ll. 3-4: This sentence suggests another total lack of understanding of the field. There is a vast literature on risks associated with different levels of alcohol consumption. Indeed, the USDA and NIAAA recommendations for moderate consumption, cited (incorrectly) previously are based upon those studies. And the authors proceed to cite some such evidence. So what does this sentence mean?</p>	<p>The reviewer is misreading or misunderstanding the sentence. The comment suggests a lack of understanding of how interventions and behavior change are assessed. The sentence actually agrees with what he states here (that there is a literature on risks associated with different levels of consumption). What it points out is that there is not literature associating the risks with <i>changes</i> in drinking behavior. It states “Little experimental evidence supports this assumption, and most epidemiologic evidence relates health outcomes to <u>existing</u> drinking behaviors rather than to <u>changes</u> in drinking behaviors.” And we go on to say: “Cross-sectional and cohort studies have consistently related high average alcohol consumption to short- or long-term health consequences.” Nevertheless, two reviewers commented on this sentence or were confused by it, and it was not necessary to make the point that we were going for in that paragraph; thus, we decided to delete the sentence.</p> <p>Regarding the NIAAA recommendations, we revised our wording of the recommendations as described in a response above (4 comments prior to this one).</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Introduction	Ibid. I 17: The term “treatment” is misleading here. “Treatment” has a fixed connotation in the alcohol field, and brief intervention by that definition is not “treatment”. Further, it is misleading to suggest that any significant number of patients identified in primary care by screening will be likely to meet criteria for dependence and thus require “treatment”. In many primary care settings of both clinical trials and SAMHSA’s SBIRT program, less than 1% do so. Note also that the term is repeated several times in the following paragraph.	We agree that this was not a good word choice. We have changed it to “interventions” to pick a term that could include behavioral intervention or interventions for those with alcohol dependence. The sentence now reads: “In everyday practice, screening and screening-related assessment procedures are necessary to identify the range of alcohol users in order to offer appropriate interventions.” We have made similar edits to address the repeated use of the word in the following paragraph and we word searched the entire report and replaced treatment with intervention or interventions whenever it had been used in the manner similar to the sentence identified by this reviewer.
Peer Reviewer #4	ES / Introduction	Ibid. II. 21-22: I seriously doubt the claim that 20% of primary care patients misuse alcohol. The figure is more likely 10-15% depending on the region of the country, socioeconomic group, and the percentage of female patients. And I also doubt the same regarding diabetes and hypertension, on which there are far better data. But certainly a better source should be used than the one cited.	We have edited the sentence to say that they represent as much as 20%... We have also added additional references to the sentence to support the statement. The range of prevalence in primary care has been reported anywhere from 4% to 29%, depending on the population (region, etc.) and we have included that detail in the report. So, saying up to 20% here in the executive summary is conservative.
Peer Reviewer #4	ES / Introduction	p. ES4, I. 12: Some explanation or example of “intermediate outcomes” seems to be required. Are these the same as in KQ 1 or something different?	We have added a footnote to this Table (the one describing the KQs that this comment refers to) that describes the intermediate outcomes of interest

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Introduction	<p>p. ES4, Figure 1: Is there an explanation of why “college students” are identified as a separate class? Is there an expectation that adults between the ages of 18 and 25 would have different behaviors or responses to medical services if they are in “college” rather than not?</p> <p>It seems that some of these categories are based upon age, medical condition (pregnancy, severity (of what?)), race/gender/ethnicity, but some are also based upon 3 locus. Some explanation seems warranted.</p>	<p>We agree with the implication of the comment, that we don’t expect substantially different behaviors between college students and young adults between the ages of 18 and 25. In fact, we have lumped those two together in the results/synthesis of the literature, with one section on Young adults and college students in each part of the report (rather than 2 separate sections for them). We have changed the analytic framework figure to read “Young adults/college students” to make this more clear.</p> <p>We have further explanation of the analytic framework and these subgroups of interest in the full report. Due to space constraints, and to focus on the results and interpretation of results, we have not included that level of detail in the executive summary.</p>
Vincent Fonseca, MD, MPH, FACPM	ES / Introduction	<p>page ES-2, top: “Though estimating the prevalence of alcohol misuse is challenging, it has been estimated that 40-50 percent of the U.S. population are affected, with the majority of these individuals engaging in what is considered risky drinking.³” This is a mistake: not even in the US adult population vs the US population does any article state a prevalence of alcohol misuse as high as 40%. The citation doesn’t state that figure except in referring ED and trauma patients: “The prevalence of unhealthy use is 7 to 20 percent or more among outpatients, 30 to 40 percent among patients in emergency departments, and 50 percent among patients with trauma.^{11,12}”.</p> <p>Please put an estimate from an article or report that actually estimates the prevalence in the US (either adult or combined adult/adolescent population). The NSDUH would be a good source.</p>	<p>See similar comment from Review 1 (comment 28). We have revised this to: “Though estimating the prevalence of alcohol misuse is challenging, it has been estimated that about 30 percent of the U.S. population is affected, with the majority of these individuals engaging in what is considered risky drinking.³” (we reference the article in NEJM by Saitz et al.)</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Introduction	Page 3, lines 39-43. The list of elements of brief intervention includes cognitive-behavioral therapy, which is a general term, along with elements that are often components of CTB (e.g., keeping diaries, action plans, etc.). The list also combines treatment components with modes of delivery. It might be better to list CBT elements in parentheses after CBT rather than list them separately. For example, "Brief alcohol interventions of varying length and number, cognitive behavioral strategies (e.g., action plans, drinking diaries, stress management, problem solving). Interventions may be delivered via face to face sessions, written self-help materials or telephone counseling."	We have edited this almost verbatim to the suggested wording, but with slight variation, taking into consideration comments of other reviewers and input from CBT experts at UNC. New version: "Brief alcohol interventions can include advice, feedback, motivational interviews of varying length and number, or cognitive behavioral strategies (e.g., self-completed action plans, written health education or self-help materials, drinking diaries, problem solving exercises to complete at home). Interventions may be delivered via face to face sessions, written self-help materials, computer, or telephone counseling."
Peer Reviewer #2	Introduction	Page 5, line 16. the text indicates that the topic was nominated by "a member" of the USPSTF. The USPSTF engages in a formal topic prioritization process. While the topic may have been nominated by a member, it was selected by the entire task force through the topic prioritization process.	Thank you, we revised that text to clarify that it was selected by the TF through the topic prioritization process.
Peer Reviewer #6	Introduction	The introduction is well written and provides helpful and well-researched background.	Thank you
Peer Reviewer #7	Introduction	The introduction and literature review is well written, concise and builds a strong argument for examining screening and brief counseling services primary care settings.	Thank you
Peer Reviewer #8	Introduction	<p>Introduction - population health versus individual health considerations</p> <p>In devising their analytic framework (described in the later Methods section), the authors address but do not fully explicate the many purposes of screening and brief intervention (and referral for treatment) for alcohol misuse. It may be useful to add some material along the lines described below, taking care not to obscure the central messages. At one level, screening and brief counselling (referral for treatment, less so) is a population health strategy that is a form of secondary intervention. It seeks to identify high-risk subgroups within populations, in this case the identification of subgroups whose alcohol intake places them at elevated risk of a range of harmful consequences.</p> <p>Following this, it seeks by using primary health care as an interface, to provide brief counselling to facilitate reduction in the overall alcohol intake of these subgroups and thereby reduce the prevalence of alcohol-related harm in the community. The rationale</p>	Thank you for these comments. It is always interesting to think about the implications of questions for the overall population and for individuals. The questions and results/conclusions in this report, being based primarily on meta-analysis of randomized controlled trials, mainly provide population level information to help determine whether we should screen and provide behavioral interventions for those with alcohol misuse (since the results are average reduction in consumption or risk difference, for example). The results, however, have relevance for individuals with alcohol misuse as well and for individual decision-making. The questions and results can have relevance for both populations and for individuals, it does not have to be one or the other; and we don't

Commentator & Affiliation	Section	Comment	Response
		<p>of working through primary care is that in many countries 80-95% of people visit their primary care physician in any 12-month period, with the average number of visits being between 2 and 5, depending on the country. In this way screening and brief counselling sits between (i) primary prevention (which targets the entire population through such approaches as legislation (e.g. minimum drinking ages and taxation), education and other community approaches) and (ii) treatment aimed at those individuals who have already experienced harm as a result of their alcohol use.</p> <p>At another level, screening, brief intervention and referral for treatment is a way of identifying individuals attending for health care who have alcohol misuse, and providing them with the opportunity through screening and brief counselling of lessening their individual risk of subsequent alcohol-related harm. Note that there are many other advantages to the individual and also to the health care provider (such as the physician). These include:</p> <ul style="list-style-type: none"> • through prompt and efficient identification of alcohol misuse, clarification of the differential diagnosis in patients who have a range of clinical presentations (e.g. epigastric pain, hypertension, anxiety, depression - and including multiple nonspecific symptoms that are often seen among those attending primary health care); • reaching a diagnosis of an alcohol-related disorder more efficiently and avoiding inappropriate or unnecessary investigations that might otherwise have to be undertaken, with their attendant costs, risks and side effects; • avoidance of inappropriate, ineffective or unnecessary treatments (for example ineffective treatment of hypertension in someone whose alcohol intake unknowingly is excessive and continues to be so); • avoiding disruption to management plans for presenting disorders because of the known interactions between alcohol and various pharmaceuticals (both at the hepatic and CNS levels) and the effects of continuing alcohol consumption on medication compliance and attendance for appointments; and • avoiding the medicolegal consequences of failing to recognise an important and remediable cause of, or contributor to, multiple human disorders. <p>It is apparent from the above that screening and brief intervention can be examined not only at the population health level and at the</p>	<p>agree that KQ1 only has population level relevance whereas KQ4 only has individual relevance; nor do we think trying to highlight the issues raised here by the reviewer would be useful to improve the report.</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
		<p>individual level, but that it may be examined according to the efficiency of diagnosis and treatment for presenting disorders.</p> <p>Examination of the research questions reveals a mix of population-level and individual-level considerations. For example, Key Question 1 addresses the question of whether the provision of screening and brief intervention, with or without referral, leads to reduced morbidity, mortality or improvement in other long-term outcomes. This is primarily a population level issue. Should programs of screening and brief intervention be established as one of the primary means of reducing alcohol-related harm in the population as a whole?</p> <p>Key Question 4a addresses the question of whether behavioural interventions improve outcomes for people with alcohol misuse who have been identified by screening within primary health care. There are many studies identified which address this issue. This has population level implications, but in addition addresses the benefits to individuals. It may therefore be helpful for the authors to identify these contrasting objectives and targets of screening and brief intervention. Please note too that for population health objectives, group means are the appropriate measures whereas for individual objectives, percentages of patients who reduce their drinking to below hazardous levels and the number needed to treat (NNT) to achieve this are the most relevant.</p>	
Peer Reviewer #8	Introduction	<p>Introduction – additional comments</p> <p>On Page 1 of the Introduction (page 33 of the entire document), there are two statements I find difficult to reconcile. The first is that 50% of adults are regular drinkers. The second is that 40-50% of the population have alcohol misuse. Some explanation of this is required, more than they are just statements from different studies.</p>	<p>We have revised the 40-50% figure as it was incorrect (see comments above from Reviewer 1 and 4). It is closer to 30% for the % with alcohol misuse. This should now make sense that 50% are regular drinkers, and they drink various amounts---some healthy amounts and some unhealthy amounts. It is a subset of the 50% that has alcohol misuse.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Introduction	On Page 3, there is a comment that it is not known whether a change in hazardous drinking causes population health benefits. This may not have been apparent from the studies examined, but there is a very large literature on (i) the factors that influence alcohol intake in a population over time , (ii) the corresponding changes in average consumption, (iii) the corresponding changes in the proportion of people in the heaviest drinking categories, and (iv) changes in morbidity and mortality. Alcohol intake, expressed as a per capita measure in a population, is influenced by legislation (e.g. minimum drinking ages and the cost and availability of alcohol <i>inter alia</i>). The authors are advised to refer to the voluminous literature on these points including the seminal studies of Seeley (1960), Cartwright et al (1974) and numerous reviews such as Bruun K et al., 1974; Edwards et al., 1994; Babor et al., 2004; Chritziks et al., 2008; Saunders and Latt, 2011. In addition there are numerous cross-sectional studies comparing cost and availability of alcohol, per capita consumption, prevalence of hazardous alcohol consumption and morbidity and mortality between states and countries. These are not as relevant to the present review. Sometimes the term prevalence is used to denote life-time prevalence, but sometimes it seems to be the 12 month prevalence or point prevalence.	This is not what the sentence says. It says: "The assumption underlying brief behavioral counseling interventions in primary care is that, for identified risky drinkers, reducing overall alcohol consumption or adopting safer drinking patterns (that is, fewer drinks per occasion and not drinking before driving) will reduce the risk for medical, social, and psychological problems. ²¹ Little experimental evidence supports this assumption, and most epidemiologic evidence relates health outcomes to existing drinking behaviors rather than to changes in drinking behaviors." However, since 2 reviewers commented (see comment 22 from Reviewer #4) on/disputed this sentence (2 nd sentence of the quote here), we realize that it must be confusing. In addition, it is not necessary to include it to make the point we're going for in that paragraph. Thus, we deleted the sentence.
Peer Reviewer #1	ES / Methods	ES Methods: Does the literature allow appropriate calculation of the NNS? The question is whether one counts those screened but not enrolled in the trials? If so, then how does one account for the many reasons screened people do not end up in the trial beyond screening negative (eg screen positives who choose not to enroll/consent)? If not, how does one know how many would really be identified and receive intervention?	See response below to the related comment about NNS (comment 91 from this reviewer). We have added more description of the approach to calculating it and we have added a number of sensitivity analyses to address these questions.
Peer Reviewer #1	ES / Methods	Could criteria for quality of studies and for strength of evidence be cited/referenced/linked?	We have added references to this portion of the ES (they were already in the full report).
Peer Reviewer #1	ES / Methods	Is it PICOTS (ES)? Or PICOS (full report)?	It should be PICOTS throughout and we have fixed this.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Methods	It may be semantics but it seems to me you did include pharmacotherapy in the report. Doesn't seem right to say you decided not to include it when what really happened was you looked for studies and found none that met your criteria (pharmacotherapy for patients identified by screening in primary care). And that is a critically important result.	<p>We believe it is correct to say that we did not include pharmacotherapy in this report. We decided before reviewing the literature that we would not be including an assessment of pharmacotherapy in this report, but determined (a priori) that we would search to see if there are any available double-blind RCTs of pharmacotherapy that enrolled subjects that were identified by screening in primary care, that treated subjects in a primary care setting, or that treated subjects with risky/hazardous drinking (as opposed to those with dependence) for the purpose of introductory/ background/ contextual information.</p> <p>We explain this in the report.</p> <p>There is a separate review currently in topic nomination development on pharmacotherapy for alcohol dependence.</p>
Peer Reviewer #1	ES / Methods	Can something be inserted re the process for decision making when there are TEP disagreements? Does the Task Order Officer make a decision? Or is it the authors of the report? Someone else? It seems that should be specified so it is clear where decisions are made.	See response to comment #1 (as this comment is directly related to that issue). In the front matter to the report, we include statements to indicate that the decision making, etc was by the authors of the report: "The findings and conclusions in this document are those of the author(s), who are responsible for its content, and do not necessarily represent the views of AHRQ."

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Methods	I wonder if development of alcohol dependence should have been considered an outcome? The idea being that early identification and brief counseling would prevent its development. Also should loss of insurance be considered a potential outcome?	Interesting questions. Our team, our key informants, and our TEP did not identify these as potential outcomes of interest during the topic refinement or the CER process. We focused on the outcomes that were determined by our team, with KI and TEP input, to be most important for clinicians, patients, and policy makers. Unfortunately, we did not include this a priori, and we don't have the time and resources to go back and conduct an additional systematic review for new potential outcomes of interest raised by 1 individual at this point in the process.
Peer Reviewer #1	ES / Methods	For KQ2 randomized trials would not be the strongest evidence, rather, unbiased comparison with a reference standard would be. Perhaps this should be specified.	We agree, and we have added this point in the methods section.
Peer Reviewer #1	ES / Methods	Is the primary outcome change in drinking from baseline to 12 months (as is stated) or is it the difference in the change? In other words, isn't the relevant outcome how much MORE people in an intervention group decrease their drinking than control subjects decrease their drinking (which invariably happens)? I was unclear in the table whether the 3 drink decrease was a 3 drink greater decrease compared to control, or if it is just what is found in the intervention groups without regard to the decrease in the control group.	<p>It is the difference between the intervention and control group. The difference is 3.6 drinks greater for the intervention group compared with the control group. That is why the Table title is "Effectiveness and strength of evidence (SOE) of behavioral interventions <u>compared with controls...</u>"</p> <p>We already included the following sentence in the Methods: "For the primary outcome of alcohol consumption (drinks per week), the effect measure was the mean difference between behavioral counseling intervention and control."</p> <p>We revised another sentence in the methods to make sure this is completely clear (new text underlined): "For our meta-analyses, our primary outcome was change in alcohol consumption (drinks per week) between baseline and 12 months <u>for intervention groups compared with control groups.</u>"</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Methods	Not clear why 5 year cutoff for reviews. Ballesteros review was useful for male/female differences and Cuijpers for mortality.	<p>Those were useful in the past, but both of the articles referenced are now outdated and we did not feel were sufficiently current to reliably reflect the latest body of literature. We conducted our own meta-analyses for those outcomes/issues (male/female differences and mortality).</p> <p>We have added the explanation to the methods section of the full report: “ We limited to the last 5 years because we wanted to ensure that findings were sufficiently current; we did not need to rely on older systematic reviews and meta-analyses because we intended to conduct our own meta-analyses that would better reflect the current body of literature.”</p>
Peer Reviewer #2	Methods	Overall the methods are fine. One question: on page 12, lines 18-19 indicated that ITT results were included if available. It's unclear whether complete cases estimates were adjusted to reflect ITT results how much of the data were ITT vs. CC.	That information is not in the Methods section. It is part of the results/findings. Whether results for each individual study were ITT or something else is included in the Appendix on Quality Criteria. There is a Table with a row for each study in that Appendix and one of the columns indicates whether the study used an ITT analysis. This information was used to determine the quality rating (internal validity) for the studies.
Peer Reviewer #3	Methods	The methodology is explicit, cogent, and easily followed.	Thank you

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #6	Methods	The methods were well executed. The search strategy was explicit and logical. The inclusion and exclusion criteria are generally justifiable. It is concerning that alcohol dependent patients were included in some of the studies as brief intervention is not an acceptable or appropriate intervention for these patients who have a chronic condition requiring a chronic condition management approach (see for example, McLellan, Lewis, et al, 2000, JAMA). Although the authors included an analysis of those papers with 10% or fewer alcohol dependent patients, more should be made of the fact that brief intervention isn't appropriate for AD and perhaps the results of studies excluding such individuals (or referring them to specialty treatment only) should only be reported or reported separately and highlighted.	Thank you. It is an empiric question—whether behavioral interventions in primary care could be beneficial for those with alcohol dependence—that we wanted to evaluate by reviewing the literature. We don't agree that it should just be assumed that behavioral interventions for primary care are not appropriate for those with dependence without doing the appropriate analyses to determine if this is supported by trial evidence. We believe that we have appropriately highlighted that the literature <i>supports the theory</i> that those with alcohol dependence require other treatments (e.g., 12-step programs). However, it does not prove this theory.
Peer Reviewer #6	Methods	The report included an analysis of the comparative effectiveness by intensity of intervention, but not by other features. Particularly, should cognitive behavioral therapy interventions (which may not be appropriate for such brief interventions and are not included in the NIAAA Clinician's guide or other standard Brief Alcohol Intervention training materials by NIAAA and others) be excluded? Is it possible to compare brief advice versus interventions utilizing motivational approaches as this is an important question for the field.	It might be an interesting question-- whether analyses could be done to compare CBT interventions vs. brief advice vs. motivational approaches. But, it was not a question we set out to answer (a priori) and would require substantial additional work to categorize and analyze the studies in this manner. We determined our analysis plan a priori, with input from a number of experts, and this comparison (CBT vs. brief advice vs. motivational approaches) was not suggested. In addition, it is not clear that all of the interventions in the included studies fit into those categories so neatly; and we have time and resource limitations.
Peer Reviewer #6	Methods	Providing results stratified by subgroups of interest (young adults, older adults, pregnant women) is very useful. If possible to provide results stratified by race and/or ethnicity, such findings would be useful.	It is not possible to provide results stratified by race and/or ethnicity because most of the studies did not report information on race or ethnicity as indicated in the Tables that summarize the characteristics of included studies.
Peer Reviewer #7	Methods	In general the methods section is well written and the methods are clearly stated.	Thank you

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	Methods	<p>Within the Study Selection section, additional searches other than systematic reviews should have been utilized for KQ2 as a general check on the included screening instruments (page 9, row 53). Without doing so, the authors have omitted a major World Health Organization (WHO) effort to develop and validate a new screening tool (further described in the results section).</p>	<p>We appreciate the suggestion about adding this WHO effort (related to the ASSIST; see related comments 129 and 142). We have added this to the report. Due to limited time and resources, and because of the intended purpose of KQ2, we relied on a review of previous reviews and supplementing the information with targeted searches and recommendations of TEP and peer reviewers. We consider this addition to be part of that supplementation by peer reviewers, thank you.</p>
Peer Reviewer #7	Methods	<p>Under subgroups of interest (page 10, rows 7-11), there is no mention individuals who are “risky” multiple substance users or who have multiple substance use disorders. Multiple substance use is common (especially tobacco and alcohol use) and outcomes are often reported in screening brief intervention trials for alcohol use. Multiple substance users would seem to be an important subgroup of interest that has not been included.</p>	<p>We list “those with co-occurring mental health disorders or chronic medical conditions” among our subgroups of interest. This includes those who have multiple substance use disorders. We have revised some of the text in the applicability section of both the ES and the Discussion of the full report to specifically mention those with multiple substance use disorders. New text underlined: “It is unclear whether our findings are applicable to people with comorbid medical or psychiatric conditions, <u>including those with multiple substance use disorders</u>, and some researchers have suggested that brief behavioral interventions may be ineffective or less effective in people with comorbid psychiatric conditions.”</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	Methods	There is little explanation on page 12 under the Data Extraction and Quality Assessment sections describing how the reviewers (two independent reviewers?) were trained or periodically monitored for reliability, etc. The two, mentioned throughout the report, seem to have had responsibility for data extraction and management, assigning USPSFT criteria to selected articles, synthesizing the findings, and grading the strength of evidence. More detail on how they were trained and supervised or perhaps their names, if senior investigators, would be appreciated.	<p>We have added more detail to the Methods as suggested. Of note, “two reviewers” does not mean that this was all done by the same two people for all of these things. A lot of people were involved in the process. It just means that each of those things was not just done by 1 person.</p> <p>In the quality rating section we added: “For each article, one of the two reviewers was always an experienced/senior investigator (DJ or RH).”</p> <p>For the SOE section we added the second sentence here: “Two reviewers assessed each domain for each key outcome, and differences were resolved by consensus. For each assessment, one of the two reviewers was always an experienced/senior investigator (DJ or RH).”</p>
Peer Reviewer #1	Results	For KQ1 these outcomes are listed (Morbidity, Mortality, Health Care Utilization, Sick Days, Costs, Legal Issues, Employment Stability, and Quality of Life) but shouldn't other be listed too? (eg there were no studies found for intermediate outcomes, primary outcome of the report (drinks) either).	Yes, we did not find any studies for those other/intermediate outcomes either for this. We have added that to the results as suggested.
Peer Reviewer #1	ES / Results	For KQ2: is the 0.98 and 0.97 sensitivity for the AUDIT and AUDIT-C an estimate summarized across studies? It seems very high and may not be representative of all studies?	No, it is not a pooled estimate. It is the highest values reported. We have revised this to make sure that is clear. We also follow that sentence with more information about the range of estimates.
Peer Reviewer #1	ES / Results	Also for KQ2, the SASQ validation studies in primary care (eg Smith et al 2009 J Gen Intern Med) and elsewhere (ERs) have tended to report sensitivity and specificity for the spectrum of misuse (as have AUDIT studies) not just for risky use. So it might be misleading to report on sensitivity and specificity just for risky use when like the AUDIT they have been validated for detecting the spectrum of misuse. Also it may be worth mention that studies of SASQ did not exclude older adults.	We have revised the presentation of that information (related to the single questions and some of the other instruments). We have created a new Table that reports sensitivity and specificity of various instruments for detecting alcohol misuse (the full spectrum) in US primary care populations---now the information from the single questions is in this Table, rather than the table for risky use.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	I don't think it is a good idea to summarize the literature on CAGE in the elderly as favorable for identifying risky amounts when Adams et al (1996) studied 5,065 consecutive ambulatory patients age >60 and found the CAGE to have a sensitivity of 31% and a specificity of 92% for detecting levels of consumption greater than the recommended limits of 14 drinks per week for men and seven drinks per week for women, and it performed poorly for detecting heavy drinking too.	We have removed that statement about CAGE being favorable in the elderly for identifying risky drinking
Peer Reviewer #1	ES / Results	Results under KQ2 do not address adolescents. It seems like they should? I know there are studies of AUDIT-C for example in that age group, one of which modified it for adolescents. There is also the CRAFFT (Knight J et al). The other issue of relevance here is that for adolescents often the goal is to screen for any use since low risk levels for adolescents are not known, and counseling should likely address any use in that age group. So a screening test should detect any use, not just higher risk/heavy use.	<p>We have added a section on adolescents, highlighting that we didn't find any studies in adolescents in primary care. We also added some text about the limitations of our approach for this KQ (relying on previous systematic reviews, and supplementing with articles from TEP, peer and public review, and personal files).</p> <p>We pulled the CRAFFT article by Knight et al ("Validity of the CRAFFT Substance Abuse Screening Test Among Adolescent Clinic Patients"; <i>Arch Pediatr Adolesc Med.</i> 2002;156:607-614). The article (and perhaps the screening instrument itself?) combines all types of alcohol and drugs/substances. This doesn't provide any information about the sensitivity and specificity for detecting alcohol misuse, it provides sensitivity and specificity for detecting problem use of anything, abuse of anything, and dependence on anything. Thus, we have not included it in this report as it is not useful for our purposes.</p>
Peer Reviewer #1	ES / Results	When statements are made in text and tables like "the one included study" a reference to that would be helpful.	We have added the references for these.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	<p>I am not sure if studies of pregnant women were not included because they were not in primary care? But I am not sure if that concept/limit applies to pregnant women who seek obstetric care for their pregnancies which is de facto their primary care. There are several studies beyond the one mentioned (one with a mortality outcome): Chang G <i>Obstet Gynecol</i> 2005;105:991-8; O'Connor and Whaley <i>Am J Public Health</i> 2007;97:252-8; Handmaker NS et al. <i>J Stud Alcohol</i> 1999;60:285-7; Jankin JR <i>Alcohol Res Health</i> 2002;26:58-65; Floyd RL et al <i>Am J Prev Med</i> 2007;32:1-10; Ingersoll K et al <i>Pediatrics</i> 2003;111:1131-35.</p>	<p>We agree that we would consider such studies primary care. We did not exclude any studies of pregnant women seeing their OB/Gyn for that reason. We have added additional text to the discussion to address the other existing studies of pregnant women and why they did not meet inclusion criteria. We also added some information about the findings of those studies.</p> <p>We added the following: "Our searches identified other studies focusing on pregnant women that did not meet our inclusion criteria (list of references). Several did not take place in a primary care setting, but instead were conducted in other settings, such as those that included jails and specialized drug and alcohol treatment centers; these included, for example, the Project CHOICES study.{Floyd, 2007, #932} Others were excluded because they did not include a control group or because they followed participants after the intervention for less than 6 months.{Handmaker, 1999, #1147; Chang, 2005, #473} Several of these studies reported benefits of behavioral interventions for pregnant women, including reduction of alcohol consumption,{Handmaker, 1999, #1147; Chang, 2005, #473} reduced risk of an alcohol-exposed pregnancy,{Floyd, 2007, #932} higher rates of abstinence,{O'Connor, 2007, #2102} and better fetal and newborn outcomes (birth weights and birth lengths, and fetal mortality rates).{O'Connor, 2007, #2102}</p>
Peer Reviewer #1	ES / Results	Table ES-4. I assume the drinks are means. Please note this.	Yes, they are means. We added this to the table.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	Please clarify: “Very brief (up to 5 minutes) and brief (more than 5, and up to 15 minutes) single-contact interventions appear to be ineffective or less effective.” Should there be a “respectively” at the end of the sentence? In other words are very brief interventions ineffective? That is important to know/report.	<p>We have revised this in the executive summary to clarify: “Our meta-analyses of studies in adults found very brief (up to 5 minutes) and brief single-contact interventions to be ineffective for some outcomes and less effective than brief multicontact interventions for others.”</p> <p>We have also added a more detailed description as a footnote (labeled Intensity of Intervention) of the relevant Table (Summary of effectiveness and strength of evidence of behavioral interventions compared with controls for improving intermediate outcomes, by population) in the executive summary and in the detailed results section of KQ4a to explain further. Briefly, from our meta-analyses in adults, very brief interventions (1 study) did not have statistically significant benefit for drinks/week, 0 studies available for heavy drinking episodes, and they were effective (but less) for achieving recommended drinking levels. The very brief contributions to the meta-analyses were just 1 study for each of those outcomes (but not the same 1 study—one was Richmond and 1 was the WHO study). Brief interventions were less effective for drinks/wk and for achieving recommended levels, but did not reach statistically significant difference for evidence of benefit for heavy episodes.</p>
Peer Reviewer #1	ES / Results	I would suggest using the term “Recommended drinking limits” (or “lower risk drinking limits”) rather than “recommended drinking levels” because the recommendation made during brief intervention is not a recommendation to drink, it is a recommendation to stay under certain limits (or not drink at all). I think this is important because there is often confusion in clinical practice between recommendations to drink moderate amounts to prevent cardiovascular disease, and advice to not exceed amounts and risk adverse health consequences.	We have made this change as suggested.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	Would liver enzymes be included in alcohol-related liver problems? Should they be separated out as outcomes in part because they are more likely to be reported in these trials?	No, liver enzyme levels are not included in alcohol-related liver abnormalities. Liver enzyme abnormalities can indicate a number of different things and don't necessarily reflect an alcohol-related problem. They were not identified as an outcome of interest because they are intermediate outcomes that were not determined to be the outcomes of greatest interest for this report.
Peer Reviewer #1	ES / Results	Was the Kristenson et al series of reports on RCT of heavy drinking men included in this review? Original report 1983 I believe but then more recent in Alcohol Clin Exp Res finding decreases in sick days, utilization and alcohol-related mortality.	No these studies were not included in the report as they did not use any of the acceptable/included screening approaches (which we defined very broadly to include any of the screening instruments we list or any questions about quantity or frequency of alcohol use) to identify subjects. Of note, these studies are listed in our Appendix (B) of excluded studies.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	I think Primary Care Visits decreasing would best be viewed as evidence of a possible harm of screening rather than an outcome for Table ES-6.	Interesting perspective, and we could speculate what that would mean, but it's not terribly relevant since there was no significant difference for primary care visits (WMD, -0.14 visits, 95% CI, -0.5 to 0.2); we did not find a decrease in PC visits as this comment suggests (that would be an overinterpretation of the point estimate without considering the CI here). Even if they did decrease, we believe it is uncertain whether this would indicate harm or benefit. It depends on the baseline number of visits for the individuals and the type of primary care visits that were decreased. For example, they might need fewer visits for blood pressure management if their drinking was reduced to the recommended limits or they might have fewer visits related to other alcohol-related morbidity that would be treated in primary care (which would all be good things). Alternatively, it could be evidence of possible harm if they stopped going to PC visits due to stigma or avoidance of further discussion of alcohol use. I there was a finding of fewer visits (which there is not), this would need to be explored further to determine the reasons.
Peer Reviewer #1	ES / Results	<p>Could ES-6 include cost per QALY (Maciosek et al) instead of a cost benefit ratio?</p> <p>In that ratio is the \$205 cost the cost for 48 months of health care (as it should be) or was it just the cost of the brief intervention? Maybe just a bit more clarification about the meaning of the costs and benefits would help in a table footnote.</p>	<p>We did not include cost per QALY as an outcome of interest as evaluation of cost-effectiveness was not an aim of this report.</p> <p>We have added a footnote to clarify what the \$205 represents and edited the text as well to indicate that this was the intervention cost. The footnote: "The \$205 per patient cost includes \$166 borne by the clinics per patient and \$39 borne by patients (for lost work time and travel costs)."</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	<p>For the cost data from Project TrEAT it is worth noting how many episodes of health consequences were counted in those costs. In other words, it is my understanding that the large cost savings occurred because of a very small number of motor vehicle related injuries that were severe and costly. If one compared the number across groups one wonders if they would be significant even though those few events led to very big cost differences (are the cost differences by randomized group actually statistically significant using appropriate statistics for highly skewed data?). My guess is that since accidents probably did not differ significantly by group, it would not be appropriate to conclude that costs related to those accidents did differ. Then again I see in a subsequent table for young adults there was a significant difference in MVCs. Maybe it would be useful to present those data for adults too.</p>	<p>We include this level of detail in the evidence tables of the full report, but not in the ES due to space constraints and to avoid obscuring the main points. There is a very detailed listing of all sorts of health consequences in the analysis in that paper. In addition, we have separate areas of the report that address those health outcomes specifically and don't repeat them here.</p> <p>We graded the strength of evidence for the cost data LOW for a number of reasons, including the type of thing mentioned by the reviewer here.</p>
Peer Reviewer #1	ES / Results	<p>When summarizing is there a way to go beyond "brief, up to 15 minutes"? It would be particularly useful to the field to know if brief or very brief advice as, more or less effective than brief motivational counseling (empathic, stage based or seeking change talk by reflective listening) because the latter requires substantial training and is much harder to disseminate (similarly, knowing how much training was done for those intervening in studies finding effectiveness would be a key point that would be useful for readers).</p>	<p>We have made revisions to address this. We now indicate in several places (including the results of the abstract and executive summary) that these brief multi-contact interventions were "generally 10-15" minutes per contact. We clarify in the more detailed sections that 4 of the 7 brief multicontact interventions used 10-15 minute interventions (TrEAT, Healthy Moms, Noknoy, and Rubio); 1 probably used 10-15 minute interventions (Wallace—not reported in article, but email to author explained that they trained them to do "up to 15 minutes" and that he expects people did 10-15 minutes); 1 did 5-10 minutes (Project Health); and 1 study used a shorter PCP intervention (Curry had a 1-5 minute intervention from the PCP) and then up to 3 longer phone calls from a graduate-level clinical psychology student (avg 14 minutes per call).</p> <p>Interestingly, there were no studies that used a very brief multi-contact intervention, with 5 min or less per contact.</p> <p>Thus, it is not something we can determine with confidence---whether a very brief (5 min or less per contact) multi-contact intervention would be sufficient. We include this in our</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
			<p>Future Research section of the full report: “We found no studies evaluating a very brief (each contact 5 minutes or less) multi-contact intervention and it is unknown whether very brief multi-contact interventions would be as effective as the brief multi-contact interventions identified in this report (generally 10-15 minutes per contact). Knowing the minimum amount of time needed for an intervention to be effective is very important for busy primary care practices, where a positive screen triggering a brief intervention could mean taking up the entire time allotted for the visit to discuss alcohol misuse, and postponing the original purpose of the visit. Future studies could possibly compare the intervention delivered in Project TrEAT (two 15 minute visits with the primary care physician and follow-up calls by a nurse) that provides some of the best available long-term evidence for the effectiveness of behavioral interventions with a shorter version of the same intervention (using 5 minute or less interventions)..”</p> <p>This is related to the comment 52 and the response to that comment above.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	<p>KQ7 in Figure ES-1 mentions Veterans, Sex and Co-occurring conditions. Are these then explicitly addressed in the results? (also, the methods does not have as extensive a list of subgroups as the Figure does).</p> <p>Should Kaner EFS et al systematic review and meta-analysis be cited in this context (Kaner EFS. Ment Health Subst Use. 2011;4(1):38–61)?</p>	<p>Yes, they are explicitly addressed in the results or discussion. We include in the report that we found no studies in Veterans; we describe our subgroup analyses by sex in the results; and we describe our limited ability to make conclusions about co-occurring conditions. (and we made sure the methods, in the study eligibility criteria table, and figure have a consistent list of subgroups).</p> <p>Our searches found the review mentioned here (“Systematic review of the impact of brief interventions on substance use and co-morbid physical and mental health conditions”) and we reviewed it. It did not meet our inclusion criteria because it lumps all substance abuse and does not aim to analyze for alcohol misuse separately. We did hand search their list of references though to make sure we identified any studies focused on alcohol that they found, so that we could evaluate those ourselves.</p>
Peer Reviewer #1	ES / Results	<p>Regarding key question #7, I wonder if it is possible to separate studies that assured screened patients of research confidentiality (e.g. result not placed in medical record) vs those in which a clinical level of confidentiality was used (e.g. result placed in medical record)? This might have an effect on the validity of screening, or effectiveness of intervention.</p>	<p>We did not set out to answer this question about whether studies doing or not doing this would have an effect. It is beyond the scope of this project.</p>
Peer Reviewer #1	ES / Results	<p>On a related note, is it worth mentioning UPPL laws (laws that allow insurers to not pay for health care of the visit is determined to be related to alcohol use)? I don't know if these could apply to primary care settings (they are usually discussed in the context of trauma or emergency care). Might alcohol misuse in the medical records increase the chance a visit wouldn't be paid for or life or disability insurance be denied? I understand there may be no data on this.</p>	<p>This is an interesting question, but we were not able to find anything that would suggest that alcohol misuse in the primary care medical records would increase the chance that a visit wouldn't be paid for or insurance denied. We can only find that they have been discussed in the context of trauma or emergency centers, as the reviewer expected.</p> <p>In the case of patient's presenting in Trauma Centers (TC) and Emergency Departments (ED) and being identified as having an injury which was related to alcohol intoxication of illicit drug use, the recording of this information</p>

Commentator & Affiliation	Section	Comment	Response
			<p>in the patient's medical record can impact their ability to receive reimbursement for medical care and in some states with the amount of disability insurance they receive.</p> <p>In 1947 the National Association of Insurance Commissioners (NAIC) adopted the Uniform Accident and Sickness Policy Provision Law (UPPL) as a model law. The law states that health insurers would not have to reimburse patients for costs incurred when an accident is a result of "the insured's being intoxicated or under the influence of any narcotic." Although the same organization reversed its position and recommended that states repeal the law in 2001, several states have taken no action on the law and as a consequence, patients presenting with injuries they sustained as the result of being intoxicated or under the influence of any narcotic still remain and could have a detrimental impact on patients and EDs and TCs.</p> <p>Chezem (2005) reports that only a portion of ED and trauma patients actually are screened for alcohol use and alcohol-related problems. She reports on several factors that serve as barriers to screening including the fact health care providers may fear that because of existing laws, insurers may deny reimbursement for medical services if a patient has a positive blood alcohol level at the time of the ED visit. Some observers have identified the legal provisions that deal with alcohol use and the insurance payment of benefits for medical care as a factor that may contribute to the failure of many medical care facilities, particularly EDs, to screen for alcohol abuse and dependence as well as other alcohol-related problems. Schemer et al. 2003 conducted a survey of trauma surgeons and found that twenty-seven percent of respondents thought that screening would</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
			<p>threaten the reimbursement of medical costs. Gentilello et al. (2005) reporting on the same survey indicated that twenty four percent of the trauma surgeons indicated that they had encountered an alcohol- or drug-related insurance denial in the past 6 months.</p> <p>Cherpitel, CJ. (2006): Alcohol-related injury and the emergency department: research and policy questions for the next decade. <i>Addiction</i>, 101: 1225-1227.</p> <p>Chezem L. (2004/2005): Legal barriers to alcohol screening in Emergency Departments and Trauma Centers. <i>Alcohol Res Health</i> 20042005; 28: 73–7.</p> <p>Gentilello LM, Donato A, Nolan S, Mackin RE, Liebich F, Hoyt DB, LaBrie RA. (2005): Effect of the Uniform Accident and Sickness Policy Provision Law on alcohol screening and intervention in trauma centers. <i>Journal of Trauma</i> Sept. 59 (3): 624-31</p> <p>Gentilello LM., Rivara R , Donovan DM, Jurkovich GJ, Daranciang E, Dunn CW et al. (1999): Alcohol interventions in a trauma center as a means of reducing the risk of injury recurrence. <i>Ann Surg</i> 230: 473–80.</p> <p>Shermer, CR, Gentilello, L, Hoyt, DB, et al (2003): National survey of trauma surgeons' use of alcohol screening and brief intervention. <i>Journal of Trauma: Injury, Infection, and Critical Care</i> 55: 849–856,</p>
Peer Reviewer #1	ES / Results	Regarding potential adverse effects, would it be useful to examine the lower bounds of confidence intervals (for results that were not statistically significant) to understand the range of potential benefits and harms? (in other words, if a study result for change in heavy drinking days ranged (95% confidence interval) from -3 to +2, the potential harm would be an increase in heavy drinking days plausible within a 95% CI).	We do not believe that such an approach has validity for estimating potential harms for the evidence in this report. That would really be reaching.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	<p>Also regarding adverse effects, I believe one tends to use lower quality data when looking for adverse effects (eg not intent to treat but instead, treatment received) in the medication study literature. That literature often includes large case series or secondary analyses of trials. My sense is that there is a fair amount of peer-reviewed literature on the topic. Here are some examples that might inform the question:</p> <ol style="list-style-type: none"> 1. Beich A, Gannik D, Malterud K. Screening and brief intervention for excessive alcohol use: qualitative interview study of the experiences of general practitioners. <i>BMJ</i> 2002;325:870 doi:10.1136/bmj.325.7369.870 (Published 19 October 2002) 2. Saitz R, Horton NJ, Cheng DM, Samet JH. 2008. Alcohol counseling reflects higher quality of primary care. <i>J Gen Intern Med</i>; 23: 1482-1486. 3. Johnson M, Jackson R, Guillaume L, Meier P, Goyder E. Barriers and facilitators to implementing screening and brief intervention for alcohol misuse: a systematic review of qualitative evidence. <i>J Public Health (Oxf)</i>. 2011 Sep;33(3):412-21. Epub 2010 Dec 17. PMID: 21169370 	<p>We have added the text below to the ES discussion and the full report discussion to mention some of this literature, noting that our review focused on trials (finding very little data), but that other types of studies may shed some light on potential adverse effects.</p> <p>“While trial data are limited regarding adverse effects of screening and behavioral interventions for alcohol misuse in primary care settings, other types of studies may offer some insights. Among a group of 24 general practitioners in Denmark who were interviewed about their participation in a screening and brief intervention program for alcohol misuse, nearly all reported experiencing negative reactions from some patients.{Beich} Such reactions ranged from feelings of uneasiness or embarrassment to finding another physician. The physicians themselves noted that the added work of screening and brief intervention was onerous and hampered the establishment of rapport with patients. They also expressed concerns that screening identified people for whom intervention was not necessary yet took valuable time and resources while at the same time failing to detect and help some for whom alcohol misuse was a real problem. However, other studies have found that patients view screening favorably, even perceiving higher quality of care when screening is followed by counseling.{Johnson} For example, one prospective cohort study found that communication and whole-person knowledge were perceived as better among patients who were counseled about their alcohol misuse compared to those who were not counseled.{Saitz}”</p>
Peer Reviewer #1	ES / Results	<p>Last para of results: It would be very helpful for translation to practice to know how much and what kind of training was done in these studies of effective brief interventions, even if just brief mention of range of hours, and type of content/teaching methods</p>	<p>We have added additional detail here about the range of hours/training. Very little detail was reported by many of the studies. We added: “When reported, training duration</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
		<p>(because a major question in the field is about how and whether these brief interventions can be disseminated widely).</p>	<p>ranged from as little as 15 minutes to as long as 6 to 8 hours, full day workshops, or a four week training in motivational interviewing principles. Nine studies reported trainings of research staff and interventionists that were thirty minutes or longer and also provided feedback, booster sessions, or weekly conference calls to maintain adherence to protocol. Five others reported trainings of thirty minutes or greater but did not provide information on booster sessions.” (References available in the full report)</p> <p>We also added additional training details in KQ7 of the Results (see comment 134 below) as follows: “Of the 23 RCTs we included in this report, 16 included at least some mention of training. Provider and/or staff trainings were reported in most studies. When reported, training duration ranged from as little as 15 minutes to as long as 6 to 8 hours, full day workshops, or a four week training in motivational interviewing principles. Nine studies reported trainings of research staff and interventionists that were thirty minutes and longer and also provided feedback, booster sessions, or weekly conference calls to maintain adherence to protocol. Five others reported trainings of thirty minutes or greater but did not provide information on booster sessions. One RCT reported that counselors completed a four week training in motivational interviewing. The type of training received was often described fairly briefly, possibly due to space limitations. For example, in Project TrEAT, physicians “were trained to administer the intervention protocol through role playing and general skills techniques in educational programs...also received additional training in booster sessions that occurred at least twice during the trial.” Some studies provided much greater detail. For example, in Project Health,</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
			<p>“training generally occurred in 2 sessions...a 2-hour small-group session and a 10 to 20 minute individual tutorial session 2 to 6 weeks after the group session. In addition, at the beginning of the recruitment period research assistants generally gave a brief (1-2 minute) refresher orientation to providers about their use of the intervention tools (i.e. goal statement, tip sheets) just before a study patient was seen. In total, providers received about 2.5 to 3 hours of training.”</p>
Peer Reviewer #1	ES / Results	<p>Also, were interventionists anything other than primary care clinicians and research personnel? If not, this is worth stating because the more common federally supported model for alcohol screening and brief intervention involves neither of those personnel types (they involve health educators whose sole job is brief intervention). Also see systematic review by Sullivan L, Tetrault J, Fiellin D 2011 in I believe Am J Addictions re provider type and brief intervention efficacy in primary care.</p>	<p>Yes, some studies had interventionists that were not primary care physicians or research personnel. We have an entire section of the results (see KQ7 section on “Personnel Involved with the Study”, over 250 words) that addresses this issue and describes the exact interventionist in each study (whether it was a PCP, nurse, PA, research personnel, etc.). It includes:</p> <p>“Fourteen of the interventions were delivered by a primary care physician alone or in conjunction with a health educator or nurse. Three were delivered by a nurse or physician’s assistant; one was conducted by a psychologist, two by a researcher; and one by unspecified interventionists.⁷³ Two interventions were provided via a computer...”</p> <p>We didn’t include all of that detail in the Executive Summary, but we did include: “Interventions delivered by primary care providers and by research personnel were both effective for reducing alcohol consumption, with data showing a trend toward greater reduction for interventions delivered mostly by primary care providers (WMD, -4.0 drinks per week, 95% CI, -5.4 to -2.6) than for those delivered primarily by research personnel (WMD, -3.0, 95% CI, -5.0 to -1.0).” ...and now we have added to the ES (with references) “ Just one intervention delivered by a nurse contributed to the drinks</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
			<p>per week meta-analysis; the reduction in drinks per week was not statistically significant for that study (WMD, -0.2, 95% CI, -8.9 to 8.6). Two other studies that did not provide sufficient data for our drinks per week meta-analysis reported benefits of interventions delivered primarily by nurses, or by nurses and physician assistants for some consumption outcomes. In addition, two interventions conducted via computer reported some evidence of effectiveness for college students.”</p>
Peer Reviewer #1	ES / Results	<p>Are there really no RCTs of adolescent alcohol screening and intervention suitable for inclusion (studies of Knight J?).</p>	<p>No, unfortunately there really are not any. There is some literature in adolescents that did not identify subjects by screening in a primary care setting (studies conducted in schools for example).</p> <p>We identified several studies with Knight J as the first author that did not meet our criteria as they generally did not include an intervention (they often compared screening approaches for operating characteristics).</p>
Peer Reviewer #1	ES / Results	<p>Figure 2: it is unclear what role PRISMA plays in this figure.</p>	<p>We deleted this abbreviation from the footnote to the Figure. It was in there because an earlier draft of the Figure included that abbreviation (PRISMA) within it (but it is unnecessary and we've deleted it).</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Results	Should this systematic review have been included? It is particularly important since it comes to a different conclusion than the current manuscript, and this paper did look at papers most likely to have reported mortality data in a high quality fashion. Also, should mortality be all-cause or alcohol-related for this current review? Cuijpers P, Riper H, Lemmers L. The effects on mortality of brief interventions for problem drinking: a meta-analysis. <i>Addiction</i> 2004;99(7):839-45.	<p>The Cuijpers meta-analysis is 7 years old now. We set a 5 year cutoff for systematic reviews and meta-analyses to ensure that they would be sufficiently up to date. We have added information in the discussion to describe the findings of the Cuijpers meta-analysis and the differences between our analyses.</p> <p>In the Cuijpers 2004 meta-analysis, they included 4 studies: Fleming et al, 1999 (Project GOAL), Fleming et al, 2002 (Project TrEAT), Wutzke et al, 2002 (WHO BI study), Chick et al, 1985.</p> <p>Our meta-analysis included the first 2 above. We excluded Chick et al because the study enrolled patients in hospital wards and was not conducted in a primary care setting by enrolling those identified by screening in primary care. We included Wutzke 2002 in a sensitivity analysis only, but not in the main analysis. We did so because of the setting (not conducted in a primary care setting and enrolling subjects identified by screening in primary care, and thus did not meet inclusion criteria). Even with the addition of Wutzke, our meta-analysis did not reach statistical significance, but it did trend in that direction (which we describe in the report).</p> <p>Further, there are 4 additional studies that we included in our meta-analysis for all adults that they did not include (Wallace et al 1998) or that were newer since their analysis (Noknoy 2010, SIP/Bischof 2008, Kypri 2004). Thus, it is not surprising that the results/conclusions would be somewhat different.</p> <p>The Cuijpers analysis yielded a 0.47 RR of mortality (95% CI: 0.25, 0.89).</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Results	p. ES7, ll. 13-40: “Sensitivity and specificity” of what? Many old validations only measured the capacity to identify dependence (or abuse and dependence) rather than those disorders plus hazardous and harmful use. Are you certain all the instruments you name are “good” in this regard with the full range of behaviors? I fail to understand why the CAGE, TWEAK, and T-ACE are even considered, as they are not intended to identify the full range of alcohol misuse. The findings are pretty clear—they don’t work well. So why include them? Or did the studies you reviewed use them?	We have revised this section and we clarify what the sensitivity and specificity pertain to. We focus on the AUDIT, AUDIT-C, and single questions as those best suited for screening in primary care for the full spectrum of alcohol misuse. We only mention the CAGE here now to highlight that it is not a good screening test for risky drinking (only good to screen for abuse/dependence); which we feel is important to highlight for practitioners. We no longer mention the TWEAK and the T-ACE in this Executive Summary b/c we agree that they are designed to focus on abuse/dependence, and addressing them here may detract from the main messages.
Peer Reviewer #4	ES / Results	Ibid., l. 51ff.: Again, providing some discussion of what is meant by “intermediate outcomes” seems to be required, as well as a justification of why certain measures were selected	We have added this as suggested.
Peer Reviewer #4	ES / Results	p. ES8, ll. 35-36: There is potentially a huge difference between “ineffective” and “less effective”. If a statement is to be made about such interventions, it would seem essential to differentiate and explain.	See response to similar comment 52 from reviewer 1 above. We have clarified this with additional detail.
Peer Reviewer #4	ES / Results	p. ES12, ll. 43-44: Please explain how the fact that providing interventions in clinical trials required support systems leads to the statement that “Such supports are likely to be required for effective screening and brief intervention”. Is there some other evidence of this or compelling logic? And in what time frame? Remember, please, that it takes “support systems” for clinicians learning to take blood pressure at first, but after they have done it a sufficient number of times, many clinicians become quite proficient at it. Do you know that over time clinicians will NEVER become proficient at SBI without continual support?	We stand by our assessment and our statement as it is worded. The current system (without supports in most practices) has resulted in many practices not offering screening or intervention. We disagree with the implications of the comment. We believe that this comment suggests a lack of understanding of how typical primary care practices operate. This is not a view that any of the clinicians on our team or our TEP share. Nor is it one that any other reviewer expressed. In addition, when this information was presented to the USPSTF, there were several individuals that recognized the importance and relevance of support systems for screening and intervention.

Commentator & Affiliation	Section	Comment	Response
Kathy Bradley	ES / Results	<p>The KQ2 review noted widely varying sensitivities for the AUDIT and low specificity for the AUDIT-C. These resulted from errors in one of the reviews used (Fiellin 2000), the use of outdated reviews with omission of recent important studies (Tables 1-2). The review for KQ2 relied on previously published reviews that had serious limitations.</p> <p>Errors in previous reviews (Fiellin 2000).</p> <p>1. The most important error was the misreporting of the AUDIT cut-point used in a key study (Steinbauer Ann Intern Med 1998). The review reported that an AUDIT threshold ≥ 8 was used when a much lower threshold was actually used (≥ 5). This error has already been propagated in another review (Reinert & Allen 2002).</p> <p>2. The Fiellin 2000 review also reported the sensitivity and specificity for the AUDIT-C for men at a cut-point (≥ 3) that is only recommended for women. Therefore, the screen appeared very sensitive and non-specific in the review, although it had lower sensitivity and higher specificity at the recommended threshold of ≥ 4 (Table 1).</p>	<p>Thank you for pointing out this error in the previous review. It is true that the original article (in footnote to their Table 3) reports using a cutoff of 5. We have reviewed the original Steinbauer article from Annals of Internal Medicine (Title: Ethnic and Sex Bias in Primary Care Screening Tests for Alcohol Use Disorders). It reported sensitivity of 70%-92% and specificity of 73%-94% for alcohol use disorders (abuse and dependence), using a cutoff of 5 or more for the AUDIT. However, this did not actually end up resulting in an error in our report because we did not include the values from this study in our range (because we knew that it used a cutoff of 5, rather than 8, which is the cutoff that we report in our Table). We reported a range from 0.61 to 0.96 for sensitivity and from 0.85 to 0.96 for specificity for the AUDIT (for abuse and dependence, cutoff 8).</p> <p>Thank you for pointing this out. We have made sure to clarify the cutpoints used for the data we report. We have also added a row to the Table for a cut-point of 4, showing the slightly lower sensitivity and higher specificity. We have also added a new Table to the Report that focuses on the sensitivity and specificity of AUDIT, AUDIT C, and single questions (the best instruments for which such data is available and the ones designed for this purpose) for identifying alcohol misuse (the full spectrum), which helps to highlight the data for the full spectrum.</p>

Commentator & Affiliation	Section	Comment	Response
Kathy Bradley	ES / Results	The reviews used for KQ2 are quite outdated (Fiellin 2000; Bradley 1998); and limited (Berner 2007 included only data on the AUDIT using a criterion standard that did not include alcohol use disorders whereas Berks (2008) focused only on the elderly and Burns (2010) only on pregnant women.	<p>We acknowledge the limitations of our approach to KQ2 in the report. Of note, we relied on the systematic reviews, but supplemented information in the KQ with articles suggested by our TEP, and peer and public reviewers, which helped to improve it and fill in some gaps. We are very grateful to this particular reviewer for making several very helpful suggestions to supplement the KQ.</p> <p>We include in the limitations, for example: “For Key Question 2 (“How do specific screening modalities compare with one another for detecting alcohol misuse?”), we did not review all individual publications assessing screening instruments. Instead, we relied on previously published systematic reviews to find information on their sensitivity and specificity and filled gaps with data from other sources.”</p>
Kathy Bradley	ES / Results	Due to the errors above, the appropriate screening thresholds are not reported in Table 5. Screening thresholds reported for the different screening questionnaires were not those that balance sensitivity and specificity, and were not comparable screening thresholds. Therefore, sensitivity and specificity for very insensitive screens (high specificity) are combined with data for very sensitive screens (low specificity). Based on adequately sized US studies with a criterion standard of risky drinking and/or AUD, the appropriate cut-points for balancing sensitivity and specificity for the AUDIT are ≥ 4 or ≥ 5 . For the AUDIT-C the appropriate cut-points for balancing sensitivity and specificity are ≥ 3 for women and ≥ 4 for men.	<p>We have added rows to the Table to show data for cut-points of 4 and 5 for AUDIT and for 3 and 4 for AUDIT-C for detecting risky drinking; we also include these cut-points in our new table showing sensitivity and specificity for detecting alcohol misuse (the full spectrum).</p> <p>We have also added the point (in relation to our new table that shows sensitivity and specificity for screening for the full spectrum) about the appropriate cut-points as mentioned here.</p>
Kathy Bradley	ES / Results	Important studies are missing. It appears no review was conducted of recent studies, which mostly excluded important studies on single item alcohol screening questionnaires (Table 2).	Thank you, we have added information from the suggested studies. Also, see related comments above about approach to identifying studies for KQ2.

Commentator & Affiliation	Section	Comment	Response
Kathy Bradley	ES / Results	Recent studies focused on identification of alcohol dependence in particular were not included. Given the goal of the review (in the legend of Table 2 on page 11), this was surprising. If the goal was to evaluate screens that distinguished patients with alcohol dependence who the review shows appear less likely to respond to brief alcohol counseling, 2 recent studies were not included (Vinson, 2007; Rubinsky, 2010).	We have revised the wording of the legend of Table 2 to clarify. The focus of the review is on screening to detect alcohol misuse (the full spectrum), and we have revised KQ2 to better address this (see responses to other comments related to KQ2). The legend now says: "For KQ2, like the previous review for the USPSTF, we assessed screening approaches using the included systematic reviews. We supplemented the findings with information from other sources to fill important gaps. We utilized TEP members and peer reviewers to help supplement findings by recommending sources."
Kathy Bradley	ES / Results	Important differences across validation samples and criterion standards were not reflected in the analyses. Table 5 (pg 19) combines data for men and women and for different criterion standards altogether. Given cultural differences in drinking patterns and drinking norms, validation studies from the US should be used. Four large US studies (3 settings: Texas, Georgia and Seattle VAs) have included appropriate detailed interview-based criterion standards for alcohol misuse (drinking above recommended limits and/or DSM-III/IV alcohol use disorders) and adequate numbers of patients to have adequate precision for sensitivity. Most of these studies found screening thresholds of 8 on the AUDIT were so insensitive that they were not even reported (Table 1).	Thank you for the very helpful comments. As described in the responses to previous comments related to KQ2, we have added a new table that addresses this issue. It summarizes the findings of the studies mentioned here. We have also added text that highlights the issue related to cut-points (e.g., that the AUDIT cut-points should perhaps be 4 or 5 and the AUDIT-C 3 for women and 4 for men), and the point about the importance of validation studies from US primary care sites.
Kathy Bradley	ES / Results	[See tables in her comment document]	We have reviewed these Tables (and the comments related to them) and incorporated the relevant studies into our report.
Peer Reviewer #1	Results	I am not sure that the tools should be described as commonly used. They may be more frequently published but 1) none are commonly used since most patients are never screened, and 2) I am not aware of data on how often the specific tools are used currently in primary care practice (though there are tools promoted and used by SAMHSA programs but those programs are not in the majority of US primary care practices). Also I have not heard of anyone using the MAST or shARPS for screening in practice (or the RAPS for that matter). So I suspect they are not commonly used.	We agree and we have deleted "commonly used". We now say "Screening tools used to identify alcohol misuse include but are not limited to..."

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	In the summary of findings for KQ2 would it be worth mentioning that some of the tools can be asked by interview, some by self-administration or interview, and some are really not feasible without either electronic forms or other aids (like the AUDIT-C or ASSIST that require selecting across multiple responses and scoring).	We have added this point to the section on Instrument Burden: "Some of the screening instruments can be asked by interview, some by self-administration or interview, and some are not feasible without either electronic forms or other aids (e.g., the AUDIT or ASSIST that require selecting across multiple responses and scoring)."
Peer Reviewer #1	Results	A serious limitation of the CAGE as a basis for brief intervention is that the CAGE identifies LIFETIME abuse/dependence. In Samet JH et al. Am J Med one finds that most people who are CAGE positive in primary care are not currently using alcohol. So it becomes extremely inefficient to use such a tool to identify people for brief intervention. Probably not a good idea to say therefore that the CAGE is a good test. The SASQ and AUDIT at least identify current (past year) misuse.	We have reworded our description of findings related to the CAGE to focus on the point that it has "very low sensitivity for detecting risky/hazardous drinking and is therefore not a good screening test for identifying risky/hazardous drinking." In addition, we have incorporated the point about most people who are CAGE positive in primary care---The article reports that most patients in whom alcohol abuse is detected in primary care using the CAGE questionnaire are either actively addressing their substance abuse or are in recovery. Samet JH, O'Connor PG. Alcohol abusers in primary care: readiness to change behavior. Am J Med 1998 Oct; 105(4):302-6?
Peer Reviewer #1	Results	A Table seems to be missing for KQ2. The most relevant question is the sensitivity and specificity for the full spectrum of alcohol misuse (risky, problem, abuse, dependence). Yet the only tables provided are sens and spec for risky/hazardous/harmful, and then for abuse/dependence. But neither of these latter are relevant for screening for alcohol misuse (full spectrum) to then do brief counseling intervention.	We have added the suggested table---one for the full spectrum---see responses to similar comments related to KQ2.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	Similarly, I am pretty sure there have been studies that look at the utility of screening questionnaires like the AUDIT and AUDIT-C (see Rubinsky A et al in Drug Alcohol Depend in the past year or two) that provide the predictive value of a score or range of scores for dependence. This is particularly important to note (if a screening tool can do this) since BI works (according to your review) for hazardous use but not dependence, so PCPs need to be able to easily make this distinction.	<p>We have added information from Rubinsky to the text to explain the probability of alcohol dependence based on scores from the AUDIT, AUDIT-C, AUDIT-3 (the 3rd question of the AUDIT), and a single question about the number of days drinking ≥ 5 drinks over the past month: "Some studies have reported the probability of alcohol dependence based on scores from screening instruments. From a family medicine clinic population including 392 men and 927 women with mean ages of 46 and 42 years, respectively, the AUDIT was found to have a post-screening probability of alcohol dependence of 87% for men for scores from 15-40 and 94% for women for scores from 13-40; the AUDIT-C was found to have a post-screening probability of alcohol dependence of 75% for men and 88% for women for scores from 10-12; AUDIT-3 (the 3rd question of the AUDIT, asking the frequency of drinking ≥ 6 drinks) was found to have a post-screening probability of alcohol dependence of 58% for men for scores from 3-4 and 88% for women for scores of 4; and a single question about the number of days drinking ≥ 5 drinks over the past month was found to have a post-screening probability of alcohol dependence of 83% for responses from 14-30 and 38% for women for responses from 3-30. The probability of alcohol dependence was much lower for lower scores.</p> <p>#2463 Rubinsky AD, Kivlahan DR, Volk RJ, et al. Estimating risk of alcohol dependence using alcohol screening scores. Drug Alcohol Depend. 2010 Apr 1;108(1-2):29-36. PMID: 20042299; excluded for wrong study design</p>
Peer Reviewer #1	Results	The summary for KQ2 reports very high sensitivity for AUDIT and AUDIT-C yet the tables (6 and 7) that summarize the evidence show much lower sensitivity. It is not clear how the summary relates to the actual evidence in the Tables.	We have revised the summary to better match the Tables and to focus on the information in the new table (on screening for the full spectrum)

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	Sensitivity of the SASQ in ref 72 was 82% (81.8%) not 81%.	Thank you, we have corrected this.
Peer Reviewer #1	Results	Can a recommendation be made about appropriate use of AUDIT and AUDIT-C cutoffs for screening in primary care? How about any need to change the number of drinks in item 3 of these questionnaires to adapt to US drink sizes?	It is not our intention to make recommendations in this report, but with all of the new revisions we have made sure to point out the sensitivity and specificity data for screening for the full spectrum in US primary care studies for various cut-points---this includes data showing better sensitivity/specificity for cutoffs that are lower than the traditionally recommended cutoffs. We also now include in the text that when focusing on adequately sized US studies that reported sensitivity and specificity of screening for the full spectrum of alcohol misuse in primary care, data suggest that the often recommended cut-points for screening (i.e., AUDIT \geq 8) may need to be revised. ... suggests lowering the cutoff score to \geq 4 or \geq 5. For the AUDIT-C, the appropriate cut-points for balancing sensitivity and specificity may need to be lowered to \geq 3 for women and \geq 4 for men.
Peer Reviewer #1	Results	Are there data to support statements that AUDIT appears to be the most widely used? Where? SASQ was 87% sensitive and 67% specific for abuse/dependence in ref 72. Why is this summarized as "not good" in the graf above Table 6? (and followed by a cryptic sentence re specificity being excellent in adults??).	We have revised this to clarify: "Among the trials included in this report, the AUDIT appears to be the most widely used screening instrument." We agree that this was not accurate (for the SASQ). We have revised the text related to the single question to describe the findings (and no longer describe it as "not good" for abuse/dependence. We now simply describe the data (stating that "The range of sensitivity reported for single question screens was from 0.77 to 0.88 for detecting alcohol abuse or dependence, depending on whether the past 3 or 12 months was considered.") We removed the cryptic sentence as well.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Results	Should the ASSIST be mentioned in the KQ2 summary? Arent there studies in primary care?	<p>We have added information about ASSIST to KQ2. The two main validation publications include a mixture of subjects from alcohol and drug treatment facilities and subjects from primary care. For one of them, 60% are from alcohol and drug treatment facilities and the other 40% were from general medical settings and from psychiatric facilities; for the other one, around a third were recruited from drug treatment. We explain in greater detail in the text. The ASSIST differs from the other screening instruments covered in this review because it was developed to screen for all psychoactive substances, rather than just for alcohol.</p> <p>See comment 142 below and our response for related details and for the citations for ASSIST that we've added to the report.</p>
Peer Reviewer #1	Results	Hasnt the AUDIT-C been studied via oral administration/computer in primary care settings in the VA? (Table 7)	Yes, it can be administered orally or by computer---we have updated the Table to reflect this.
Peer Reviewer #1	Results	KQ5 Could the \$39 cost calculated in 1997 be made more meaningful by updating it to 2012 dollars?	<p>We have added that this is 1997 dollars and added that it is worth approximately \$53 in 2011-12. We used the conversion factor available from http://oregonstate.edu/cia/polisci/download-conversion-factors</p> <p>The conversion factor is 0.736 to go from 1997 dollars to present dollars.</p>
Peer Reviewer #1	Results	KQ6: could this section clarify whether it is referring to health care costs or all costs from a societal perspective? (this is clear in one sentence but not in many places)	We have added this to the results section on "Costs" to clarify what the 2 studies reported. The overall numbers are from a societal perspective.
Peer Reviewer #1	Results	From Project Treat, the nonsignificant differences in emergency department utilization, nonsignificant differences in motor vehicle crashes, and the odd significant difference in hospital utilization (where one controls increase and intervention stays flat—from the original papers) raises questions about whether the cost differences are significant/believable or due to outliers.	We agree and we have taken this type of logic into consideration in our grading of the strength of evidence; it is one of several things that contributed to the LOW strength of evidence grade.
Peer Reviewer #1	Results	KQ7: A bit more detail would be very useful here—regarding the	We have added additional training details in

Commentator & Affiliation	Section	Comment	Response
		<p>amount of training required. Please provide the amount of training required for the most effective brief interventions (eg brief multi) and even for comparison, for the brief or very brief. This information is critical for determining if these effective practices are translatable to the real world.</p>	<p>KQ7 of the Results (quoted here below). We have included the range across studies, with references, as well as some more specific details about Project Health and Project TrEAT (the two brief multicontact studies that provide the longest followup evidence)</p> <p>“Of the 23 RCTs we included in this report, 16 included at least some mention of training. Provider and/or staff trainings were reported in most studies. When reported, training duration ranged from as little as 15 minutes{ref} to as long as 6 to 8 hours,{ref} full day workshops,{ref} or a four week training in motivational interviewing principles.{ref} Nine studies^{41-43,45,48-50,52,55,56,83,86,88-94} reported trainings of research staff and interventionists that were thirty minutes and longer and also provided feedback, booster sessions, or weekly conference calls to maintain adherence to protocol. Five others^{81,84,87,96,97,99,100} reported trainings of thirty minutes or greater but did not provide information on booster sessions. One RCT^{55,56,86} reported that counselors completed a four week training in motivational interviewing.</p> <p>The type of training received was often described fairly briefly, possibly due to space limitations. For example, in Project TrEAT, physicians “were trained to administer the intervention protocol through role playing and general skills techniques in educational programs...also received additional training in booster sessions that occurred at least twice during the trial.” Some studies provided much greater detail. For example, in Project Health, “training generally occurred in 2 sessions...a 2-hour small-group session and a 10 to 20 minute individual tutorial session 2 to 6 weeks after the group session. In addition, at the beginning of the recruitment period research assistants generally gave a brief (1-2 minute)</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
			refresher orientation to providers about their use of the intervention tools (ie goal statement, tip sheets) just before a study patient was seen. In total, providers received about 2.5 to 3 hours of training.” ”
Peer Reviewer #1	Results	Also, related to that, I think it would be important for readers, and important to the USPSTF who may make use of this document, to have the facts that in the best evidence papers, screening took >30 mins and was done by research staff, and that substantial training is required in studies in which interventions were found to be effective. The reason this is important, is...see the most recent USPSFT statement on Depression screening that notes that it is recommended if staff and supports are in place.	We agree that this is important and we address it in the discussion: “All interventions required support systems to provide screening and screening-related assessment, and in some cases, provider prompting. Screenings to identify subjects for the included studies were often extensive, multi-step processes that included face-to-face interviews lasting up to 30 minutes by research personnel. Less time would be required for screening and screening-related assessments in primary care practice; we estimate less than 2 minutes for negative screens and 5-10 minutes for positive screens, with most of the time for screening-related assessment to determine whether the patient has an alcohol use disorder as opposed to risky/hazardous drinking.”
Peer Reviewer #2	Results	Overall the results section is well done and the answers to the above questions are yes. With regard to interpreting the results for older adults it would be helpful to have information on how their baseline drinks per week compared to adults.	Thank you. We have added the information on baseline drinks to footnotes for the Table that summarizes the outcomes in both the ES and Results (“Table. Summary of effectiveness and strength of evidence of behavioral interventions compared with controls for improving intermediate outcomes, by population”).
Peer Reviewer #3	Results	Page 61, Line 20: Grammatical error: Text should read, "It reported a statistically significant difference in hospital days in the last 6 months for the intervention group *compared with* the control group at 6, 12, and 48 months..." Otherwise, flawless recounting of the relevant literature.	We fixed the grammatical error. Thank you
Peer Reviewer #6	Results	The detail provided is appropriate and studies are generally well described. Tables and figures and appendices are adequate and descriptive.	Thank you

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #6	Results	The well known "Cut-Back" study by Babor, Higgins-Biddle and others, (Alcohol and Alcoholism, 2006) should be considered for inclusion.	This study (refid #6417) was identified in our literature search, was reviewed, and was rated poor quality for the following reasons: High risk of selection bias and confounding due to attrition; 65% or more of those eligible for 3 month follow up did not complete 3-month follow up; less than half of those intended to be sampled for 3-month follow up completed 3-month follow up; also, the study was randomized by clinic/practice, but analyzed at the individual level; unable to determine if groups were similar at baseline for important potential confounders; unable to separate results for zone 1 vs zone 2 vs zone 3; by design, only some of those enrolled were actually contacted for follow-up. This information is included in the appendix that describes the quality ratings.
Peer Reviewer #7	Results	The results are clearly presented by question and key messages are concisely stated.	Thank you
		Figures, tables and appendices are adequate and descriptive; however, there are missing data points throughout on a number of tables (for example, Table 10, page 31 includes 2 components that are "not reported"). Since most studies reviewed are recent and there are relatively few included, it seems reasonable that the study authors could be contacted to supply missing data for the report.	We have emailed the authors to fill in the missing data points mentioned here. For example, we now reference personal communication (email) with Wallace about the duration of the intervention—they trained providers to intervene for "up to 15 minutes" and the authors believe that the interventions were generally 10-15 minutes in duration, per contact.
Peer Reviewer #7	Results	The results of the sub-analyses of some of the sub-groups of interest (identified in rows 7-10 on page 10) seem to be missing (i.e., racial/ethnic minorities, people with co-occurring mental health disorders or chronic medical conditions, veterans, etc.) ?	There are none, because of insufficient data. This information is in the results in the section that describes the characteristics of studies and applicability. It is also highlighted in the discussion of the report.
Peer Reviewer #8	Results	Key Question 1. I note that the authors did not identify any studies directly addressing this question. I understand that a study undertaken by Drummond, Kaner and their colleagues in the United Kingdom has examined this.	Our searches identified several publications by Drummond or Kaner. None of them address KQ1.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Results	<p>Key Question 2.</p> <p>Some screening instruments (e.g. the AUDIT) can also offer a brief assessment and facilitate diagnosis, and also provide a platform for intervention, given the high face validity of the questions.</p> <p>On Page 19 (page 51), are there clues as to the reason for the wide range of values for sensitivity for the different screening instruments? Were these studies all undertaken in primary care – so that they would represent the person’s current alcohol use or were some undertaken in settings where the person might have last consumed alcohol some weeks before? The extreme range of values does not pass the “test of commonsense”.</p>	<p>We have added text that explains that some of the instruments can facilitate diagnosis (based on the post-screening probability of alcohol dependence). We have re-written KQ2 as described in previous responses.</p> <p>We have limited our review to studies in primary care or primary care-like settings, but differences in populations, settings, use patterns, and demographic characteristics are likely related to the wide ranges for some of the values. In our new table focused on screening for the full spectrum of alcohol misuse, the ranges are much tighter, likely because the table is more focused (limited to US primary care studies). The issue of what an appropriate cutoff is may also be related to the wide range for some of these.</p>
Peer Reviewer #8	Results	<p>Please could the authors specify whether the sensitivity and specificity analyses in Table 5 are based on a comparison of at-risk + hazardous + harmful drinking versus non-hazardous use, with subjects with alcohol abuse and dependence excluded.</p> <p>Likewise, did they exclude subjects with at-risk + hazardous + harmful drinking from the analyses presented in Table 6?</p> <p>It is relevant that the authors identify a cut-off point of 5 as providing the best combination of sensitivity and specificity for the AUDIT. Is this based on an ROC analysis? Would the authors like to make a recommendation that the cut-off point of the AUDIT is changed to 5?</p>	<p>We have clarified the basis of the analyses for those Tables as suggested. The addition of a table related to the full spectrum is the most useful thing to understanding the underlying spirit of these questions though.</p> <p>See previous comments and responses related to the cut-off point of the AUDIT. We have added text to address this (but don’t make a formal recommendation)</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Results	<p>Key Question 3.</p> <p>I note that there are no clearly identified adverse effects of screening. I am surprised there are not qualitative data on the response of people who have experienced screening for alcohol misuse. Beich and his colleagues working in Denmark have stated that physicians there are concerned about the intrusiveness of screening among their primary care populations. There are also philosophical issues raised about screening in primary health care such as: Who owns the consultation – the patient or the physician?</p>	<p>See similar comment above (comment #64) from reviewer #1.</p> <p>We have added the text below to the ES discussion and the full report discussion to mention some of this literature, noting that our review focused on trials (finding very little data), but that other types of studies may shed some light on potential adverse effects.</p> <p>“While trial data are limited regarding adverse effects of screening and behavioral interventions for alcohol misuse in primary care settings, other types of studies may offer some insights. Among a group of 24 general practitioners in Denmark who were interviewed about their participation in a screening and brief intervention program for alcohol misuse, nearly all reported experiencing negative reactions from some patients.{Beich} Such reactions ranged from feelings of uneasiness or embarrassment to finding another physician. The physicians themselves noted that the added work of screening and brief intervention was onerous and hampered the establishment of rapport with patients. They also expressed concerns that screening identified people for whom intervention was not necessary yet took valuable time and resources while at the same time failing to detect and help some for whom alcohol misuse was a real problem. However, other studies have found that patients view screening favorably, even perceiving higher quality of care when screening is followed by counseling.{Johnson} For example, one prospective cohort study found that communication and whole-person knowledge were perceived as better among patients who were counseled about their alcohol misuse compared to those who were not counseled.{Saitz}”</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Results	Key Question 4a. There is a risk of confusing the analyses and the message with the findings of KQ 4b.	Thank you, we appreciate this risk. We have tried to word the findings in such a way that the findings are not confused.
Peer Reviewer #8	Results	Key Question 4b. Is it possible to test a dose-response relationship by allocating a numerical value to the intensity of treatment. This could of course include a value allocated to the control condition. As the authors say, the control condition in many studies was a minimalist intervention in which patients received advice and/or written material about alcohol. The authors may wish to contact Dr. Tom Babor at the University of Connecticut who examined equally intense but non-content-relevant in surveys for alcohol misuse.	We don't believe this would be a good approach and we believe it might be attempting to achieve a level of precision that the literature for this question does not provide/allow for without making several leaps/assumptions.
Peer Reviewer #8	Results	Key Question 5. Was there any evidence of reduced effect in any of the more extended interventions compared with briefer ones?	No, and that is described in KQ4a because the outcomes are intermediate outcomes of benefit (e.g. ,reduction in consumption). We explain those findings in that section, and in the Discussion (in relation to the intensity of the intervention). We explain that, if anything, it's the other way around---that the very brief interventions have less evidence of benefit. We're not sure what this question implies in relation to KQ5---we assume the reviewer is implying that reduced effect in extended interventions would have implied harms/adverse effects (if we had found that). That is somewhat of a leap to assume that less effect is equivalent to harm (and we didn't find what he is suggesting anyway).
Peer Reviewer #8	Results	Key Question 6. In reading Page 53 (page 85 of the whole document), I was surprised that the authors' meta-analyses did not find a reduction in all-cause mortality for adults. This is in contrast to the findings of the meta-analysis conducted by the World Health Organization in 2009-10 as part of the global Mental Health Gap Action Program (mhGAP), which found a significant effect on all-cause mortality. I accept that metaanalyses may generate different results but it is puzzling (to say the least) that two almost contemporaneous meta-analyses come to two different conclusions. Did the authors identify the following paper, which examined the	Comment 69 above from Reviewer 1 is somewhat related to this comment. We have added information in the discussion to describe the findings of the Cuijpers meta-analysis and the differences between our analyses. They (the WHO) don't appear to have conducted their own mortality meta-analysis; rather, they cited data from another publication (Cuijpers et al, 2004). The meta-analysis they used is 7 years old now. In the

Commentator & Affiliation	Section	Comment	Response
		<p>long-term outcomes of the WHO collaborative brief intervention study? Wutzke S.E., Conigrave K.M., Saunders J.B. and Hall W.D. The long-term effectiveness of brief interventions for unsafe alcohol consumption: a 10-year follow-up. <i>Addiction</i> 2002; 97:665-675.</p> <p>As mortality is such an important end-point and given the fact that the authors concluded that evidence was insufficient to draw conclusions on morbidity, I would suggest that the authors contact WHO and try to identify the reasons for this difference. Communication would properly be via the Director of the Department of Mental Health and Substance Abuse, Dr Shekhar Saxena at the WHO Headquarters in Geneva.</p>	<p>Cuijpers 2004 meta-analysis, they included 4 studies: Fleming et al, 1999 (Project GOAL), Fleming et al, 2002 (Project TrEAT), Wutzke et al, 2002 (WHO BI study), Chick et al, 1985.</p> <p>Our meta-analysis included the first 2 above. We excluded Chick et al because the study enrolled patients in hospital wards and was not conducted in a primary care setting by enrolling those identified by screening in primary care. We included Wutzke 2002 in a sensitivity analysis only, but not in the main analysis. We did so because of the setting (not conducted in a primary care setting and enrolling subjects identified by screening in primary care, and thus did not meet inclusion criteria). Even with the addition of Wutzke, the meta-analysis did not reach statistical significance, but it did trend in that direction (which we describe in the report).</p> <p>Further, there are 4 additional studies that we included in our meta-analysis for all adults that they missed (Wallace et al 1998) or that were newer since their analysis (Noknoy 2010, SIP/Bischof 2008, Kypri 2004). Thus, it is not surprising that the results/conclusions would be somewhat different.</p> <p>The Cuijpers analysis yielded a 0.47 RR of mortality (95% CI: 0.25, 0.89). The Wutzke study was weighted very heavily (50%) and had the narrowest CI.</p> <p>The evidence used by WHO: http://www.who.int/mental_health/mhgap/evidence/alcohol/mh_evidence_prof_alcohol_q1_screening_2010_en.pdf</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Results	<p>With regard to morbidity in pregnant women, do the authors conclude that there were no reports of pregnancy or birth outcomes amongst the literature examined? I suggest that if this is the case, a statement is made specifically to this effect. These are more relevant outcomes during pregnancy than considering alcohol-related liver problems or other longer-term morbidities. It would also be worth making a comment about the absence of any data on prevention of fetal alcohol spectrum disorder if this is indeed the case.</p>	<p>See related comment #50 from reviewer #1 and the response to that comment.</p> <p>We have added text to the Exec summary discussion (and full report discussion) that includes mentioning the 1 study we're aware of that reported fetal and newborn outcomes: "Our searches identified other studies focusing on pregnant women that did not meet our inclusion criteria.{references} Several did not take place in a primary care setting, but instead were conducted in other settings, such as those that included jails and specialized drug and alcohol treatment centers; these included, for example, the Project CHOICES study.{Floyd, 2007, #932} Others were excluded because they did not include a control group or because they followed participants after the intervention for less than 6 months.{Handmaker, 1999, #1147; Chang, 2005, #473} Several of these studies reported benefits of behavioral interventions for pregnant women, including reduction of alcohol consumption,{Handmaker, 1999, #1147; Chang, 2005, #473} reduced risk of an alcohol-exposed pregnancy,{Floyd, 2007, #932} higher rates of abstinence,{O'Connor, 2007, #2102} and better fetal and newborn outcomes (birth weights and birth lengths, and fetal mortality rates).{O'Connor, 2007, #2102}</p>
Peer Reviewer #8	Results	<p>Key Question 7.</p> <p>This is a highly relevant question and one that has not been addressed in most previous systematic reviews. There is however a body of literature that has been omitted from the analyses. For example, I found no reference to the work of the World Health Organisation Phase III or Phase IV Multicentre Collaborative Studies, which examined specifically these issues. These included randomised controlled trials and other studies adopting experimental design.</p> <p>Some relevant papers are the following publications (admittedly, the first in a WHO report, peer-reviewed but not appearing in its entirety in a scientific journal).</p>	<p>We didn't include most of these because we were not addressing questions about the dissemination and implementation of interventions, and these studies did not meet our inclusion/exclusion criteria (for example, we didn't include cost-effectiveness analyses). We have added text to clarify that this report does not address the dissemination and implementation literature and that it may shed further light on health care system influences that promote or hinder effective screening and interventions for alcohol misuse. Our KQ7 is indeed confined to examining RCTs primarily</p>

<http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1134>

Published Online: July 10, 2012

Commentator & Affiliation	Section	Comment	Response
		<p>Saunders, J. & Wutzke, SE (eds). <i>World Health Organization Phase III Collaborative Study on implementing and supporting intervention strategies in primary health care. Report on Strand I: General Practitioners current practices and perceptions of preventive medicine and intervention for hazardous alcohol use; a 16- country study</i>. Copenhagen: World Health Organization, Alcohol, Drugs and Tobacco Programme (ref: EUR/ICP/LVNG 02 05 02), 2004 (235 pages).</p> <p>Gomel, M. K., Wutzke, S. E., Hardcastle, D. M., Lapsley, H. & Reznik, R. B. "Costeffectiveness of strategies to market and train primary health care physicians in brief intervention techniques for hazardous alcohol use", <i>Social Science and Medicine</i> (1998), 47, 203-211.</p> <p>Funk M, Wutzke S, Kaner E, Anderson P, Pas L, McCormick R, Gual A, Barfod S, Saunders JB. A multi-country controlled trial of strategies to promote dissemination and implementation of brief alcohol intervention in primary health care: Findings of a World Health Organization Collaborative Study. <i>Journal of Studies on Alcohol</i> 2005; 66: 379-388.</p> <p>Because of this omission I doubt whether the approach adopted by the authors has fairly ascertained and examined the literature on health service interventions. It seems that the analyses have been confined to examining randomised controlled trials primarily examining the efficacy or effectiveness of screening and brief intervention.</p> <p>There is an entirely separate scientific literature on implementation studies, which really needs to be assessed before definitive statements are made on health service factors facilitating or hindering brief interventions.</p> <p>Given my concern that the literature examined for this particular key question is not representative, I would suggest that Key Question 7 is deleted from the report. It is certainly an important question but needs to include an additional major body of literature. The review and analyses of the combined literature could be published as a supplementary report in order not to delay publication of the findings of the other key questions.</p>	<p>examining the efficacy or effectiveness of screening and brief intervention, as described in the methods section. We have added text to KQ7 and to the limitations section to call this to the reader's attention.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Discussion	Note that “harmful” is almost identical to “abuse”. The distinction made between harmful and abuse in the first paragraph of the discussion should not be made. I think the authors really mean to lump risky/hazardous/problem together and then the alcohol use disorders/diagnoses harmful/abuse and then dependence. Which is why it would be useful to add “problem” to the table in the intro, to avoid this confusion. What is uncertain is whether the trials apply to abuse/harmful.	We agree and have fixed this (no longer lumping harmful with risky/hazardous, and treating it as most similar to abuse). See responses to previous and later comments from this reviewer that address this (comments 2, 27, 33, 89, 95)
Peer Reviewer #1	ES / Discussion	Worth noting that not only did studies exclude patients with alcohol dependence but they very often also excluded people who drank very heavily (various definitions) meaning that the effectiveness results apply to those who have risky/hazardous use who do not drink very heavily (and do not have dependence).	This seems to be introducing yet another definition of a new subgroup (people who drank very heavily). It is not clear to us how this subgroup fits in with the defined terminology or how this is a useful subgroup to evaluate, when considering practical implications (beyond already considering harmful drinkers, abuse, and dependence). Some studies did and some did not exclude those who drank very heavily. We reviewed all of the studies and just 5/16 RCTs in adults used a quantity/frequency cutoff to exclude those who drank very heavily (e.g., excluded those who drink >50 drinks per week). Another 4 excluded those with AUDIT scores above a certain threshold (ranged from 13 to 21), which might exclude most of those who drank heavily. But, 7 of the 16 did not have an exclusion criteria that would likely exclude those who drank very heavily. Thus, we have not added the suggested point about those who drank very heavily.
Peer Reviewer #1	ES / Discussion	The “Screening for alcohol misuse” paragraph in the Discussion seems verbatim copied from the results re KQ2 and therefore my earlier comments about that para also apply here.	We have made all edits to both sections.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Discussion	<p>ES-9 is a very important Table. I now see that NNS is really a most optimistic scenario—that screening identifies all with misuse, and that all identified with misuse get a brief intervention and that all brief interventions are of the quality of those delivered in the trial. I think some text pointing this out would be useful because since the absolute risk reductions are relatively small, decrements in these assumptions could have major impact on NNS in particular (see Beich A et al BMJ 2003 systematic review and meta-analysis).</p> <p>The other thing to point out in the table is that it may not be accurate reading down the column. For example, no data suggest that if the prevalence of risky use is 4% then the prevalence of dependence would be 50% (2%)...</p>	<p>We have added a footnote to the Table to explain this: “The scenario in this table is optimistic, as it assumes that screening identifies all those with alcohol misuse (100% sensitive) and that all those identified with misuse potentially get an intervention. We conducted sensitivity analyses to explore how NNT and NNS would change using other assumptions. The NNT does not change much using a variety of different assumptions; it ranges from 6.7 to 18.2. Using a sensitivity of 81% for the screening instrument (representative of the single question) changes the NNS range to 39 to 281 (from 31 to 227). If only half of all those with a positive screening test receive an intervention, then the NNS ranges increases to 63 to 455. If 90% of those with a positive screen receive an intervention, the NNS ranges increases to 35 to 253. If the screening instrument sensitivity is 81% and only half of those with a positive screen receive an intervention, then the NNS range increases to 155 to 1122.”</p> <p>We added a footnote to address the last point about reading down the column to clarify that the prevalence of risky use and dependence in the 2 columns provide the range of estimates and are not linked to each other here.</p>
Peer Reviewer #1	ES / Discussion	<p>Again as in the results, in the discussion please note if very brief are ineffective or just less effective. Also when less effective, how much less (4%? 11% vs 15%)? A I drawing the correct conclusion from what is presented in the paper? Please clarify.</p>	<p>See response to comment 52 from this reviewer. We have clarified this sentence in all places it appeared in the document and added additional detail.</p>
Peer Reviewer #1	ES / Discussion	<p>The authors are to be commended for how they handled reporting evidence regarding dependence. It is clear, addresses an important issue, and yet given what is available in the literature is handled in a way that does not distract from the focus of the report where the evidence is (the preventive service re counseling for nondependent misuse).</p>	<p>Thank you</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Discussion	The last paragraph under “treatments for alcohol dependence” may require some editing. I think it is true that no studies exist where people identified with dependence by screening in primary care had pharmacotherapy tested. But O’Malley Arch Intern Med. 2003;163:1695-1704 did test the efficacy of naltrexone in primary care. Also, it is largely true that pharmacotherapy has not been tested for hazardous use or abuse but Kranzler et al included such individuals in his study: Journal of Clinical Psychopharmacology: June 2003 - Volume 23 - Issue 3 - pp 294-304.	We have looked at those publications and we made some revision to that paragraph to clarify, it now states: “Studies of pharmacotherapy for patients with alcohol dependence have generally enrolled subjects responding to advertisements or those being treated in specialty alcohol treatment centers. We were unable to identify any double-blind randomized controlled trials of pharmacotherapy that identified subjects by screening in a primary care setting or that assessed the efficacy or comparative effectiveness of pharmacotherapy in a primary care setting. Further, we were unable to identify any studies of pharmacotherapy for people with risky/hazardous drinking.”
Peer Reviewer #1	ES / Discussion	Applicability: Again, it isn’t clear these studies are applicable to abuse (as stated by the authors) which means it isn’t clear they are applicable to those with the ICD-10 diagnosis harmful use. The results may apply to those with risky/hazardous use, and maybe even to those with use and some consequences that do not meet criteria for harmful/abuse/dependence (often referred to as “problem use”) but not to those with dependence, and uncertain for those with very heavy drinking, abuse, harmful use or dependence. Third para under Applicability again excludes “harmful” from disorder. This is really not controversial---see Hasin D Alcohol Health Res 2003 at NIAAA website who has written on diagnostic classifications. It is also confirmed in WHO and ICD-10 publications (as well as my Saitz R 2005 NEJM paper that includes references to these definitions). (also note misuse of “harmful” in the Conclusions section)	See response to comment 2 above. We agree with the comment and we have revised the document to reflect that harmful use is most similar to abuse and we no longer lump it with risky/hazardous use in our conclusions about applicability or elsewhere (except when describing previous publications that did lump those things together for their population of interest). We have clarified that applicability to abuse/harmful use isn’t clear from the literature.
Peer Reviewer #1	ES / Discussion	In the Applicability paragraph referral to specialty treatment is mentioned. It is worth adding there that the review did not identify studies that demonstrated the effectiveness, however, of referral (or other treatments) for people identified by screening as having dependence.	We have added this point to the future research section. We’ve also included related information in several other sections of the report
Peer Reviewer #1	ES / Discussion	I am surprised there were no studies among veterans. Probably you are correct but it just seems surprising so I might suggest contacting Katharine Bradley or Daniel Kivlahan at the VA to see if they know of any that might meet your criteria.	Thank you, we have discussed this with experts in the field (on our team and with TEP) and have done additional targeted searches to look again; we found none meeting our criteria.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Discussion	<p>Regarding frequency of screening I agree there isn't much out there. This may be useful to you; it is an abstract published in J Gen Intern Med 2009: Alford DP, Almeida AB, Saitz R, Brolin M, Kim TW, Shanahan CW, Botticelli M, Samet JH. Should adults who screen negative for unhealthy substance use be rescreened annually? National SGIM Meeting, May 13-16, 2009, Miami, FL. (oral presentation) J Gen Intern Med 2009; 24: (S1) S169 AND Presented at AMERSA, November 5-7, 2009, Bethesda, MD.</p>	<p>Thank you, this is an interesting abstract. It is the abstract for an oral presentation from SGIM. But, we don't think it is sufficient evidence to determine the recommended frequency of screening. The fact that the abstract only has follow up data for 1,014 out of the 34,412 that had an initial negative screen (and these are certainly not representative) means we really don't know the answer.</p> <p>In addition, (1) it's not clear that it's limited to primary care, (2) there is no information about rescreening those with a positive screen, and (3) they were rescreened "at least 1 year after their initial screen" but it isn't clear when the rescreening was done.</p> <p>From the abstract: From March 2007 through December 2008, 41,302 adult patients were screened for unhealthy substance use in general health care settings; 6,890 (17%) screened positive, and 34,412 (83%) screened negative. Of the patients who screened negative, 1,014 reappeared to medical care at least 1 year after their initial screen and were rescreened. Of those rescreened, 34 (3.4%) screened positive for unhealthy substance use. Of patients positive at rescreening, 59% (20/34) reported unhealthy alcohol use only, 32% (11/34) reported unhealthy drug (prescription or illicit) use only, and 9% (3/34) reported both unhealthy alcohol and drug use. In a logistic regression model, patients under 60 years of age (aOR 35.6 [CI 4.8–263.1]) and men (aOR 2.3 [CI 1.1–4.7]) were more likely to report incident unhealthy use. CONCLUSION: Annual rescreening of patients for unhealthy substance use who initially screened negative identified a modest number of incident cases. Men and younger adults were at higher risk.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Discussion	Limitations: I would add that although self-report can be accurate and there is little alternative, it remains a concern that social desirability bias could play a role in those results—because although self-report is from both randomized groups in these studies, clearly the group that gets more attention and advice to decrease their drinking may be more likely to report they decreased their drinking...	We agree and we have added this to the limitations: “Nevertheless, it remains a concern that social desirability bias could play a role in the results of the included studies (i.e., although self-report is from both randomized groups in these studies, the group that gets more attention and advice to decrease their drinking may be more likely to report that they decreased their drinking).”
Peer Reviewer #1	ES / Discussion	I realize this may be repetitive but the report is repetitive...The first sentence of the Discussion says the report identified 4 categories of misuse. But harmful and abuse are almost identical and are diagnoses. Risky/hazardous is another category and not a diagnosis; dependence is another category and is a diagnosis.	We have revised the full document to reflect this. See responses to previous similar comments.
Peer Reviewer #1	ES / Discussion	The discussion talks about motivational interventions but on review of the studies included in the review it becomes clear that many of the interventions tested were not motivational interviewing. Some were just feedback and advice. This is important because the latter will be much more likely to be able to be disseminated in practice.	We have revised the discussion to clarify that such interventions can be advice, feedback, motivational interviewing, or can use other cognitive behavioral strategies.
Peer Reviewer #1	ES / Discussion	For future research might I suggest that since the report identified no studies of pharmacotherapy for dependence among people identified by screening in primary care that such be recommended. Similarly, studies of referral to treatment (do they go, does it work when they get there) of people identified by screening in primary care would be important. The reason I think these studies should be recommended is that screening identifies people with dependence and there appears to be no evidence re what to do with those people that would be effective.	We have added both of these suggestions to the Future Research section: "...we found no double-blind randomized controlled trials of pharmacotherapy for alcohol dependence that identified subjects by screening in a primary care setting or that assessed the efficacy or comparative effectiveness of pharmacotherapy in a primary care setting. Future studies could fill this void in the literature." And "...We found insufficient evidence to draw conclusions about the impact of screening and behavioral interventions on follow-up with referrals. Future studies could assess referral to treatment for alcohol dependence for people identified by screening in primary care, evaluating whether they follow-up with referrals and if it works when they get there."

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	ES / Discussion	Note that the recommendation to do a screening vs no screening study is not easily translatable to a study because it would likely mean not collecting baseline data and difficulty defining the control group. It could be done but not easily and review groups would likely not rate it highly. I do think it needs to be done though...It could be done as a cluster RCT of health centers and perhaps that should be recommended.	We agree with the comment and we have added that it could perhaps be done as a cluster RCT of practices/health centers.
Peer Reviewer #4	ES / Discussion	p. ES17, ll. 25-26: So what is the evidence that age (adolescence vs. adulthood) is equivalent to prior military service in predicting the efficacy of a medical service? I doubt that you found any studies of left-handed adults, but you did not conclude that the results have uncertain applicability for them. Why not? How about people who played basketball as children? Just how are veterans medically different from other adults?	We do not suggest that age is equivalent to prior military service for predicting the efficacy. The sentence related to this comments is in the Applicability section and states: "We did not identify any studies in adolescent populations or any conducted exclusively among veterans, and the results thus have uncertain applicability to these populations." The comment suggests that Veterans are not an important subgroup to search for evidence about the efficacy of screening and behavioral interventions. It also suggests that there is no reason to think that efficacy would differ for Veterans compared with non-Veterans. We disagree with this implication. Our team of researchers and our KIs and TEP agreed with the importance of searching for evidence in Veterans to help determine whether findings were applicable to them. Some of the theory is based on Veterans having higher rates of psychiatric comorbidities (e.g., depression, PTSD) that could contribute to decreased effectiveness of behavioral counseling interventions for alcohol misuse.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Discussion	Ibid., ll. 31-32: This implies that the results may be different with people with HIV, because you did not find studies of that population. Did you find studies limited to people with Actinic Keratoses? Would you conclude that the results would be different for them?	<p>We have revised this sentence in the Applicability section to clarify it. It previously stated: “ Although we searched for studies conducted in settings with primary care-like relationships (e.g., infectious disease clinics for people with HIV), we did not find any, and our results have uncertain applicability to such settings/populations.”</p> <p>We see how this could be misinterpreted to focus on those with HIV, but the focus was intended to be on non-traditional primary care settings that have a primary care-like relationship. In other words, this would be an infectious disease specialist providing primary care rather than a GP or family physician or internists. We have revised the sentence to clarify it (by deleting populations from the end of the sentence to make it clear that the focus was on the setting and not on a person with HIV; and by adding “non-traditional primary care settings such as...” to the parenthetical statement.</p>
Peer Reviewer #4	ES / Discussion	Ibid., ll. 42-45: See above on support systems. Can you suggest here a contrasting clinical preventive service that does NOT require or benefit from training clinicians to provide it? Or would you, rather, feel more accurate in prefacing these statements by something like, “As with other clinical preventive services, . . .”	See response to comment 73 above from this reviewer.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Discussion	p. ES18, ll. 3-4: I seem to recall your citing some data about how long the effects of interventions last. Would this be a clue?	We don't believe that information is very useful for determining how often screening should be done. Additional information would be needed to make a valid conclusion. The sentence related to this comment said "We did not find any evidence that would inform decisions about the appropriate frequency of screening (i.e., whether it should be done annually, every five years, or something else)." The data we have about how long interventions last is essentially from 2 studies that followed subjects out to 4 years (Project TrEAT and Project Health). Both of those found intervention subjects to maintain or continue to accrue additional benefit compared to their baseline, but the control group had late improvement so that differences were no longer statistically significant between the two groups. This does not help us to determine whether people should be re-screened at a particular time.
Peer Reviewer #4	ES / Discussion	p. ES14, l. 5: See above on "alcohol-related liver problems".	See response to comment 10

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	ES / Discussion	p. ES16, Table 10: What is the logic of mixing the various categories mentioned here together. You have 3 types of therapy, 2 loci and methods of providing therapy (outpatient and residential), a non-therapy using pharmaceuticals, and a service provided to safely withdraw patients from intoxication (which usually involves pharmaceuticals). These are not comparable services or categories. More precision in describing what is offered to patients with dependence is required of an academic paper than is provided by America's treatment industry.	<p>We don't see a problem in laying out the different forms of treatment for alcohol dependence in this way--this is informational/contextual to provide the reader with a sense of the available options that have been reportedly effective. A detailed comparative effectiveness review of the treatments for alcohol dependence is beyond the scope of this review.</p> <p>We have edited the text in this section to make it clear: "Although we did not systematically examine the efficacy/effectiveness of various treatments for alcohol dependence (AD) (e.g., pharmacotherapy, 12-step programs, and specialized outpatient treatment programs), we provide contextual information regarding such treatments in this section. Because screening for alcohol misuse will inevitably identify some individuals with AD, providers and those making recommendations need some information about whether there are effective interventions available for such individuals. However, a detailed review and comparison of treatments for alcohol dependence is beyond the scope of this review."</p>
Peer Reviewer #4	ES / Discussion	Ibid., ll. 47-57: You might well cite here any evidence you found about referral to treatment or state that no evidence was found.	This section is focused on treatments for alcohol dependence. We added the suggested point about lack of evidence for referral to the abstract, results section, and future research section of the report (but not this section on treatment for alcohol dependence in the ES), noting that none of the included studies reported follow-up with referrals as an outcome...and that evidence was insufficient to draw conclusions for follow-up with referrals. We have added more about the insufficient evidence for referrals in other places as well (see related comments 9, 96, and 102)

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Discussion	The discussion is well done, including coverage of applicability and limitations. One limitation that could be discussed more is that most studies were not designed or powered to assess the impact of the intervention on morbidity and mortality. Their focus was primarily on behavioral outcomes. Future research recommendations could include more health systems research.	We have added this to the limitations as suggested: "Studies were generally not designed to assess the impact of the interventions on morbidity and mortality; their focus was primarily on behavioral outcomes." ...and to the future research section: "We also found very few studies that measured health or utilization outcomes, with overall insufficient or low strength of evidence for the impact of behavioral interventions on mortality, morbidity, utilization, costs, and quality of life."
Peer Reviewer #3	Discussion	Page 80, Line 5: The authors may wish to add that the data in support of OPRM1 polymorphisms predicting naltrexone responsiveness is actually quite controversial and inconsistent. I have attached Gelertner et al's (2007) contradictory study for your review.	We agree and we have revised the text in this section to indicate the uncertainty/controversy, and reference that 1 study did not support those findings related to OPRM1 polymorphisms as suggested. We have also added description of additional studies that did support the findings.
Peer Reviewer #3	Discussion	The authors write, "While there are no biomarkers accurate enough to be widely accepted to measure changes in alcohol use..." There is in fact a robust literature on well-validated biomarkers, most notably for the use of PEth testing. I have included a recent example below (Marques et al, 2011).	The paper referenced here does not directly address the point. It is about PEth in the blood of drivers. This may be useful for detecting when someone has a high blood alcohol content (and thus should not drive), but it has nothing to do with screening for alcohol misuse in primary care, nor anything to do with measuring changes in alcohol use for people with risky drinking. Marques P, Hansson T, Isaksson A, Walther L, Jones J, Lewis D, Jones M. Detection of phosphatidylethanol (PEth) in the blood of drivers in an alcohol ignition interlock program. <i>Traffic Inj Prev</i> 2011 Apr;12(2):136-41.

Commentator & Affiliation	Section	Comment	Response
Julio Casoy Alkermes	Discussion	<p>Page 80 of the Draft Report states “We were unable to identify any studies that ...treated subjects with a pharmacotherapy in a primary care setting.” However, there is a published report of a primary care alcohol dependence pharmacotherapy study conducted by Lee et al (2010), the abstract for which is pasted below: The feasibility of using extended-release injectable naltrexone (XR-NTX) to treat alcohol dependence in routine primary care settings is unknown. An open-label, observational cohort study evaluated 3-month treatment retention, patient satisfaction, and alcohol use among alcohol-dependent patients in two urban public hospital medical clinics. Adults seeking treatment were offered monthly medical management (MM) and three XR-NTX injections (380 mg, intramuscular). Physicians delivered MM emphasized alcohol abstinence, medication effects, and accessing mutual help and counseling resources. Seventy-two alcohol-dependent patients were enrolled; 90% (65 of 72) of eligible subjects received the first XR-NTX injection; 75% (49 of 65) initiating treatment received the second XR-NTX injection; 62% (40 of 65), the third. Among the 56% (n = 40) receiving three injections, median drinks per day decreased from 4.1 (95% confidence interval = 2.9–6) at baseline to 0.5 (0–1.7) during Month 3. Extended-release naltrexone delivered in a primary care MM model appears a feasible and acceptable treatment for alcohol dependence. Reference: Lee, J.D., Grossman, E. DiRocco, D., Truncali, A., Hanely, K., Stevens, D., Rotrosen, J., and Gourevitch, M.N. (2010). Extended-release naltrexone for treatment of alcohol dependence in primary care. <i>Journal of Substance Abuse Treatment</i>, 39(1):14-21, 2010</p>	<p>There were also some other reviewers commenting on this point and we revised this section to be more accurate and clear (since some have mentioned this open-label cohort study). It now states: “We were unable to identify any double-blind randomized controlled trials of pharmacotherapy that identified subjects by screening in a primary care setting or that assessed the efficacy or comparative effectiveness of pharmacotherapy in a primary care setting. Further, we were unable to identify any studies of pharmacotherapy for people with risky/hazardous drinking.”</p>

Commentator & Affiliation	Section	Comment	Response
<p>Julio Casoy Alkermes</p>	<p>Discussion</p>	<p>Page 79 of the Draft Report states that alcohol dependence medications have a “low to moderate effect size.” However, O’Malley et al (2007), Yale University School of Medicine, published a post-hoc subgroup analysis of a randomized, placebo controlled study of extended-release naltrexone (XR-NTX) which found a large effect for this medication. The abstract for this publication is pasted below:</p> <p>Extended-release naltrexone (XR-NTX) is a once-a month injectable formulation that is Food and Drug Administration– approved for the treatment of alcohol dependence in patients able to abstain from alcohol before treatment initiation. This paper presents the results of an analysis of efficacy data from a subgroup of patients with 4 days or more of voluntary abstinence before treatment initiation (n = 82) on a wide range of drinkingrelated outcomes. In these patients, all of whom received counseling, the rate of abstinence was several fold higher for XRNTX 380 mg compared with placebo: median time to first drink was 41 days versus 12 days, respectively; rate of continuous abstinence at end of the study was 32% versus 11% (P = 0.02). Extended-release naltrexone 380 mg, compared with placebo, substantially increased time to first heavy drinking event (>180 days vs 20 days; P = 0.04) and decreased the median number of any drinking days per month by 90% (0.7 vs 7.2; P = 0.005) and heavy drinking days per month by 93% (0.2 days vs 2.9 days; P = 0.007). The XRNTX 380 mg group also had more than twice as many responders compared with placebo (70% vs 30%; P = 0.006; responder defined as having no more than 2 heavy drinking days in any consecutive 28-day period) and experienced greater improvement in +- glutamyl transpeptidase levels (P = 0.03). Outcomes for XR-NTX 190 mg (n = 26) were generally intermediate, demonstrating a doseresponse effect. In conclusion, XR-NTX 380 mg prolonged abstinence and reduced the number of heavy drinking days and drinking days in patients who were abstinent for as few as 4 days before treatment initiation. Reference O’Malley, S.S., Garbutt, J.C., Gastfriend, D.R., Dong, Q., Kranzler, H.R. (2007). Efficacy of extended-release naltrexone in alcohol-dependent patients who are abstinent before treatment. J Clin Psychopharmacol 2007;27:507–512. For important safety information, please refer to Vivitrol full prescribing information, including boxed warning at: http://www.vivitrol.com/pdf_docs/prescribing_info.pdf</p>	<p>We appreciate Dr. Casoy’s comment that the effect size for the subgroup analysis of XR-NTX was larger than the low to moderate effect size we report for naltrexone from the Kranzler and Van Kirk review. However, the effect sizes we reference are “overall” effect sizes calculated from a number of studies. Indeed, some of the initial <i>individual</i> oral naltrexone studies had much larger effect sizes as well. Our goal here was to give the reader a global sense of the effectiveness of naltrexone and acamprosate based on a broad range of studies rather than dig into more subtle aspects of this body of work and highlight one study with a larger effect size. For additional clarity, We have changed the sentence to note that this is an average estimate.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #6	Discussion/ Conclusion	The implications and major findings are clearly stated and well written. Although there was mention of assessment reactivity in the limitations section, the subject of research reactivity more broadly could be considered in more detail. This is an important consideration in understanding the results of this field of research and it should be more directly stated (and in the abstract and executive summary) that the findings of prior studies are likely very conservative because of this important issue.	<p>We have included a paragraph about the assessment issue in the ES as well as in the Discussion. From the ES: "It is possible that the assessments of alcohol misuse conducted in the included trials conceal therapeutic benefits of the behavioral interventions (i.e., bias results toward the null). Many studies included extensive assessment of alcohol-related behaviors, which could directly result in behavior changes. The control groups in the included studies generally reduced alcohol consumption. Some possible explanations for changes in behavior as a result of the screening and screening-related assessment include (1) increased awareness of the extent of their drinking; (2) the screening questions prompted them to discuss drinking with their primary care provider at a subsequent visit; (3) receipt of some minimal intervention, such as printed educational materials about general health or about alcohol specifically (control groups in the included studies often received some printed materials); or (4) regression to the mean. One study empirically tested whether brief assessment (without a behavioral intervention) reduces hazardous drinking by comparing brief assessment with a control that did not include assessment. The study concluded that assessment appears to reduce hazardous drinking but noted a potential limitation of measurement artifact due to social desirability bias.⁴³"</p> <p>We also have included this in the Abstract of the draft paper for Annals of Internal Medicine that we submitted. "Limitations: Assessments of alcohol misuse in the included trials could bias results toward the null"</p>
Peer Reviewer #7	Discussion/ Conclusion	The implications of the findings are clearly stated.	Thank you

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	Discussion/ Conclusion	The limitations of the report could be expanded to include those of the meta-analysis methodology in general ('best evidence meta-analysis' as used for this report vs. casting a wider net to avoid unwanted subjectivity in methodological selection criteria).	We do not agree that this should be included in the limitations. We also don't agree that there is unwanted subjectivity. It is not clear what limitation the reviewer is referring to. In addition, we have been transparent with our methodology.
Peer Reviewer #7	Discussion/ Conclusion	<p>Although the questions, findings and discussion are by necessity narrowly focused on alcohol use interventions in primary care settings, the discussion might be broadened to include the more progressive direction of screening, brief intervention and referral research and implementation activities which are less focused on single substance targeting (e.g., alcohol only). As noted earlier, the overlap of multiple substance use is large in at-risk populations and research that identifies simultaneous or sequential interventions for co-occurring alcohol, tobacco or other drug use should be a primary focus of future activities.</p> <p>Although not randomized trials, the recent initiatives implemented by SAMHSA for example, with regard to its SBIRT state, campus and medical residency programs, has provided the opportunity to understand the feasibility of screening for multiple substance use and the opportunity to develop interventions that address interrelated substance use risk factors. These initiatives might be discussed in relation to informing policy and practice decisions as they have moved forward not only the initiation of CPT code approval for the reimbursement of tobacco and alcohol brief counseling services, but other means of sustaining services within primary care settings as well. As stated in the first paragraph of the introduction of this report, tobacco use remains the primary cause of preventable mortality in this country. Providing tobacco screening and counseling as one component of an alcohol screening and counseling package represents a much easier "sell" to policy makers and providers.</p>	<p>We appreciate the importance of these issues, but this report is focused on screening and interventions for alcohol misuse. We believe that a discussion of screening for multiple substance use is beyond the scope of this report and would detract from the important findings by bringing in more discussion of tobacco and other drugs. These issues, while interesting, should be the subject of other works---both for the purposes of keeping this report focused on the topic/questions and for giving appropriate attention to all of the issues surrounding screening and intervening for co-occurring problems (as this would be a substantial undertaking).</p> <p>Of note, our report does address the lack of evidence on people with multiple substance use problems in the context of screening and behavioral interventions for alcohol misuse</p>
Peer Reviewer #7	Discussion/ Conclusion	Although the report discusses an expansion of the definition of primary care, the definition does not stretch far enough. Many at-risk populations utilize primary care services offered by community health centers or Federally Qualified Health Centers which have multiple "primary care" services housed under a single roof. Screening and brief counseling services offered in primary care dental sites, for example, represents a growing line of research and implementation activities.	Thank you for the comment. Community health centers and FQHCs would generally fit within our definition of primary care. Dental sites, however, would not and we did not intend to stretch the definition further.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	Discussion/ Conclusion	The authors might also mention in the discussion about the growing use of electronic medical records which often include electronic screening assessments or prompts for brief counseling services. The time to administer most screening tools, one of the major concerns of busy primary care providers, is greatly shortened by the use of electronic methodology.	Interesting point, but this is speculation. The time to administer the instruments that should be used (by our assessment) is already as short as 1-2 minutes (AUDIT-C and single question). So, it isn't really conceivable to "greatly shorten" that time. Further, there is not empiric evidence that electronic methodology would shorten the time to administer. In fact, it would require logging on to a computer, which could lengthen the time to administer some of these, depending on the characteristics of the patients, practice, and provider and how much already is done through the EMR.
Peer Reviewer #1	Clarity and Usability	Yes. Particularly if a bit more info is added re how to translate—eg training investment and the fact that screening and assessment were also done by staff specifically hired to do so. Also the report will have its usability greatly improved if the terminology used is corrected and updated to contemporary use and understanding of these conditions.	Thank you. See responses to previous comments from this reviewer for the more detailed description and our responses.
Peer Reviewer #2	Clarity and Usability	Yes, the report is well-structured and useful.	Thank you
Peer Reviewer #3	Clarity and Usability	Excellent and timely review.	Thank you
Peer Reviewer #6	Clarity and Usability	The report structure and organization is excellent. Conclusions are written in a manner that is informative to policy, practice, and future research.	Thank you
Peer Reviewer #7	Clarity and Usability	The report is well structured and organized. The conclusions and main points might be broadened (as described above or in other ways) to provide a larger framework in which to inform policy and practice decisions.	Thank you See responses to previous comments from this reviewer.
Peer Reviewer #1	General	The target population is "alcohol misuse." This terminology has the advantage of being the same term used in the USPSTF's last report on this topic. However, it causes great confusion about the population defined. This is not simply an academic point. It has real consequences for patients and the public because the term is used to mean different things by different users of the term. For example, the VA uses "misuse" to mean the full spectrum of alcohol use that risks or causes consequences (including dependence). Labeling someone with dependence (a diagnosable disease according to the International Classification of Diseases	We appreciate these comments. We believe that the most important thing is to be clear in how we've defined the terminology, and that it is clear that alcohol misuse is an "umbrella"/overarching term, as we've defined it. Further, we believe it is actually synonymous with unhealthy alcohol use---as we describe in the report. We have discussed the differences in VA and CMS use of the term misuse with our TOO---

Commentator & Affiliation	Section	Comment	Response
		<p>and the APA's Diagnostic and Statistical Manual) as having "misuse" perpetuates stigma and is inaccurate. On the other hand, the VA is to be commended for recognizing that the whole spectrum from risky use through dependent use of alcohol is what is clinically important (they simply have the name wrong).</p> <p>In another part of the Federal government, the Center for Medicare and Medicaid Services (CMS), the term is used completely differently, and with serious and unfortunate consequences. This month (October 2011) CMS announced a coverage decision for alcohol screening and intervention. It used the term "alcohol misuse." CMS's use of the term specifically excluded dependence. As a result, the national coverage decision is to pay for screening for misuse and then intervention for misuse but not dependence. Clinicians in the real world cannot screen only for nondependent misuse. Screening tests identify the spectrum of unhealthy use from risky use through dependence and as such clinicians who screen identify all such patients and then need to address them whether they have dependence or not. When they identify people with dependence, their work falls beyond the national coverage decision.</p> <p>AHRQ and the USPSTF have an opportunity to be clear about what screening identifies. The current report is clear in part by actually addressing the full spectrum, which is very important. But retaining a dated and confusing term for that full spectrum has already had and will continue to have harmful consequences for people with the spectrum of unhealthy alcohol use defined by the Institute of Medicine in 1990 (which is the only published and defined term that includes the spectrum from risky use through problem use through abuse and dependence). Using the term "misuse" simply to be consistent with the last USPSTF report seems not a good enough reason to perpetuate confusion and risk bad care as a result when a simple correction and update could solve the problem with little if any downside.</p>	<p>our understanding is that the two will no longer be in conflict after this report comes out, because CMS will use the USPSTF recommendation (and the terminology in it), making the two consistent. Thus, it will actually be beneficial for this report to use the term alcohol misuse---because it should result in the VA and CMS having consistent terminology and definitions.</p> <p>The points that have been raised about bad care or about stigma (i.e., concern with considering people with dependence to be in the broader category of alcohol misuse) are theoretical and are not widely agreed upon. There is no evidence to suggest that these things would be improved by using the term "unhealthy alcohol use"---we're not convinced that would have any less stigma or that it would improve care somehow.</p>

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	General	<p>Another terminology issue is confusing. The report uses the term “binge” to refer to heavy drinking episode or heavy episodic drinking. NIAAA defines binge as a number of drinks per unit time. The Journal of Studies on Alcohol and Drugs does not allow use of the term binge in its peer-reviewed publication because of the confusion it generates (meaning anything from 4-5 drinks on an occasion, to 4-5 drinks in a day, to 4-5 drinks in 2 hours, to drinking continuously for several days). As such it would be advisable to use a clearer more accurate term in this report such as heavy drinking episode. (It is also notable that the term “binge” was not defined in the report).</p> <p>Lastly re terminology, another reason it is so critical for this report to get the terminology correct is demonstrated by the report itself. The report talks about harmful drinking as if it is not a diagnosable disorder. Yet leading writings in the field (Hasin D), and the International Classification of Diseases, note that it IS a diagnosis and it is almost identical to DSM IV abuse. Yet this report lumps harmful with hazardous/risky and separates it from abuse. This report will have great impact in the field and how these issues are discussed so the terminology really needs to be accurate. This is not a small semantic or grammatical issue.</p>	<p>We now use the term heavy drinking episodes or heavy episodic drinking throughout most of the report. However, we have retained the use of “binge” in a few places where it makes sense to do so. For example, when describing the details of a study that used and defined “binge drinking” [e.g. from Saitz, “...percentage of subjects with any binge drinking, defined as more than three drinks per occasion for women and older adults and more than four for men, was among the secondary outcomes reported at 6 months. Among patients in the intervention group who saw resident physicians, 44 percent (95% CI, to 30, 58) reported any binge drinking compared with 64 percent (95% CI, 45 to 79) in the control group ...”]</p> <p>We agree that harmful drinking/harmful use is an ICD diagnosis (and have defined it by the ICD definition in the report) and that it is most similar to alcohol abuse. We have edited the report to reflect this and no longer lump it with risky/hazardous throughout the narrative.</p>
Peer Reviewer #2	General	The report is clinically meaningful. The tables and text define the continuum of alcohol misuse appropriately.	Thank you
Peer Reviewer #3	General	This is an exceptionally well-designed and well-written systematic review of the literature, with very little room for criticism. The key questions are quite explicitly addressed, and are wholly consistent with the purpose of the review.	Thank you
Peer Reviewer #6	General	The report is very well written and well-organized and the key questions are appropriate and explicitly. The report is, for the most part, highly clinically useful with the exceptions and concerns listed below (in specific comments).	Thank you

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	General	The systematic review is well-organized and should be clinically useful to the variety of targeted audience members specified in report including providers, purchasers, and government programs. The findings are applicable to medical patients with risky, hazardous or harmful drinking identified by screening in primary care. Key questions have been expanded from the previous USPSTF2004 questions to include screening and behavioral interventions for the full spectrum of alcohol misuse including dependence. Referral as an intervention was also included.	Thank you
Peer Reviewer #7	General	The key questions are clearly stated, however the questions might be reordered to avoid repetitive text and confusion when reading the questions. For example, could the “adverse” effects KQs (3,5) be placed together (especially since there were few findings for these items), and perhaps answer the intermediate outcome questions prior to the longer-term outcomes questions? KQ 1 seems like a distraction, out of order with the rest of the questions and might better be combined with question 6.	We understand that the KQs could be placed in a variety of different orders. However, we explored the wording and order of questions in great detail during the early stages of the project the current order is the one preferred by the vast majority of our team of researchers, our TEP, and the USPSTF. KQ1 is first because it is the overarching question that would included any studies that randomized people or practices to screening vs. no screening (or usual care), whereas KQ6 is downstream from this and is focused on studies that take people that have already screened positive and then randomizes them to behavioral intervention vs. an eligible control group.
Peer Reviewer #8	General	This is a comprehensive and well-written systematic review/meta-analysis of the key studies on screening and brief interventions in primary care for alcohol misuse. Such reviews are regarded as crucial to the evidence-based approach to policy, treatment and service delivery in health care. It is highly systematised, along the lines of reviews undertaken by the Cochrane Collaboration and (in recent years) the World Health Organization. Overall, I would judge the present review to be at least the equal of the reviews undertaken by previous authors in this area and in certain respects is superior to them.	Thank you.
Peer Reviewer #8	General	However, I have reservations about the coverage of the scientific literature addressing some of the key questions. It seems that the central search strategy has elicited adequately the literature relevant to Questions 1, 2, 4a, 4b, and 6 (with caveats), but less so to Questions 3 and 5. I do not think it has accessed the literature on Question 7 satisfactorily, and suggest that this key question is deleted from the resent review (see later comments).	See responses to the later, more detailed, comments from this reviewer that explain the reviewer’s perspective.

