ADDICTION

SSA SOCIETY FOR THE STUDY OF ADDICTION

FOR DEBATE

doi:10.1111/add.13136

Reviewing and interpreting the effects of brief alcohol interventions: comment on a Cochrane review about motivational interviewing for young adults

Sean Grant, Eric R. Pedersen, Karen Chan Osilla, Magdalena Kulesza & Elizabeth J. D'Amico

RAND Corporation, Santa Monica, CA, USA

ABSTRACT

Background Cochrane recently published a systematic review on motivational interviewing (MI) for alcohol misuse in young adults. The review authors concluded that 'there are no substantive, meaningful benefits of MI interventions for the prevention of alcohol misuse' (p. 2), as effect sizes were 'small and unlikely to be of any meaningful benefit in practice' (p. 27). As most of these interventions were quite brief, we wish to open a dialogue about interpreting effect sizes in this review and of (brief) alcohol interventions more generally. **Analysis** We analyze four methodological aspects of the review that likely influenced the author's conclusions about intervention effects: (1) risk of bias assessments, (2) search strategies, (3) assessing the quality of the body of evidence and (4) definitions of sustainability and clinical significance. **Conclusions** We interpret the effect sizes found in this review to indicate modest yet beneficial and potentially meaningful effects of these interventions, given their brevity and low cost. This interpretation is consistent with other reviews on brief, MI-based interventions and brief interventions more generally. We therefore encourage the field to re-open dialogue about the clinical importance of the effects of MI on alcohol misuse by young adults. Rather than dismissing interventions with small effects, we believe a more fruitful way forward for the field would be to catalogue effect sizes for various alcohol interventions. Such a catalogue would help stakeholders themselves to choose which interventions meet their minimum desired impact, and thus may be suitable given their targeted populations, setting and resources.

Keywords Brief intervention, Cochrane, effect size, GRADE approach, minimal clinically important difference, motivational interviewing, risk of bias, systematic review.

Correspondence to: Sean Grant, RAND Corporation, 1776 Main Street, Santa Monica, CA 90407, USA. E-mail: sgrant@rand.org
Submitted 3 February 2015; initial review completed 5 May 2015; final version accepted 17 August 2015

INTRODUCTION

Cochrane has published a systematic review on Motivational Interviewing (MI) for alcohol misuse in young adults [1]. Systematic reviews and meta-analyses of alcohol use interventions provide robust summaries of intervention effects, setting agendas for future research while informing policy and practice [2,3]. We commend the review authors, along with the oversight of the Cochrane Drugs and Alcohol editorial group, for conducting a review on this important topic. MI-based interventions are common alternatives to more didactic approaches that are more lecture-based and provide education about alcohol [4,5]. Given the prevalence of MI approaches, understanding their effectiveness is clearly important for the field. The review authors have a notable history of producing systematic reviews instrumental to understanding the effectiveness of various alcohol preventive interventions for young people [6–12]. Similarly, Cochrane is methodologically exemplary in setting international standards for conducting systematic reviews [13].

The review authors concluded that 'there are no substantive, meaningful benefits of MI interventions for the prevention of alcohol misuse' (p. 2), stating that effect sizes were 'small and unlikely to be of any meaningful benefit in practice' (p. 27) [1]. These conclusions were surprising, as they are discrepant with conclusions of other reviews on similar interventions with similar effect sizes [1,14,15]. This review's conclusions indicate that the field should discuss what should be expected from

brief interventions and how best to use them—or to caution their dissemination and focus more research on alternative intervention approaches, as suggested by a recent individual participant data (IPD) meta-analysis of a subset of trials evaluating brief motivational interventions for college student drinking [16,17]. Careful interpretation of intervention effects is crucial to using systematic reviews for informing future research, policy and practice, as summary conclusions are more likely to attract attention than specific effect sizes from meta-analyses [18–21]. Critical appraisal of this systematic review is pertinent to understanding its conclusions and why they may be discordant from other reviews: the specific review methods employed can influence review findings and how they are interpreted. Any 'disciplinary dialogue' arising from such a critical appraisal and responses to it are also likely to apply to future studies and reviews of brief interventions.

In this paper, we wish to open a dialogue about interpreting effect sizes in this review and of behavioral interventions more generally. Although the review authors' interpretations of MI effects may eventually be considered the most appropriate conclusions, we believe some of the review's methods prompt constraint in making definitive assertions that MI does not have meaningful benefits for young adult alcohol misuse. We will discuss how the following methodological aspects of the review likely influenced the effect sizes found and their interpretation: (a) risk of bias assessments [22], (b) search strategy, (c) application of the GRADE approach to assess quality of the body of evidence [23] and (d) definitions of sustainability and clinical significance of intervention effects.

RISK OF BIAS

Assessing risk of bias is a key step in conducting systematic reviews. Notably, these assessments play an important role in grading the quality of evidence reviewed and thus directly influence final interpretation of review findings [23]. One question to ask when appraising a systematic review is: were appropriate risk of bias assessment procedures used?

This review used the Cochrane risk of bias tool, which prompts review authors to judge risk of bias as high, low or unclear for a series of items that cover established domains of bias in randomized trials (e.g. proper randomization), and to provide their rationale for each judgment [22]. We agree with the need for conducting risk of bias assessments in systematic reviews. However, we believe there are some issues concerning how the Cochrane risk of bias tool was used in this review that affected interpretations of review findings.

In re-analyzing the information reported in the review, it appears that 11% of the risk of bias assessments were inconsistent with either Cochrane guidance (e.g. rating card-shuffling as 'unclear risk' rather than 'low risk' for random sequence generation), or with other assessments within the review (e.g. rating reporting of all outcomes as 'unclear risk' for selective outcome reporting for some studies and 'low risk' for others). In addition, 39% of risk of bias assessments were 'unclear risk', meaning the review authors did not have sufficient information to tell whether risks for particular biases were high or low. Information needed to complete various aspects of a systematic review is often missing from published reports of primary studies. To address unclear risks of bias and other missing information, Cochrane guidance requires review authors to contact trial authors for missing data [13]. However, the review authors did not report that they contacted the trial authors, and our own contact of the trial authors supported this—of 69 publications, 58 contact authors (84%) replied to our inquiries; of those, 51 (88% of respondents; 74% of all contact authors) either do not recall or have no record of being contacted by the review authors. Despite the responsiveness of trial authors to our e-mails, it is still possible that the review authors attempted to contact them and simply did not report this information. Regardless, the percentage of 'unclear' risk of bias assessments should prompt caution in making definitive assertions about MI effects. It is possible that if more studies were found to be at low risk of bias the quality of evidence would have been graded higher which, in turn, would impact interpretations of effects, as discussed below.

In addition, 94% of trials were judged 'high risk' for performance bias because participants and providers were not blind to their assigned intervention condition. Performance bias is standard practice for reviews on pharmaceuticals, and double-blind procedures are an important indicator of study quality. However, behavioral intervention trials are generally 'open-label' trials [24]: it is generally impossible to blind providers and participants because clinicians must know what type of intervention they are delivering, and recipients' awareness of and engagement with intervention content is often necessary for the intervention to work [25]. The review authors even note, in some risk of bias judgments, that it is 'not possible [to] blind participants to intervention' (e.g. p. 37). Inability to blind participants and providers is a methodological challenge inherent in most behavioral interventions. Because risk of bias assessments should be relevant to the content area, we propose that lack of blinding providers and participants should be assessed as a limitation only if such blinding is possible. We encourage Cochrane to evaluate whether formal inclusion of performance bias is appropriate for behavioral intervention trials, or make the distinction between 'lack of blinding' providers and participants and 'inability

to blind' them. This precedent has been set by other Cochrane reviews on health behaviors [26,27], although criteria alternative to blinding are needed to judge performance bias in this area, rather than not judging studies by this risk of bias at all [28]. It is worth emphasizing, however, that blinding of outcome assessors is encouraged in behavioral intervention trials, especially given that subjective outcome measures are used prevalently in this area of research [25].

SEARCH STRATEGY

Comprehensive searches for eligible studies from various information sources are a key element of systematic reviews [29]. Another question in appraising a systematic review is: were relevant studies missed by the search strategy? Although challenging to identify, 'grey' literature—reports that are not controlled by commercial publishers—can provide a more complete view of available evidence by identifying documents linked to journal articles (e.g. dissertations with additional data) and new studies altogether (e.g. studies only reported in dissertations). Searching the grey literature is particularly important in this area, as many MI interventions for young adult alcohol use are evaluated in dissertations [30].

The grey literature was hypothetically eligible in this review. However, the reported search strategy is limited to bibliographic databases of commercially published research and reference lists of included studies and topic-related systematic reviews. Moreover, the review's references section lists 'published data only' as the data source for every included study. This could lead to a significant amount of missing data for studies included in the review, and leave eligible studies unidentified. For example, one author of our paper is conducting a systematic review on Brief Alcohol Screening and Intervention for College Students (BASICS) [31]: a manualized, MI-based intervention for young adult alcohol misuse that meets this Cochrane review's eligibility criteria. Consequently, all randomized trials evaluating BASICS that were published before October 2013 should be included in this review. The Cochrane review identified some manuscripts on BASICS; however, there are 20 studies (11 of which are reported in graduate theses/dissertations) published before October 2013 that are not included in this Cochrane review, as well as 14 papers that provide additional information to studies included in this review that were not cited or referenced in the review (see Supporting information, Supplement 1). The above-mentioned concerns with risk of bias assessments relying on reported data only are compounded by the lack of specifically searching grey literature sources, where further information about 'unclear' risks of bias could potentially be found. Moreover, at least two reports thought by review authors to be from different studies appear to be the same study (Barnett *et al.*, 2010 [32] and Monti *et al.*, 2007 [33]). Limitations in the search strategy's ability to identify and properly classify the primary literature in this area should further temper definitive conclusions.

QUALITY OF EVIDENCE

Grading the quality of evidence aims to indicate the likelihood that future research could change the direction or magnitude of estimated intervention effects for the population and outcomes reviewed [34]. A relevant appraisal question is: was grading quality of evidence conducted appropriately? Cochrane encourages using the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) working group guidelines [23]. This approach assesses quality of evidence for each individual review outcome and details these assessments in a Summary of Findings table [13]. Risk of bias assessments are a component—along with directness of evidence, effect size heterogeneity, effect estimate precision and risk of publication bias—that determines quality of evidence. Consequently, concerns about risk of bias assessments and search strategies are not trivial, as these formally influence interpretation of effects using GRADE quality ratings [23]. Our concern lies not with the GRADE approach, but with the need for more detailed justifications for downgrading evidence.

The review authors primarily discounted the statistically significant yet small-to-moderate effect sizes favoring MI in the Summary of Findings table because of risk of bias assessments. Of the eight outcomes summarized in this table, seven had their quality of evidence downgraded due to 'risk of bias' or 'potential risk of bias', with no explanation of which biases led to downgrading. Systematic reviewers should delineate the component risks of bias that led to evidence downgrading rather than reporting 'holistic' judgments (e.g. downgrading due to 'risk of bias'). As transparency is a key feature of both the Cochrane risk of bias tool and the GRADE approach, reviewers need to document carefully and justify the procedures used to assess the quality of the evidence (e.g. explicit decision rules for combining risk of bias domains in quality of evidence grades) to allow readers to appraise decisions underlying evidence grades [35].

For this review, greater detail on the use of risk of bias assessments to determine GRADE quality ratings would have been helpful to appraise interpretation of effect sizes critically. Currently, readers cannot discern which risks of bias were used to downgrade evidence, as the review authors reported only 'risk of bias' for the rationale where applicable. Moreover, quality of evidence ratings led the review authors to conclude that they 'cannot rule out the

possibility that the effects observed in this review may be exaggerated due to methodological limitations' (p. 28) [1]. Without details on decisions leading to quality-of-evidence ratings, we cannot assess clearly whether conclusions are imprecise or even misleading [36], as the aforementioned issues with risk of bias assessments call these evidence grades into question. For example, the review authors may have downgraded evidence for inability to blind providers or participants, which would mean that, at best, behavioral interventions could have moderate quality of evidence if discounted for this reason. We ask experts involved in brief alcohol interventions—and behavioral interventions more generally—to discuss publicly whether lack of blinding participants or providers is appropriate to include in risk of bias or quality of evidence assessments for such interventions.

SUSTAINABILITY AND CLINICAL SIGNIFICANCE OF INTERVENTION EFFECTS

Summary of Findings tables in systematic reviews intend to present main findings in a transparent, simple format. Sustainability of effects post-intervention is one important consideration for summarizing interventions with potential public health impact [37], and finding sustainable interventions that have long-term effects on preventing young adult alcohol misuse is a key goal for the field. However, it is important to reflect upon the nature and purpose of an intervention when defining and assessing its sustainability. MI approaches for young adult populations are often used as brief preventive interventions targeting escalation of drinking during particular high-risk times. We therefore pose one final critical appraisal question: was the primary focus on effects after 4 months or longer appropriate?

The Summary of Findings table and conclusions for this review—which reported only on effects at 4 months and after—describe MI as not producing effects that reach a minimal clinically important difference (MCID). However, we contend that the review authors have used a particular perspective on MCIDs and expected sustainability of this intervention with which other stakeholders may disagree. First, the MI interventions in the Cochrane review were quite brief: in 49 of 66 included studies (74%) the MI intervention was one session, and seven studies (11%) had two sessions. The 'longest' intervention included in the review involved five sessions over a 9-hour period, while the shortest was 15 minutes, and in 43 trials sessions took 1 hour or less (65%). Given the purpose, brevity and probable cost-effectiveness of these interventions [38,39], it may be questionable to rest chief conclusions about effectiveness on 'sustainable effects' and define sustainability at 4 months or longer, as many practitioners may not expect or use these brief interventions to have large and long-lasting effects. For example, most of these interventions were conducted with college students; reducing drinking over a 4-month period in college (which is typically a semester in the United States) is important, as it may prevent the escalation to heavier drinking patterns and subsequent consequences, such as risky sexual behavior and academic problems [40].

This methodological decision is not trivial, as outcomes differ significantly based on follow-up period. Intervention effects up to 4 months were larger than effects after 4 months for seven of the eight outcomes in this review —and were at least twice as large for four of these outcomes. As written, it is unclear why review authors defined 4 months as the cut-off for sustainability. One might argue that it is clinically meaningful for a brief intervention that does not address structural determinants of alcohol misuse (e.g. culture of drinking on college campuses, outlet density and access) to demonstrate observable effects up to 4 months, and smaller yet still observable effects beyond 4 months. Moreover, the effect size from every meta-analysis in this review favors MI, 19 of 34 (56%) effect sizes demonstrate statistically significant effects, and no studies reported harms for the intervention (see 'Comparison 1' in the review). Given the above, we were surprised by the review's strong conclusions unfavorable to MI, especially as the review authors noted that their *conclusions* differed from another Cochrane review on MI for drug use, even though the other review had similar *effect sizes* [41]. Discrepancies in interpreting similar effect sizes for similar interventions indicate that MCIDs are not defined uniformly in this area, and the review authors' conclusions represent a particular *interpretation* of effects (rather than the effects themselves) that differs from other review authors' interpretations [1,14,15].

SUGGESTIONS MOVING FORWARD

We propose several suggestions moving forward. First, the newly developed Methodological Expectations of Cochrane Intervention Reviews (MECIR) provides promising standards for the conduct and reporting of Cochrane reviews [42]. We support recommendations to describe the application of the GRADE approach in the Methods section more fully, explain decisions about quality of evidence in reporting results and incorporate information about quality of evidence in discussion sections of Cochrane reviews. For example, authors reviewing behavioral interventions could use a checklist to improve the consistency, reproducibility and transparency of GRADE assessments [43]. In addition, although MECIR does not make searching for grey literature mandatory, we recommend that the Cochrane Drug and Alcohol Group consult experts on a given review

topic to determine whether grey literature sources (e.g. dissertations) may provide relevant evidence.

There is also a pressing need to adapt existing 'risk of bias' and 'quality of evidence' assessment tools for behavioral interventions [44–46]. For example, alternative strategies to minimize performance bias (such as preventing participant and provider awareness of primary research hypotheses) could be developed by trial researchers and then used by systematic reviewers in risk of bias assessments when blinding participants and providers is not an option [28]. Assessing risk of performance bias becomes more challenging when evaluating automated online interventions that reduce risk of provider performance bias, but not participant performance bias [47]. Challenges may also arise when comparing a behavioral intervention to pharmacological interventions —although only when a placebo control is offered, as participants and providers would be aware of the presence of a pill in assigned intervention conditions [48]. Cochrane is well positioned to take a leading role in developing and enforcing evidence-based guidance on assessing risk of bias and quality of evidence tailored to behavioral interventions [44].

Lastly, we believe that further work can be conducted to develop benchmarks for MCIDs. Consensus on MCIDs for various outcomes targeted by substance use interventions is currently lacking. As such, interpreting the meaning of various effect sizes in discussion sections of systematic reviews is left largely to the views of researchers—and to a certain degree peer-reviewers and editors. Depending on their background and experiences, stakeholders may have different views about the clinical significance of a given effect size. Some providers may find a small effect size meaningful for particular type of intervention used with specific clients, whereas some researchers may find this same effect size as not meaningful if they are focused on long-term outcomes or widespread dissemination.

Rather than dismissing interventions with small effects as insubstantial, a more fruitful way forward for the field may be to engage stakeholders from various groups in establishing MCIDs for important outcomes in substance use research, given that MI interventions—brief or otherwise—are only one of many interventions available to address alcohol misuse. For example, groups developing core outcome sets could conduct Delphi processes [49] with researchers, practitioners and clients [50] to determine which outcomes each finds most important, the smallest clinical effect that would be meaningful to them and reasons for differences in the selection of core outcomes and definitions of MCIDs by stakeholder group [51]. These consensus-based MCIDs could then be used to inform interpretations of effect sizes in systematic reviews that are sensitive to the views of important stakeholder groups. They could also be used by evidence-based program registers—such as the US Substance Abuse and Mental Health Services Administration's National Registry of Evidence-based Programs and Practices [52]—to catalogue intervention effect sizes. With hundreds of RCTs on alcohol treatment alone [53], and more than 70 reviews published just by the Cochrane Drugs and Alcohol Group [54], such a catalogue would help stakeholders to choose which interventions meet their own definitions of MCIDs and thus may be suitable given their targeted populations, setting and resources.

CONCLUDING THOUGHTS

As a product of Cochrane, this review has potentially far-reaching impact. Its conclusions have already been disseminated and discussed in research circles [18], mental health practice and consumer resource circles [19,20] and lay media [21]. Several reviews on MI-based interventions indicate that these interventions have modest effect sizes. Other reviews in this area have interpreted such effects as beneficial and potentially meaningful, given the brevity and inexpensiveness of these interventions, as well as the focus on changing behavior during high-risk periods [14,15]. Moreover, several recent reviews provide support for the theory of change underlying MI interventions [55,56]. Effect sizes common to these reviews suggest that these brief interventions are one useful tool in the public health toolkit for preventing young adult alcohol misuse. As such, we encourage the field to re-open the dialogue about the effects of MI on alcohol misuse in young adult populations. Namely, we ask interested stakeholders to reconsider whether conclusions are commensurate with the information provided in the Cochrane review, effect sizes of these brief interventions are of clinical importance, and future research and policy should prioritize these interventions or others for young adult alcohol misuse. Given the widespread use of MI and brief interventions, such discussion is imperative before making changes to future research, policy and practice.

Declaration of interests

S.G.'s fiancée is a salaried-employee of Eli Lilly and Company, and owns stock. S.G. has accompanied his fiancée on company-sponsored travel. E.P., K.O., M.K. and E.J.D. have received and been supported on several grants from NIH to conduct trials on motivational interviewing for substance use. K.O. and E.J.D. are members of the Motivational Interviewing Network of Trainers (MINT). K.O. and E.J.D. train practitioners in motivational interviewing for substance use. K.O. and E.J.D. consult for several organizations regarding motivational interviewing for substance use. All authors have been and/or are currently involved

in research evaluating motivational interviewing (MI) and MI-based interventions.

Acknowledgements

The authors wish to acknowledge David Humphreys and Emily Tanner-Smith for helpful comments on earlier drafts of this manuscript.

References

1. Foxcroft D. R., Coombes L., Wood S., Allen D., Almeida Santimano N. M. L. Motivational interviewing for alcohol misuse in young adults. *Cochrane Database Syst Rev* 2014; **8**CD007025.
2. Wilson D. Meta-analyses in alcohol and other drug abuse treatment research. *Addiction* 2000; **95**: S419–38.
3. Amato L., Davoli M., Vecchi S., Ali R., Farrell M., Faggiano F., *et al.* Cochrane systematic reviews in the field of addiction: what's there and what should be. *Drug Alcohol Depend* 2010; **113**: 96–103.
4. Larimer M. E., Cronce J. M. Identification, prevention, and treatment revisited: individual-focused college drinking prevention strategies 1999–2006. *Addict Behav* 2007; **32**: 2439–68.
5. Cronce J. M., Larimer M. E. Individual-focused approaches to the prevention of college student drinking. *Alcohol Res Health* 2011; **34**: 210–21.
6. Moreira M., Smith L., Foxcroft D. Social norms interventions to reduce alcohol misuse in university or college students. *Cochrane Database Syst Rev* 2009; **3**: CD006748.
7. Foxcroft D., Tsertsvadze A. Universal alcohol misuse prevention programmes for children and adolescents: Cochrane systematic reviews. *Perspect Public Health* 2012; **132**: 128–34.
8. Foxcroft D., Tsertsvadze A. Universal family-based prevention programs for alcohol misuse in young people. *Cochrane Database Syst Rev* 2011; **9**: CD009308.
9. Foxcroft D., Tsertsvadze A. Universal multi-component prevention programs for alcohol misuse in young people. *Cochrane Database Syst Rev* 2011; **9**: CD009307.
10. Foxcroft D., Tsertsvadze A. Universal school-based prevention programs for alcohol misuse in young people. *Cochrane Database Syst Rev* 2011; **5**: CD009113.
11. Foxcroft D., Lister-Sharp D., Lowe G. Alcohol misuse prevention for young people: a systematic review reveals methodological concerns and lack of reliable evidence of effectiveness. *Addiction* 1997; **92**: 531–8.
12. Foxcroft D., Ireland D., Lister-Sharp D., Lowe G., Breen R. Longer-term primary prevention for alcohol misuse in young people: a systematic review. *Addiction* 2003; **98**: 397–411.
13. Higgins J., Green S., editors. Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0 5.1.0updated March 2011. Chichester, UK: Wiley-Blackwell; 2011.
14. Tanner-Smith E., Lipsey M. Brief alcohol interventions for adolescents and young adults: a systematic review and meta-analysis. *J Subst Abuse Treat* 2014; **51**: 1–18.
15. Cushing C., Jensen C., Miller M., Leffingwell T. Meta-analysis of motivational interviewing for adolescent health behavior: efficacy beyond substance use. *J Consult Clin Psychol* 2014; **82**: 1212–8.
16. Mun E. Y., de la Torre J., Atkins D. C., White H. R., Ray A. E., Kim S. Y. *et al.* Project INTEGRATE: an integrative study of brief alcohol interventions for college students. *Psychol Addict Behav* 2014; **29**: 34–48.
17. Huh D., Mun E. Y., Larimer M. E., White H. R., Ray A. E., Rhew I. C. *et al.* Brief motivational interventions for college student drinking may not be as powerful as we think: an individual participant-level data meta-analysis. *Alcohol Clin Exp Res* 2015; **39**: 919–31.
18. Davoli M., Amato L. Large evidence base, small effects: motivational interviewing for alcohol misuse in young adults (editorial). *Cochrane Database Syst Rev* 2014; **9**: ED000088.
19. Center for Advancing Health. Counseling has limited benefit on young people drinking alcohol 2014. Available at: http://www.webcitation.org/6YhjFWVGF. (accessed on 21 May 2015)
20. Mad in America. Specialized counseling approach does not help with drinking problems 2014. Available at: http://www.webcitation.org/6YhjPi9lR. (accessed on 21 May 2015)
21. Science Daily. Counseling has limited benefit on young people drinking alcohol, study suggests 2014. Available at: http://www.webcitation.org/6YhjSe0Iy. (accessed on 21 May 2015)
22. Higgins J., Altman D. G., Gøtzsche P. C., Jüni P., Moher D., Oxman A. D. *et al.* The Cochrane's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928.
23. Guyatt G. H., Oxman A. D., Schünemann H. J., Tugwell P., Knotterus A. GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol* 2010; **64**: 380–2.
24. Tarquinio C., Kivits J., Minary L., Coste J., Alla F. Evaluating complex interventions: perspectives and issues for health behaviour change interventions. *Psychol Health* 2015; **30**: 35–51.
25. Grant S. P., Mayo-Wilson E., Melendez-Torres G. J., Montgomery P. Reporting quality of social and psychological intervention trials: a systematic review of reporting guidelines and trial publications. *PLoS One* 2013; **8**: e65442.
26. Rice V. H., Hartmann-Boyce J., Stead L. F. Nursing interventions for smoking cessation. *Cochrane Database Syst Rev* 2013; **8**: CD001188.
27. Hillsdon M., Foster C., Thorogood M. Interventions for promoting physical activity. *Cochrane Database Syst Rev* 2005CD003180.
28. de Bruin M., McCambridge J., Prins J. M. Reducing the risk of bias in health behaviour change trials: improving trial design, reporting or bias assessment criteria? A review and case study. *Psychol Health* 2015; **30**: 8–34.
29. Peters S., Johnston V., Coppieters M. Interpreting systematic reviews: looking beyond the all too familiar conclusion. *J Hand Ther* 2014; **27**: 1–3.
30. Mahood Q., Eerd D., Irvin E. Searching for grey literature for systematic reviews: challenges and benefits. *Res Synth Meth* 2014; **5**: 221–34.
31. Dimeff L. A., Baer J. S., Kivlahan D. R., Marlatt G. A. Brief Alcohol Screening and Intervention for College Students. New York, NY: Guilford Press; 1999.
32. Barnett N., Apodaca T., Magill M., Colby S., Gwaltney C., Rohsenow D. *et al.* Moderators and mediators of two brief interventions for alcohol in the emergency department. *Addiction* 2010; **105**: 452–65.
33. Monti P., Barnett N., Colby S., Gwaltney C., Spirito A., Rohsenow D. *et al.* Motivational interviewing versus feedback only in emergency care for young adult problem drinking. *Addiction* 2007; **102**: 1234–43.

34. Davoli M., Amato L., Clark N., Farrell M., Hickman M., Hill S. *et al*. The role of Cochrane reviews in informing international guidelines: a case study of using the Grading of Recommendations, Assessment, Development and Evaluation system to develop World Health Organization guidelines for the psychosocially assisted pharmacological treatment of opioid dependence. *Addiction* 2014; **110**: 891–898.
35. Berkman N., Lohr K., Ansari M., McDonagh M., Balk E., Whitlock E. *et al*. Grading the strength of a body of evidence when assessing health care interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: an update. In: AHRQ, editor. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*, Chapter 15. AHRQ publication no. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014, pp. 314–49.
36. Katikireddi S., Egan M., Petticrew M. How do systematic reviews incorporate risk of bias assessments into the synthesis of evidence? A methodological study. *J Epidemiol Community Health* 2014 Online First PublicationDOI: 10.1136/jech-2014-204711.
37. Armstrong R., Waters E., Moore L., Riggs E., Cuervo L., Lumbiganon P. *et al.* Improving the reporting of public health intervention research: advancing TREND and CONSORT. *J Public Health* 2008; **30**: 103–9.
38. Neighbors C. J., Barnett N. P., Rohsenow D. J., Colby S. M., Monti P. M. Cost-effectiveness of a motivational lntervention for alcohol-involved youth in a hospital emergency department. *J Stud Alcohol Drugs* 2010; **71**: 384–94.
39. Cowell A. J., Brown J. M., Mills M. J., Bender R. H., Wedehase B. J. Cost-effectiveness analysis of motivational interviewing with feedback to reduce drinking among a sample of college students. *J Stud Alcohol Drugs* 2012; **73**: 226–37.
40. Sher K. J., Rutledge P. C. Heavy drinking across the transition to college: predicting first-semester heavy drinking from precollege variables. *Addict Behav* 2007; **32**: 819–35.
41. Smedslund G., Berg R. C., Hammerstrøm K. T., Steiro A., Leiknes K. A., Dahl H. M. *et al.* Motivational interviewing for substance abuse. *Cochrane Database Syst Rev* 2011; **5**: CD008063.
42. Chandler J., Churchill R., Higgins J., Lasserson T., Tovey D. Methodological standards for the conduct of new Cochrane Intervention Reviews, version 2.3: Cochrane; 2013. Available at: http://www.webcitation.org/6beErfCKl (accessed on 18 September 2015)
43. Meader N., King K., Llewellyn A., Norman G., Brown J., Rodgers M. *et al.* A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Syst Rev* 2014; **3**: 1–9.
44. de Bruin M. Risk of bias in randomised controlled trials of health behaviour change interventions: evidence, practices and challenges. *Psychol Health* 2015; **30**: 1–7.
45. Rehfuess E. A., Bruce N., Prüss-Üstün A. GRADE for the advancement of public health. *J Epidemiol Community Health* 2011; **65**: 559.
46. Gugiu P. C. Hierarchy of evidence and appraisal of limitations (HEAL) grading system. *Eval Program Plann* 2015; **48**: 149–59.
47. Eysenbach G. CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. *J Med Internet Res* 2011; **13**: e126.
48. Keller M. B., McCullough J. P., Klein D. N., Arnow B., Dunner D. L., Gelenberg A. J. *et al.* A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med* 2000; **342**: 1462–70.
49. Bellamy N., Carette S., Ford P. M., Kean W. F., le Riche N. G., Lussier A. *et al.* Osteoarthritis antirheumatic drug trials. III. Setting the delta for clinical trials—results of a consensus development (Delphi) exercise. *J Rheumatol* 1992; **19**: 451–7.
50. McGlothlin A. E., Lewis R. J. Minimal clinically important difference: defining what really matters to patients. *JAMA* 2014; **312**: 1342–3.
51. Waters A. M., Smith C. T., Young B., Jones T. M. The CONSENSUS study: protocol for a mixed methods study to establish which outcomes should be included in a core outcome set for oropharyngeal cancer. *Trials* 2014; **15**: 168.
52. Burkhardt J. T., Schröter D. C., Magura S., Means S. N., Coryn C. L. An overview of evidence-based program registers (EBPRs) for behavioral health. *Eval Program Plann* 2015; **48**: 92–9.
53. Miller W. R., Wilbourne P. L., Hettema J. What works? A summary of alcohol treatment outcome research. In: Hester R. K., Miller W. R., editors. Handbook of Alcoholism Treatment Approaches: Effective Alternatives, 3rd edn. Boston, MA: Allyn & Bacon; 2003, pp. 13–63.
54. Davoli M., Amato L. Do Cochrane reviews provide useful information to guide policy and practice? The experience of the Cochrane Drugs and Alcohol Group. *Epidemiol Psychiatr Sci* 2011; **20**: 219–23.
55. Ray A., Kim S., White H., Larimer M., Mun E., Clarke N. *et al.* When less is more and more is less in brief motivational interventions: characteristics of intervention content and their associations with drinking outcomes. *Psychol Addict Behav* 2014; **28**: 1026–1040.
56. Magill M., Gaume J., Apodaca T., Walthers J., Mastroleo N., Borsari B. *et al.* The technical hypothesis of motivational interviewing: a meta-analysis of MI's key causal model. *J Consult Clin Psychol* 2014; **82**: 973–83.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site:

Online Supplement 1 References of studies identified in BASICS systematic review that meet Foxcroft et al's criteria but are not included.

