
Copyright:

This is the peer reviewed version of the following article: Schmidt C, Schulte B, Seo H, Kuhn S, O'Donnell A, Kriston L, Verthein U, Reimer J. Meta-analysis on the effectiveness of alcohol screening with brief interventions for patients in emergency care settings. *Addiction* 2016, 111(5), 783-794., which has been published in final form at http://dx.doi.org/10.1111/add.13263. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Date deposited:

18/01/2018

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License
Title:
Meta-analysis on the Effectiveness of Alcohol Screening with Brief Interventions for Patients in Emergency Care Settings

Authors and Affiliations:
Schmidt CS¹, Schulte B¹, Seo H¹, Kuhn S¹, O’Donnell A², Kriston L³, Verthein U¹, Reimer J¹

¹ Centre for Interdisciplinary Addiction Research of Hamburg University, University Medical Center Hamburg-Eppendorf
² Institute of Health & Society, Newcastle University
³ Department of Medical Psychology, University Medical Center Hamburg-Eppendorf

Running title:
Meta-analysis on alcohol brief interventions in emergency care

Word count (excluding abstract, references, tables and figures): 3,737

Declaration of Interest:
This work was supported by the European Union as part of the BISTAIRS (Brief interventions in the treatment of alcohol use disorders in relevant settings) research project (Agreement number 2011_1204).

All authors declare that they have no conflicts of interest.
ABSTRACT

Background and Aims

Despite ambiguous evidence for the effectiveness of alcohol screening with brief interventions (BI) in emergency departments (ED), ambition for their widespread implementation continues to grow. To clarify whether such an application of BI is justifiable, we conducted a systematic review and meta-analysis on studies testing the impact of BI on alcohol consumption.

Methods

We included peer-reviewed, randomized controlled studies investigating the effects of BI on alcohol consumption in injured and/or intoxicated patients, published January 2002 - September 2015. Changes from baseline in consumption quantity, intensity and number of heavy drinking episodes were assessed at 3, 6 and 12 month follow-up, resulting in 9 separate random-effects meta-analyses of standardized mean differences (SMD). Moderation effects of intervention mode, length, type of interventionist, intensity of control intervention and study quality were assessed using subgroup comparisons and meta-regression.

Results

We considered 33 publications (28 separate studies) including 14,456 patients. Six out of nine comparisons revealed small significant effects in favour of BI, with the highest SMD at 0.19 (95% CI: 0.08-0.31). No significant moderators could be identified, and statistical heterogeneity ($I^2$) was below 40%.

Conclusions

In a large meta-analysis of randomized controlled trials in emergency care settings, there was evidence for very small effects of brief interventions on alcohol consumption reductions. More intensive interventions showed no benefit over shorter approaches. Non-
face-to-face interventions appear comparably effective, but this finding remains tentative due to the low number of non-face-to-face studies.
BACKGROUND

Stimulated by the large and robust evidence base supporting the implementation of brief interventions (BI) for alcohol in primary health care (1-3), there is mounting interest in the role they might play in helping address alcohol-related harm in a wider range of delivery settings, such as emergency health care (4, 5). Emergency departments (EDs) or trauma centers may provide valuable opportunities for BI provision (6). More than a third of alcohol-attributable deaths result from unintentional injuries (7) and a systematic review in western countries found that approximately one out of seven patients showing up in emergency EDs report harmful drinking levels (8). There is also evidence that the “teachable moments” inherent in alcohol-related emergency admissions may result in a powerful motivation to change amongst patients (9). At the same time, previous studies have identified a range of barriers to BI implementation in these settings, with low training rates, limited knowledge and awareness, time pressure and lack of financial support, all found to impact negatively on the motivation and compliance of practitioners (10, 11). Such evidence questions the potential for more widespread implementation efforts in ED settings unless a clear benefit of BI is demonstrated.

Yet since the first comprehensive systematic review of Nilsen et al. (12), the evidence for the effectiveness of BI in ED settings remains inconclusive, partly due to the comparable drinking reductions often found in both BI and control groups. The systematic review and meta-analysis from Havard et al. (13), which included 11 studies on BI in ED, also found no clear benefit of BI on alcohol use and associated outcomes, although some positive impact in reducing alcohol-related injuries was shown. Results from the considerable amount of studies that have accumulated since these early reviews, continue to pose more questions than they answer, in particular given the substantial clinical heterogeneity evident amongst studies in terms of type of BI, participants and outcome measures. Moreover, more recent reviews in this field have either lacked a specific ED focus in terms of setting (such as McQueen et al. (14), which assessed general hospital inpatient environments), or have restricted their inclusion criteria to specific populations or intervention modalities. For
example, two recent reviews focused on ED studies employing motivational interviewing (MI) in adolescents and young adults only (15, 16). Although the decision for such “splitting” meta-analyses is a legitimate one, depending on the question asked by the review (17), there are several good reasons why a broader “lumping” approach should be preferred. For example, “lumping” a larger number of studies, from a wider variety of clinical settings, can boost both the power and relevance of the findings (18). Importantly, as more selective analyses provide only partial evidence for the effectiveness of BI in ED settings, it remains challenging to develop robust recommendations for clinical practice.

At the same time, there have been a number of important developments in the field in recent years, including the publication of findings from two large studies of BI in ED (19-23), alongside the potential offered by rapid progress in computer, internet and mobile technology. In addition, there is emerging evidence from research in primary health care settings to suggest that short, simple interventions may be as effective as longer, more complex BI (24), which might also constitute a possibility for such time-limited settings as EDs (25).

As such, we felt that an updated systematic review and meta-analysis of the BI literature in emergency settings was timely. Our primary aim was to provide a comprehensive and up to date overview of the impact of BI in ED on alcohol consumption. In addition, to address both the need to systematize the highly heterogeneous data, and to account for recent developments in computer-based BI approaches, we also sought to assess the impact of potential moderators on BI effectiveness in this setting.

Our objectives were:

1) To obtain a comprehensive and up to date estimate of BI effects on alcohol consumption in emergency care settings.

2) To investigate the impact of intervention type (face-to-face versus computer-based), intervention length, type of professionals involved, study quality (based on the Cochrane risk of bias definitions (26)), and intensity of control intervention on BI effectiveness.
3) To conduct exploratory analyses on further potential moderators of effect, including attrition rate, timing of BI in relation to screening, country and site characteristics.

4) To investigate the role of an “assessment reactivity” effect (i.e. whether drinking reductions can be attributed to the study assessment, being an intervention in itself).

This paper presents the main results from our review and meta-analysis.

METHODS

The meta-analysis is based on the studies selected for the narrative systematic review originally conducted in 2013 within the European BISTAIRS project (Brief Interventions in the Treatment of Alcohol Use Disorders in Relevant Settings), (27), and updated in September 2015. Randomized controlled trials (RCTs, including cluster randomization or randomization by time sequence) were included, consisting of one to four BI sessions, each lasting 5 to 30 minutes (or 40 min, if only one session was provided). All studies, including intoxicated and/or injured patients over all age groups were considered, except for studies which focused solely on adolescents (i.e. excluded patients > 17 years). Studies published between January 2002 and August 2015 and written in English language were included. Grey literature was searched by screening websites of relevant alcohol and brief intervention organizations and networks (e.g. INEBRIA) and reference lists of included articles were also considered (27). A summary of extracted data is listed in Table S1.

For this meta-analysis, one reviewer extracted quantitative data under the supervision of a second reviewer, and effect sizes were calculated by a team of two reviewers. All categorizations and quantitative analyses, including the assessment of moderators and exploratory investigations, were performed exclusively for this study and were not part of the BISTAIRS review (27).

Outcomes
The primary outcome measure was change in alcohol consumption from baseline to 3-, 6- or 12-month follow-up. Data was categorized into measures of quantity (including drinks/units per week or month), intensity (drinks per drinking day/occasion) and number of heavy drinking episodes (“binge drinking”). Studies with insufficient quantitative data on these outcome measures were excluded.

**Statistical analysis**

Computation of effect sizes out of the available data was facilitated using a set of calculators provided by the Campbell Collaboration (28). The effect measure was the standardized mean difference (SMD) between intervention and control groups for change from baseline. We assumed to have greater precision by comparing change scores because baseline drinking levels of BI and control groups were not always comparable. Missing parameters for effect size calculation were computed or imputed from available information according to the guidance of the Cochrane and Campbell Collaborations (26, 28). If the standard deviation of the change score was not reported or calculable it was estimated from the baseline and follow-up standard deviations (26, Section 16.1.3.2). Their intercorrelations were assumed being 0.5 following the procedure in the meta-analysis of Jonas et al. (29). If a trial included more than one intervention and/or control group, sample sizes, mean scores and standard deviations of the associated groups were pooled.

Using the program Review Manager (RevMan) provided by the Cochrane Collaboration (30), we employed random-effects meta-analyses according to DerSimonian and Laird (31). Heterogeneity was assessed with the Cochran’s Q test and quantified with the I²-value (32). Random-effects models were used because real-world data in social sciences are likely to have variable population parameters and the effect size is expected to vary from study to study causing heterogeneity (17, 26, 31, 33).

**Risk of bias**
Study quality was assessed using the Cochrane risk of bias tool (26), evaluating the adequacy of random sequence generation (selection bias), allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment, handling of incomplete outcome data (attrition bias), and selective reporting (reporting bias).

Publication bias was assessed by visual inspection of funnel plots (34).

Moderator analyses

Moderator analyses were conducted to investigate potential influences of “intervention type”, “intervention length”, “interventionist”, study quality and type/intensity of control intervention. “Intervention type” compares face-to-face against non face-to-face BI (e.g. computer- or text-message-based interventions, or printed feedback). Intervention length was dichotomized, comparing brief (< 15 min) vs. extended interventions (≥ 15 min, including multiple contacts), which is compatible with the categories defined in the meta-analysis of Jonas et al. (29). The variable “interventionist” also consisted of two categories, “internal” referring to the use of ED staff (e.g. trained nurses) who performed BI, vs. “external” professionals, e.g. research team members. For study quality, two categories of high and low quality were applied. The category of high quality included studies which fulfilled each defined criterion for low risk of bias. The studies showing at least one unclear or unfulfilled criterion were classified as low quality. The variable “type/intensity of control intervention” comprises the categories “treatment as usual” (standard care or no specific description), “leaflet only” (e.g. alcohol information leaflet), “specific intervention” (e.g. unspecific, empathic advice), and “brief advice” (personalized feedback on drinking levels and face-to-face contact). For each potential moderator, we performed subgroup analysis using RevMan (30), and also random-effects meta-regression (employing a generalized least squares estimate), according to the methods described in Field & Gillet (33) and using the
SPSS syntax files and macros supplied with their article. We did separate meta-regressions for each follow-up measurement point and for each moderator.

**Further exploratory analyses**

In addition, we performed exploratory subgroup analyses for the quantity outcome or the 12-month follow-up (the analyses with the highest number of studies included), investigating associations with attrition rate (up to 75% vs. above 75% retention, determined by median split), timing of BI (after screening/in ED vs. separate appointment), site characteristics (Level 1 trauma center vs. other EDs) and country (USA vs. rest of the world). We also screened our included studies if they systematically investigated “assessment reactivity” by employing additional, “no-assessment” or “minimal-assessment” control groups, to summarize their results narratively.

**RESULTS**

**Characteristics of the included studies**

Thirty-three publications from 28 different RCTs were identified (see Figure 1), including 14,456 patients aged 13 years or older. The majority of the studies included individuals over 18, without setting an upper age limit. Five studies included also adolescents under 18 years (35-39). All studies included both genders, except for one investigating only men (40).

The majority of the 28 included studies employed face-to-face interventions, with six investigating non-face to face interventions (36, 41-45). Non-face-to-face interventions comprised an interactive computer program (36), printed computer-generated feedback (41, 43), leaflets (44), or mobile phone text messages (42, 45). Among the face-to-face interventions, 8 studies (9 publications) were categorized as “brief” BI (22, 23, 25, 46-51) and 14 studies (18 publications) as extended BI (19-21, 35, 37-40, 52-61). “Brief” BI ranged
between 5 and 10 minutes (median 5), and typically consisted of individual feedback with brief advice, or a brief motivational interview. “Extended” interventions ranged between 15 and 40 minutes (median 30) with a stronger focus on motivational elements. Eight employed one “booster” session (35, 37, 56-61) of 5-30 minutes (median 15) duration, between 2 weeks and 3 months after discharge. BI was typically conducted directly after assessment, before patients left the ED. Only 7 publications reported that BI was scheduled on separate appointments (38, 40, 54-58). In 12 of the 20 face-to-face BI studies, external interventionists were employed (19-21, 35, 37-39, 46-48, 56-59, 61), mostly research staff with postgraduate qualification in clinical psychology or counselling. The 9 studies (11 publications) classified as using “internal” interventionists either reported to use ED personnel (22, 23, 40, 49-53, 60) or trained nurses (54, 55). One study was excluded from this comparison (25) because both internal and external staff were used. Results of the risk of bias assessment are shown in Table S2. We categorized 19 publications with low risk of bias (19-21, 25, 35, 36, 38, 42, 44, 46, 48, 50-55, 59, 61) and 14 publications with high/unclear risk of bias (22, 23, 37, 39-41, 43, 45, 47, 49, 56-58, 60). Control interventions were grouped in four categories: “treatment as usual” or “no intervention” was the largest category including 15 publications (36, 40, 41, 44-46, 48-50, 52, 53, 56, 57, 60, 61). Provision of general information leaflets served as the control intervention in 10 publications (19-23, 25, 35, 38, 54, 55). The category “specific intervention” (37, 39, 42, 43, 47, 51, 58) comprises either handout plus brief counsellor contact (e.g. unspecific, empathic advice), weekly text message reminders, or personalized feedback. One study employed a short form of BI (brief advice) as control intervention (59). Retention rates ranged between 38% and 89.5%, with a median of 75%. Around two thirds of all studies were conducted in the US, mostly in level 1 trauma centres. More detailed descriptions of study samples, interventions and outcomes are provided in Table S2. Length of stay, intoxication levels, or injury type and severity were not systematically reported and were therefore not amenable to analysis.

Quantity
Twenty-two studies enrolling 12,613 participants at entry presented data on mean alcohol consumption per week and month at 3-, 6- and/or 12-month follow-up. Comparison at 12-month follow-up found a small but significant superiority of the BI conditions, indicating that participants receiving BI reduced their alcohol use significantly more than those in control groups (Figure 2).

**Intensity**

Fourteen studies enrolling 8,507 participants at entry presented data on mean alcohol consumption per day or occasion at 3-, 6- or 12-month follow-up. As shown in Figure 3, all comparisons indicate small but significant superiority of the BI conditions, with the highest SMD at 3-month follow-up.

**Number of heavy drinking episodes**

Eighteen studies enrolling 7,895 participants at entry presented data on number of binge drinking occasions at 3-, 6- or 12-month follow-up. As shown in Figure 4, the comparisons for 6-month and 12-month follow-up indicate slightly higher reductions in the BI condition.

**Publication bias**

Funnel plots were investigated for each outcome to detect possible publication bias (see Appendix S4). Although, in most comparisons, the number of studies was low for a reliable assessment, some plots suggested that mild to moderate publication bias might be present. An example is given in Figure 5, presenting a slightly asymmetrical funnel plot for the analysis with the most studies included.

**Statistical heterogeneity**
In 5 out of 9 comparisons, the $I^2$-value was 5% or lower (see figures 2-4). In the remaining comparisons, moderate heterogeneity was observed ($I^2$ between 21% and 37%), but without reaching statistical significance.

**Analyses of moderators**

For none of the five assessed moderators ("intervention type", "intervention length", "interventionist" study quality, type/intensity of control intervention) influenced the pooled outcomes statistically significantly, either in meta-regressions or in the chi-squared tests for subgroup differences. Figures 6-8 are representative examples illustrating the lack of systematic differences. For the remaining forest plots, see Appendix S3.

**Further exploratory analyses**

We found no evidence on systematic differences depending on attrition rate, timing of BI, site or country (see Appendix S3).

To investigate potential screening reactivity effects, five of the 28 studies additionally employed not assessed or minimally assessed control groups (35, 48, 50, 52, 61). In all five studies, these control groups showed consumption reductions comparable to the fully assessed controls.

**DISCUSSION**

This meta-analysis presents an examination of the literature on the effectiveness of BI in ED settings, together with an assessment of potential moderators. To date, this represents the most comprehensive meta-analysis in this field, based on 33 publications covering 28 individual studies. Six out of 9 meta-analyses, comparing change from baseline score differences between BI and control conditions, present significant results favoring BI. However effect sizes are small, with the highest SMD amounting to 0.19.
Statistical heterogeneity was not significant and of moderate size ($I^2$ below 40%), suggesting that the findings are robust across different populations, intervention types, implementation modes, settings and outcomes. There was little variation between the different types of drinking outcomes and follow-up measurement points, and none of the potential moderators appeared to influence intervention effects. In this context, it needs to be emphasized that this meta-analysis sought to assess sources of variability between studies (e.g. intervention characteristics), not within studies. Clinical population heterogeneity within studies (e.g. risky vs. dependent drinkers, self-harm vs. motor vehicle accident, high vs. low motivation to change, etc.) is likely to impact BI effectiveness (62), but is challenging to assess within meta-analyses. For such an attempt, effect sizes would need to be reported separately for different patient groups (which is rarely the case), and these groups would need to be comparable across studies. The qualitative review of Field et al. (62) suggests important potential moderators in the ED setting (alcohol severity, readiness to change, attribution of the injury to alcohol), but came to inconclusive results with regard to BI effectiveness.

Examining the results with regard to follow-up measurement times, summary effects on quantity tended to be higher in studies with larger follow-up intervals. In contrast, SMDs for drinks per day/occasion were lowest at 12-month follow-up. However, no definite conclusions can be drawn on longitudinal trends nor on the longevity of the (still small) intervention effects, because most studies did not assess data on all three follow-up points.

The lack of any moderation effect suggests that for ED settings there is no clear superiority of either longer or shorter BI approaches, face-to-face vs. no-face-to-face interventions, nor the employment of external counselors vs. ED staff. The impression that BI in this setting leads to only marginally higher drinking reductions, compared to control groups, seems to be comparable for all intervention modalities and intensities. Evidence also suggests that impacts of different intensities of control interventions are also comparable, however conclusions are limited as “treatment as usual” is rarely well defined and might comprise very different conditions, including advice on drinking or provision of (referral)
information. Contextual factors might also influence control condition outcomes, such as site characteristics (e.g. inpatient trauma wards vs. ED), as differences in length of stay, injury severity, available time to deliver BI, and the extent of already established mandatory alcohol interventions (such as in the US). However, these characteristics were not systematically reported in the included studies and could therefore not be included in moderator analyses. Previously, mainly conducted in primary care, have found that assessment reactivity, whereby the act of screening a patient for an alcohol use disorder stimulates a desire to address their drinking, may also account for some BI effect (63). Although not examined quantitatively, our findings did not suggest this to be an important factor for control group improvements in ED settings: none of the four studies that had employed minimally or not assessed controls could find evidence for such an effect. As already stated by Daeppen et al. (48), the observed decreases in control groups should not be attributed to effects of the assessment, and might speak more in favour that the reason for ED admission itself (e.g. accident) might play a major role in changing drinking behaviours. Compared to previous systematic reviews in this field, the number of studies included in the present meta-analytic review was much higher. However, the overall small effect sizes we found are in line with the ambiguous and inconclusive results of earlier work in this setting (12, 13, 15, 16), with changes in alcohol consumption either only slightly or not significantly different between intervention and control groups.

Our analyses of moderators could not determine a superiority for longer vs. shorter interventions, a finding which echoes previous work in primary health care and ED settings (24, 25). However, it contradicts the observation of Merz et al (16) that MI studies employing additional sessions tended to be more effective, and the meta-analysis of Jonas et al. for BI in primary health care, which suggested a better impact for brief multisession interventions (29). A number of factors may contribute to this trend. For example, intervention length is not necessarily linked to content or intervention fidelity, and especially in the time-limited nature of ED settings, psychological or motivational-based interventions might not be delivered with the same quality as in specialized addiction treatment or counselling settings. However, this
issue remains speculative, as literature is lacking systematic comparisons of BI fidelity across settings.

At the same time, one should bear in mind that the effectiveness of BI is not restricted to drinking reduction outcomes, but can also comprise other dimensions, such as reduction of alcohol-related injuries or driving-under-the-influence. This meta-analysis sought to focus on changes in alcohol consumption, but we acknowledge that this could also be seen as a limitation of this study. However, a meta-analysis of other broader outcomes (such as injuries, alcohol-related problems, readiness to change, driving behavior, etc.) remains challenging due to the heterogeneous and limited reporting of such outcomes in the existing literature.

Other limitations of this study are related to our chosen methodology. We decided to calculate meta-analyses using mean change scores from baseline to follow-up time. The advantage of changes scores is the enhanced interpretation of the effectiveness over time, compared to analysis on follow-up data alone, especially in case of baseline differences between intervention and control groups. However, a disadvantage of using change scores is the need to impute some parameters (e.g. the standard deviation of the change), as most studies did not deliver this data. Though the procedure proposed by Jonas et al. (29, chapter 3.5) aims to control for risks of artifacts, we cannot exclude the possibility that some effect sizes might have been different, if all data would have been available.

Moreover, some of the reported effects might be somewhat overestimated due to the presence of a low to moderate publication bias, which can be assumed in some of our funnel plots, especially because we did not find any grey literature meeting our inclusion criteria (e.g. the RCT criterion). However, confidence in these findings is limited due to the low number of studies in some comparisons.

**Implications for Research and Practice**

Although results generally favour BI, most summary effects are very small. It may be necessary to consider a more cautious approach to widespread implementation of BI in ED
settings, especially with regard to longer interventions, given limited time and financial resources. However, ED settings might encourage brief, and/or computer-assisted approaches, since moderator analyses point to a comparable effectiveness of these modalities. This could not only present a more cost- and time-effective approach, but may also be more appropriate for the fast-paced and stressful nature of ED settings. There is a need for systematic cost-effectiveness analyses of such brief approaches compared to current practice. Moreover, the small number of non-face-to-face BI studies included in the present meta-analysis calls for further research in this area. The impact of clinical population heterogeneity on BI effectiveness is another important issue, to determine which patients benefit most from BI. In addition, further research should focus on consistently observed control group improvements. If not due to assessment reactivity (as suggested by our results and other findings), these improvements might be attributed to an alcohol-related accident or the ED admission itself, the potential extent of reporting (social desirability) bias, or maybe the elements and active ingredients that ED personnel might already have implemented within their standard treatment of risky drinking patients.

Overall, our study is in line with the observation that BI delivered in ED settings might be less effective than BI in PHC. This could be due to various factors: the fast-paced and transient nature of ED settings may inhibit the formation of an efficient and trustworthy working cooperation between patient and clinician, not least as ED attendance is a critical situation in which stress, as well as alcohol intoxication may limit a patient’s receptiveness and motivation (43, 62). Moreover, there is generally greater acceptance for the provision of preventive lifestyle interventions (e.g. for diets and smoking) in PHC settings than in ED (25).

In conclusion, decision makers in ED settings should carefully weigh costs and benefits related to BI implementation, and might consider prioritising the implementation of non-face-to-face BI, or very short forms of BI delivery (e.g. screening with feedback only) in routine practice.
Figure 1: Study selection process
Figure 2: Comparison of change score differences between BI and control groups: Quantity (drinks per week/month)
Figure 3: Comparison of change score differences between BI and control groups: Intensity (drinks per drinking day/occasion)
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blow 2006</td>
<td>0.038</td>
<td>0.08</td>
<td>22.0%</td>
<td>-0.08</td>
<td>0.08</td>
<td>10.1%</td>
</tr>
<tr>
<td>Male 2005</td>
<td>0.093</td>
<td>0.08</td>
<td>24.8%</td>
<td>0.04</td>
<td>0.04</td>
<td>13.2%</td>
</tr>
<tr>
<td>Mello 2006</td>
<td>0.052</td>
<td>0.12</td>
<td>12.6%</td>
<td>0.03</td>
<td>0.03</td>
<td>21.6%</td>
</tr>
<tr>
<td>Segatto 2011</td>
<td>0.093</td>
<td>0.16</td>
<td>7.0%</td>
<td>0.10</td>
<td>0.22</td>
<td>22.0%</td>
</tr>
<tr>
<td>Spirito 2004</td>
<td>0.085</td>
<td>0.17</td>
<td>6.9%</td>
<td>0.08</td>
<td>0.27</td>
<td>4.4%</td>
</tr>
<tr>
<td>Suffoletto 2012</td>
<td>0.495</td>
<td>0.39</td>
<td>5.9%</td>
<td>0.56</td>
<td>0.37</td>
<td>1.2%</td>
</tr>
<tr>
<td>Suffoletto 2014</td>
<td>0.294</td>
<td>0.10</td>
<td>17.8%</td>
<td>0.26</td>
<td>0.05</td>
<td>4.5%</td>
</tr>
<tr>
<td>Wang 2010</td>
<td>0.039</td>
<td>0.15</td>
<td>8.8%</td>
<td>0.04</td>
<td>0.25</td>
<td>3.2%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>100.0%</td>
<td>0.06</td>
<td>0.03</td>
<td>0.15%</td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.00; \chi^2 = 7.35, df = 7 \ (P = 0.39); P = 5\%$
Test for overall effect: $Z = 1.38 \ (P = 0.17)$

1.5.4 Follow up: 4 to 6 months
- Apodaca 2007: 0.067, 0.084, 1.6%, 0.87 [0.11, 1.62]
- Donohue 2009: 0.130, 0.082, 17.7%, 0.13 [0.05, 0.31]
- Donohue 2012: 0.110, 0.130, 11.3%, 0.11 [0.14, 0.20]
- Field 2007, 2010, 2010: 0.091, 0.065, 24.7%, 0.08 [0.04, 0.22]
- Monti 2007: 0.458, 0.157, 8.2%, 0.46 [0.15, 0.77]
- Shorty 2011: 0.225, 0.183, 6.5%, 0.23 [0.13, 0.58]
- Soderstrom 2007: 0.049, 0.089, 18.4%, 0.05 [0.13, 0.23]
- Spirito 2004: 0.002, 0.179, 6.6%, 0.00 [0.44, 0.27]
- Thrinks 2010: 0.070, 0.208, 5.0%, 0.07 [0.34, 0.48]

Subtotal (95% CI): 100.0%, 0.13 [0.03, 0.23]

Heterogeneity: $\tau^2 = 0.01; \chi^2 = 10.79, df = 8 \ (P = 0.21); P = 26\%$
Test for overall effect: $Z = 2.64 \ (P = 0.009)$

1.5.5 Follow up: 12 months
- Blow 2006: 0.041, 0.08, 11.3%, 0.04 [0.14, 0.22]
- Donohue 2008: 0.073, 0.083, 10.5%, 0.08 [0.10, 0.25]
- Donohue 2012: 0.105, 0.120, 6.4%, 0.10 [0.13, 0.34]
- Daepen 2007: 0.095, 0.090, 11.3%, 0.10 [0.08, 0.27]
- Field 2009, 2010, 2010: 0.097, 0.066, 19.9%, 0.10 [0.04, 0.23]
- Male 2005: 0.038, 0.084, 13.0%, 0.03 [0.17, 0.17]
- Mello 2012: 0.053, 0.126, 5.7%, 0.05 [0.19, 0.30]
- Monti 2007: 0.465, 0.159, 3.6%, 0.47 [0.18, 0.76]
- Shorty 2011: -0.119, 0.102, 2.5%, -0.12 [-0.50, 0.25]
- Soderstrom 2007: -0.007, 0.089, 11.4%, -0.01 [-0.18, 0.17]
- Sommers 2006: 0.085, 0.237, 1.5%, 0.09 [0.38, 0.55]
- Spirito 2004: 0.178, 0.108, 2.9%, 0.18 [0.17, 0.53]

Subtotal (95% CI): 100.0%, 0.07 [0.01, 0.13]

Heterogeneity: $\tau^2 = 0.00; \chi^2 = 9.44, df = 11 \ (P = 0.58); P = 9\%$
Test for overall effect: $Z = 2.38 \ (P = 0.02)$

Figure 4: Comparison of change score differences between BI and control groups: Number of heavy drinking episodes.
Figure 5: Funnel plot for Comparison of change score differences between BI and control groups: Quantity (drinks per week/month), 12-months follow-up.
Figure 6: Forest plot for Number of binge drinking occasions, 3-months follow-up, grouped by intervention type.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight IV, Random, 95% CI Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith 2003</td>
<td>0.3869</td>
<td>0.1037</td>
<td>2.0%</td>
</tr>
<tr>
<td>Crawford 2004</td>
<td>0.1712</td>
<td>0.1023</td>
<td>6.3%</td>
</tr>
<tr>
<td>Spirito 2004</td>
<td>0.2693</td>
<td>0.1807</td>
<td>2.0%</td>
</tr>
<tr>
<td>Sommers 2006</td>
<td>0.0896</td>
<td>0.2377</td>
<td>1.2%</td>
</tr>
<tr>
<td>Soderstrom 2007</td>
<td>-0.0015</td>
<td>0.0997</td>
<td>8.1%</td>
</tr>
<tr>
<td>Murali 2007</td>
<td>0.4905</td>
<td>0.1631</td>
<td>2.8%</td>
</tr>
<tr>
<td>Field 2003, 2010, 2010</td>
<td>0.0562</td>
<td>0.0588</td>
<td>13.8%</td>
</tr>
<tr>
<td>Bernstein 2010</td>
<td>0.0049</td>
<td>0.0993</td>
<td>6.5%</td>
</tr>
<tr>
<td>Chehab 2010</td>
<td>0.2644</td>
<td>0.1215</td>
<td>3.5%</td>
</tr>
<tr>
<td>Sherry 2011</td>
<td>0.0375</td>
<td>0.1962</td>
<td>1.9%</td>
</tr>
<tr>
<td>Mello 2012</td>
<td>0.0206</td>
<td>0.1268</td>
<td>4.1%</td>
</tr>
<tr>
<td>Sommers 2013</td>
<td>0.1396</td>
<td>0.0343</td>
<td>3.4%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>0.13 [0.04, 0.21]</td>
<td>55.9%</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.01; Chi² = 14.77, df = 11 (P = 0.19), I² = 26%

Test for overall effect: Z = 2.00 (P = 0.05)

3.3.3 brief

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight IV, Random, 95% CI Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blow 2006</td>
<td>0.1720</td>
<td>0.0902</td>
<td>8.0%</td>
</tr>
<tr>
<td>Daleiden 2007</td>
<td>0.0704</td>
<td>0.0902</td>
<td>8.0%</td>
</tr>
<tr>
<td>Acedaniel ED 2010</td>
<td>0.0623</td>
<td>0.0993</td>
<td>7.0%</td>
</tr>
<tr>
<td>D'Orlando 2006</td>
<td>0.1032</td>
<td>0.0933</td>
<td>7.5%</td>
</tr>
<tr>
<td>D'Orlando 2012</td>
<td>0.1036</td>
<td>0.1201</td>
<td>4.8%</td>
</tr>
<tr>
<td>Drummond 2014</td>
<td>0.0049</td>
<td>0.0843</td>
<td>9.0%</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>0.08 [0.01, 0.16]</td>
<td>44.1%</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 2.00, df = 5 (P = 0.84), I² = 0%

Test for overall effect: Z = 2.14 (P = 0.03)

Total (95% CI) 100.0% 0.10 [0.05, 0.15]

Heterogeneity: Tau² = 0.00; Chi² = 17.21, df = 17 (P = 0.44), I² = 1%

Test for overall effect: Z = 3.88 (P = 0.001)

Test for subgroup differences: Chi² = 0.88, df = 1 (P = 0.40), I² = 0%\n
Figure 7: Forest plot for Quantity, 12-months follow-up, grouped by intervention length.
### 5.3.2 Tail-assessment only

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith 2003</td>
<td>0.3963</td>
<td>0.1037</td>
<td>1.9%</td>
<td>0.40 [0.04, 0.76]</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Mcke 2005</td>
<td>0.0237</td>
<td>0.0941</td>
<td>8.5%</td>
<td>0.02 [-0.14, 0.19]</td>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>Simmers 2006</td>
<td>0.0838</td>
<td>0.2377</td>
<td>1.1%</td>
<td>0.08 [-0.38, 0.55]</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>Haggard 2007</td>
<td>0.0734</td>
<td>0.0932</td>
<td>7.4%</td>
<td>0.07 [-0.11, 0.26]</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>Cheethamb 2010</td>
<td>0.2044</td>
<td>0.1315</td>
<td>3.5%</td>
<td>0.26 [-0.01, 0.52]</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Mello 2012</td>
<td>0.0206</td>
<td>0.1288</td>
<td>3.7%</td>
<td>0.02 [-0.23, 0.27]</td>
<td>2012</td>
<td></td>
</tr>
<tr>
<td>D'Costa 2012</td>
<td>0.1038</td>
<td>0.1201</td>
<td>4.1%</td>
<td>0.10 [-0.13, 0.34]</td>
<td>2012</td>
<td></td>
</tr>
<tr>
<td>Simmers 2013</td>
<td>0.1396</td>
<td>0.1996</td>
<td>3.1%</td>
<td>0.00 [-0.27, 0.27]</td>
<td>2013</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>0.13</td>
<td>33.6%</td>
<td>0.09 [0.01, 0.17]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.00$, $Q = 5.95$, df = 7 ($P = 0.55$); $I^2 = 0$
Test for overall effect: $Z = 2.09$ ($P = 0.04$)

### 5.3.10 leaflet only

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crawford 2004</td>
<td>0.1712</td>
<td>0.1023</td>
<td>5.7%</td>
<td>0.17 [-0.03, 0.37]</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>Acatecman ED 2010</td>
<td>0.623</td>
<td>0.0868</td>
<td>8.4%</td>
<td>0.06 [-0.13, 0.26]</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>Field 2006, 2010, 2010</td>
<td>0.0532</td>
<td>0.0565</td>
<td>12.9%</td>
<td>0.06 [-0.08, 0.19]</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Bernstein 2010</td>
<td>0.1045</td>
<td>0.0981</td>
<td>8.2%</td>
<td>0.11 [-0.09, 0.30]</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Drummond 2014</td>
<td>0.0845</td>
<td>0.0845</td>
<td>8.3%</td>
<td>0.00 [-0.17, 0.17]</td>
<td>2014</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>0.3900</td>
<td>39.6%</td>
<td>0.07 [0.01, 0.15]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.00$, $Q = 1.87$, df = 4 ($P = 0.76$); $I^2 = 0$
Test for overall effect: $Z = 1.81$ ($P = 0.07$)

### 5.3.11 specific intervention

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spirit 2004</td>
<td>0.2160</td>
<td>0.1807</td>
<td>1.5%</td>
<td>0.20 [-0.06, 0.46]</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>Elbow 2006</td>
<td>0.1726</td>
<td>0.0902</td>
<td>7.4%</td>
<td>0.17 [-0.00, 0.35]</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>Monb 2008</td>
<td>0.4905</td>
<td>0.1609</td>
<td>2.5%</td>
<td>0.49 [-0.08, 0.56]</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>D'Costa 2010</td>
<td>0.0390</td>
<td>0.0808</td>
<td>8.6%</td>
<td>0.10 [-0.08, 0.19]</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Sherry 2011</td>
<td>-0.0325</td>
<td>0.1928</td>
<td>1.6%</td>
<td>-0.04 [-0.42, 0.34]</td>
<td>2011</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>0.19 [0.05, 0.33]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.01$, $Q = 8.12$, df = 4 ($P = 0.19$); $I^2 = 36$
Test for overall effect: $Z = 5.65$ ($P = 0.00$)

### 5.3.13 brief advice

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>SE</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanderson 2007</td>
<td>-0.0015</td>
<td>0.0037</td>
<td>7.4%</td>
<td>-0.00 [-0.10, 0.10]</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>0.18 [0.09, 0.37]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Not applicable
Test for overall effect: $Z = 0.02$ ($P = 0.99$)

Total (95% CI) 100.0% 0.09 [0.05, 0.14]

Heterogeneity: $I^2 = 0.00$, $Q = 17.96$, df = 18 ($P = 0.48$); $I^2 = 0$
Test for overall effect: $Z = 2.81$ ($P = 0.001$)
Test for sub-group differences: $Chi^2 = 3.27$, df = 3 ($P = 0.36$); $I^2 = 0.2$

---

Figure 8: Forest plot for Quantity, 12-months follow-up, grouped by type/intensity of control intervention.
REFERENCES


53. Cherpitel CJ, Moskalewicz J, Swiatkiewicz G, Ye Y, Bond J. Screening, Brief Intervention, and Referral to Treatment (SBIRT) in a Polish Emergency Department: Three-


61. Sommers MS, Lyons MS, Fargo JD, Sommers BD, McDonald CC, Shope JT, et al. Emergency Department-Based Brief Intervention to Reduce Risky Driving and
