The pursuit of evidence-based practice: Comparisons of three guidelines on psychosocial interventions for alcohol problems

ANDERS BERGMARK & LISA SKOGENS & NINIVE VON GREIFF

AIMS – In this article we scrutinise three prominent guidelines on psychosocial interventions for alcohol problems. We pay special attention to how congruent the guidelines are in terms of the interventions recommended, and the processes used in order to identify and rank the "evidence" underpinning these recommendations. DATA – The analysed guidelines are: 1) Practice Guideline for the Treatment of Patients with Substance Use Disorders, American Psychiatric Association (2006); 2) Alcohol-Use Disorders. The NICE Guideline on Diagnosis, Assessment and Management of Harmful Drinking and Alcohol Dependence (2011), National Collaborating Centre for Mental Health, UK; 3) Guidelines for the Treatment of Alcohol Problems, Australian Government, Department of Health and Ageing (2009). The purpose is not to review the three guidelines as such, but to study them as an example of the production of evidence. All report to be based on “best available evidence”, so the guidelines were compared both regarding the actual recommendations and the guideline production procedures and differences in these procedures with related consequences.

RESULTS & CONCLUSIONS – Prestigious organisations in different national contexts have reached divergent conclusions about evidence-based practice and the quality of the scientific studies underpinning these conclusions. Differences in the guidelines regarding interpretations, limitations and grading illustrate the difficulties with the dilemmas of sensitivity (to include factors that are significant for how a psychological intervention is to be judged) and specificity (that irrelevant studies are cleared off) in the recommendations presented.

KEYWORDS – national guidelines, alcohol, evidence, production of evidence

Submitted 25.11.2013 Final version accepted 15.04.2014

Introduction

Since the end of the 1990s, proposals on the need to establish an evidence-based practice (EBP) have constituted a growing international trend in several professional fields. The demand for a scientific base for practice is most elaborated in the medical field, but is also apparent for psychosocial, non-medical interventions.

At an overall level, the development of the EBP movement has produced two radically different tendencies; on the one hand, we see international organisations (such as the Cochrane Collaboration, Campbell Collaboration, AGREE Collaboration, CONSORT Group, and GRADE working group) continuously reinforcing

Acknowledgement

This study was funded by The Swedish Research Council.
and elaborating their methodological specifications and demands for standardisation through systematic reviews and other procedures that enhance the production of unbiased evidence. These organisations aim at establishing globally valid knowledge. On the other hand, it is quite clear that many national and regional organisations seeking to establish EBP have by and large developed their regimes without consideration of important international developments on standardisation and synthesisisation. The difference between the radical strive for standardisation and control of bias that characterises the international organisations active in the EBP field and more local regimes on different levels are likely to produce different interpretations of what de facto constitutes evidence, not only between different international organisations and national/local bodies, but also between different national or local organisations. Such differences are the main focus in this article partly because they can be seen as a threat to the whole idea of establishing valid evidence on the effects of specific psychosocial interventions. We are well aware of the fact that there are other factors than the actual processing of evidence that shapes the form and content of different guidelines (guidelines are generally to some extent context-dependent). But in the following analysis we are mainly concerned with the different steps and procedures that identify and evaluate the evidence underpinning specific recommendations (i.e. which databases are used in order to identify original studies and systematic reviews, which studies and reviews are identified and selected as valid evidence, and what type of grading of evidence is used).

Within the field of evidence-based medicine (EBM), Timmermans and Mauk have observed that

So many parties have jumped on the EBM bandwagon and so many clinical practice guidelines are churned out by individuals, professional organisations, insurers and others that the benefits of uniformity may disappear in the cacophony of overlapping, conflicting and poorly constructed guidelines. With more than 1,000 guidelines created annually, calls for guidelines for clinical guidelines have been issued. (2005, p. 19)

Clinical guidelines are often defined as “systematically developed statements that assist clinicians and service users in making decisions about appropriate treatment for specific conditions” (Mann, 1996). It is also frequently pointed out that guidelines by definition are not mandatory (i.e. they are not rules), but they are also viewed as one of the most powerful tools for addressing the persistent problem of clinical practice variation. They can thus be seen as an attempt to standardise practice and therefore constitute a threat to professional autonomy.

In this article we scrutinise three prominent guidelines on psychosocial interventions for alcohol problems. As pointed out above, we pay special attention to how congruent the guidelines are in terms of the interventions recommended, and the processes used in order to identify and rank the “evidence” underpinning these recommendations. The analysed guidelines are:
• Practice Guideline for the Treatment of Patients with Substance Use Disorders, American Psychiatric Association (2006);
• Alcohol-Use Disorders. The NICE Guideline on Diagnosis, Assessment and Management of Harmful Drinking and Alcohol Dependence (2011), National Collaborating Centre for Mental Health, published by The British Psychological Society and The Royal College of Psychiatrists;
• Guidelines for the Treatment of Alcohol Problems, Australian Government, Department of Health and Ageing (2009).

The purpose of this article is not to review the three guidelines as such, but to study them as an example of the effort of establishing EBP. The guidelines serve to synthesise and critically appraise available evidence in practical recommendations for treatment providers. The three studied guidelines all claim to be based on “best available evidence”. It therefore seems reasonable to compare both the actual recommendations and the guideline production procedures and to examine the differences in these procedures with related consequences.

Data and methods

We have studied guidelines from the United States, the United Kingdom and Australia. All of these were the latest updated versions available in April 2013 and deal with the identification, assessment and management of alcohol dependence and harmful alcohol use/alcohol problems. The US guideline also includes treatment for other substance abuse than alcohol. All three guidelines address pharmacological and psychological/psychosocial treatment interventions.

These particular guidelines were chosen, for they were produced by significant institutions in the field. They are all English-language guidelines, which makes them available to the international scientific community as well as to practitioners and clinicians of alcohol treatment. While other guidelines could have been selected for analysis, we found that possible alternatives in general had more drawbacks than advantages: the guidelines issued by WHO have an international origin (WHO) and are therefore not a suitable object for the present analysis; the guidelines from ASAM (American Society of Addiction) and BAP (British Association for Psychopharmacology) are much less developed and are produced by organisations geared mainly towards guidelines for pharmacological rather than psychosocial interventions.

This study’s focus is on the introductory sections where the methods used to develop the guidelines are described and on the parts that concern psychosocial interventions for treating alcohol problems.

All three guidelines were also briefly overviewed as a whole to evaluate whether other parts should be examined more closely. Some examples of areas close to psychosocial treatment interventions, such as chapters on “experience of care” by clients/patients, relatives and carers, have been read more closely and will be discussed in relation to the main results.

After the first reading of the guidelines, we created a template to organise the first part of the data material, i.e. methods used to develop the guidelines. The development procedure of each guideline, the
composition of expert and other advisory
groups, search protocols, and grading and
evaluation systems of evidence and of in-
terventions were compiled within these
templates. We then compared similarities
and differences, and analysed possible
consequences of the recommendations in
the guidelines.

In a next step, we closely read the chap-
ters on psychosocial interventions and
summarised in matrices which interven-
tions were included in the guidelines,
how these interventions were evaluated
and graded, what scientific evidence the
evaluation was based on and how the evi-
dence was evaluated and graded. During
this process the matrices also included a
column for comments on the studies and
interventions included.

Finally, we chose two psychosocial in-
terventions for a more detailed analysis,
comparing and discussing the grading and
recommendation of these specific inter-
ventions.

**Results**

**Guideline procedures**

The guideline from the American Psychi-
tric Association (APA) – developed by
psychiatrists in active clinical practice or
in research or other academic endeavours
(p. 5) – was published in August 2006 and
covers several types of substance use dis-
orders. It consists of three parts: “Treat-
ment Recommendations for Patients with
Substance Use Disorders” (also published
as a supplement to the American Journal
of Psychiatry), “Background Information
and Review of Available Evidence” and
“Future Research Needs”. The first part
contains both general and specific recom-
definations on treatment of patients with
specific substance disorders. This part
also includes a general discussion of the
formulation and implementation of treat-
ment plans. The second part provides an
overview of substance use disorders with a
structured review and synthesis of the evi-
dence base for the recommendations made
in the first part. The third part, summarising
areas for future research needs, is not
used or analysed in this article.

The key features of the APA Guideline
development process are described as fol-
low:

- A comprehensive literature review to
  identify all relevant randomised clinical
trials as well as less rigorously designed
  clinical trials and case series when evi-
dence from randomised trials was una-
vailable.
- The development of evidence tables that
  summarise the key features of each iden-
tified study, including funding source,
study design, sample sizes, subject char-
acteristics, treatment characteristics,
and treatment outcomes.
- Initial drafting of the guideline by a
  work group that included psychiatrists
  with clinical and research expertise in
  substance use disorders.
- The production of multiple revised
drafts with widespread review (23 or-
ganisations and 70 individuals submit-
ted significant comments).
- Approval by the APA Assembly and
  Board of Trustees.
- Planned revisions at regular intervals.
  (American Psychiatric
  Association, 2006, p. 7)

As stated above, revisions are planned. So-called Guideline Watches are intended
to “summarize significant developments in practice since publication of an APA practice guideline...watches represent opinions of the authors and approval of the Executive Committee but not policy of the APA (Smith Connery & Kleber, 2007, p. 1). But when data was collected for this article in April 2013, only one Guideline Watch concerning treatment of nicotine dependence and treatment of alcohol dependence had been published (Smith Connery & Kleber, 2007).

The Australian guidelines focus, as the title shows, only on treatment of alcohol problems. The guideline was published in 2009 in two parts, as actual guidelines and as a review of the evidence. Treatment recommendations are included in both these documents, and the reader is encouraged to read the review, which provides more details and a full list of references on the evidence base for guideline recommendations (p. 1). Both documents have four named editors from the University of Sidney. Specific authors are responsible for each chapter in the actual guideline but not for the review. An Advisory Committee, mainly with senior researchers from several Australian universities, has given advice throughout the process of developing both documents. A Guideline Group, some members of which have also written chapters in both documents, has provided advice on content and format of both the guidelines and the review (Guidelines, p. IV). The guideline group includes researchers, clinicians and administrators from relevant institutions. The review covers the major interventions currently available for treating alcohol use disorders (p. 5) and includes screening, assessment and treatment planning.

When developing the review, focus has been on evidence that has emerged since a previous literature review (Shand et al., 2003). Where appropriate, material from the 2003 edition and its accompanying literature review has been included (Guidelines, p. 2).

The development process of the Australian guidelines is described as follows:

- Updating the review of the evidence for treatment of alcohol problems and published as a companion document.
- Consulting with an expert panel.
- Seeking feedback from clinicians concerning the previous edition (reported separately).

(Australian Government, Department of Health and Ageing, 2009, p. 2)

The UK guideline, published in 2011, tackles only alcohol problems. Commissioned by the National Institute for Health & Clinical Excellence (NICE), the UK guidelines were developed through the National Collaborating Centre for Mental Health (NCCMH). The NCCMH is a collaboration of professional organisations involved in the field of mental health, along with national service user and carer organisations, a number of academic institutions and NICE (p. 13). The Guideline Development Group that produced the guidelines included alcohol misusers and carers as well as professionals in psychiatry, clinical psychology, general practice, nursing and psychiatric pharmacy. Staff from the NCCMH provided leadership and support throughout the guideline development process. This process is described in six basic steps:

- Define the scope, which sets the param-
eters of the guideline and provides a focus and steer for the development work.

– Define review questions considered important for practitioners and service users.

– Develop criteria for evidence searching and search for evidence.

– Design validated protocols for systematic review and apply to evidence recovered by search.

– Synthesise and (meta-)analyse data retrieved, guided by the review questions and produce Grading of Recommendations: Assessment, Development and Evaluation (GRADE) evidence profiles and summaries.

– Answer review questions with evidence-based recommendations for clinical practice.

(NCCMH 2011, p. 41)

The systematic literature searches were in all restricted to clinical guidelines, health technology assessment reports, key systematic reviews and randomised controlled trials (RCTs). However, in the specific search protocols for psychological and psychosocial interventions, the searches are restricted to RCT studies only.

The UK guideline was initiated and produced in a national healthcare setting, with an implementation plan described in the introductory part:

local healthcare groups will be expected to produce a plan and identify resources for implementation alongside with appropriate timetables. Subsequently, a multidisciplinary group involving commissioners of healthcare, primary care and specialist mental health professionals, service users and carers should undertake the translation of the implementation plan into local protocols (p. 13–14).

This procedure could to some extent be seen as an alternative to the graded recommendations found in the other two guidelines. There is also a “short version” of the UK clinical guideline (2011) with clinical recommendations, but again, without a grading system like the ones used by APA and in the Australian guidelines.

The Australian guideline is national, commissioned by the Australian Government and produced by the National Drug and Alcohol Research Centre, but in contrast to the UK guideline, no implementation plan is described. Instead, the recommendations are graded on a scale directly related to the grading of the evidence.

In the US, there are no national frames organising treatment of alcohol problems. Instead there is a wide range of substance abuse programmes funded through different sources, where administration of policy and services takes place through state authorities (McCrady, 2012). There are thus no nationally initiated or produced guidelines but many organisations producing guidelines in their various areas (Bergmark & Lundström, 2011), such as the CSAT (Centre for Substance Abuse Treatment) that produces treatment improvement protocols. McCrady (2012) states that “the USA has no systematic approach to incorporating research findings into alcohol treatment, nor are there US-wide mandated standards of care” (p. 231), but points out APA as an important producer of practice guidelines. Thus, the US guideline selected for this study is the one produced by the APA, a well-es-
Table 1. Overview of the studied guidelines

| Country   | USA                                      | Australia                                      | United Kingdom                                      |
|-----------|------------------------------------------|------------------------------------------------|-----------------------------------------------------|
| Initiator / Producer | American Psychiatric Association / American Psychiatric Association | Australian Government; Department of Health and Aging / University of Sydney | National Institute for Health & Clinical Excellence (NICE) / National Collaborating Centre for Mental Health (NCCMH) |
| Scope     | Substance use disorders                  | Alcohol problems                               | Alcohol-use disorders                               |
| Structure | Part A: Treatment Recommendations        | Treatment of Alcohol Problems; Review and Guidelines are presented in separate parts | Presents recommendations and review in separate chapters within the same document. Own meta-analyses are presented in an appendix. |
| Databases used in searches | MEDLINE/PubMed, Cochrane, + informal search | MEDLINE, Cochrane, ISI Web of Knowledge, PsycINFO, EBM Reviews + reference lists and guidelines | CINAHL, EMBASE, MEDLINE, PsycINFO, Cochrane¹ |
| Grading of evidence/ Evidence-hierarchy | A: Double-blind randomised clinical trial | lb: Systematic review or meta-analysis of RCT | GRADE-system used for summarising and synthesising of evidence: Moderate Further research is likely to have an important impact on confidence in the estimate of the effect and is likely to change the estimate. |
| Recommendation structure | Three categories of endorsement: | Directly based on: | Recommendations given without grading² |
|           | [I] Recommended with substantial clinical confidence | A: Category I evidence | |
|           | [II] Recommended with moderate clinical confidence | B: Category II evidence or extrapolation from I | |
|           | [III] May be recommended on the basis of individual circumstances | C: Category III evidence or extrapolation from I or II | |
|           |                                           | D: Category IV evidence or extrapolation from I, II or III | |
|           |                                           | E: Standard of care (practical or ethical consensus) | |

¹ The databases listed refer to searches on psychosocial interventions. CINAHL (Cumulative Index to Nursing and Allied Health Literature), EMBASE (Biomedical Database), MEDLINE (the U.S. National Library of Medicine's bibliographic database), PsycINFO (Psychological Information Database).

² To be recommended in the NICE guidelines, an active psychological intervention should "show evidence of effectiveness against no treatment control/waitlist in the first instance, then against treatment as usual, and preferably should be more effective than other active interventions" (p. 344).

tablished non-governmental organisation with great influence and reputation.

Besides the organisational differences between the US on the one hand and the UK and Australia on the other, the APA guideline also differs in range, since it covers several forms of substance use disorders while the others only deal with alcohol. The APA recommendations are based on graded support from evidence. One interesting and possibly important difference between the evidence hierar-
chies used in the APA and the Australian guidelines is that meta-analysis and other systematic reviews are ranked as the highest form of evidence in the Australian guidelines, while this type of structured secondary data analysis is much lower ranked in the APA evidence grading.

The UK guideline is the most transparent and elaborated of the three. The development methods are described extensively and search protocols are disclosed for each section. For review questions about interventions, the PICO² frame was used as search protocol and meta-analyses were conducted when possible. The search protocols for psychological and psychosocial interventions reveal that only RCT studies were used as evidence base. The UK guideline is also stricter than the others in its inclusion criteria, since only studies on treatment-seeking populations were selected (p. 230). Studies on treatments addressing other problems than the actual drinking problem (such as anxiety and depression) were also excluded.

The APA and the Australian guidelines primarily use an evidence hierarchy where the evidence is valued with reference to its internal validity (Petticrew & Roberts, 2002; Hansen & Rieper, 2012). This type of evidence hierarchy is the most common methodological norm among evidence-producing organisations and has its origins in EBM (Sackett et al., 1997).

The ranking of studies in the guidelines differ. As mentioned above, the UK guideline uses only RCT studies, and the producers performed their own meta-analyses when evaluating psychological and psychosocial interventions. The Australian guideline producers did not make their own meta-analyses but include earlier meta-analyses in their review and give them, as well as other systematic reviews, the highest rank in their evidence hierarchy. In the APA guidelines, RCT studies are given the highest ranking while systematic reviews are number six (E) in the evidence hierarchy.

The strict evidence grading used otherwise is abandoned in the UK guidelines in the chapter covering care experiences of people who misuse alcohol and their families and carers. This part of the review focuses on qualitative research because it was “felt to be most appropriate to answer questions about experience of care” (p. 64). It is also stated that “Because good quality qualitative research exists within the literature, quantitative and survey studies were excluded” (p. 64). Thus, instead of an evidence hierarchy to evaluate the evidence, the producers of the UK guidelines here chose to use an evidence typology where the evaluating of the evidence is based on the study design’s appropriateness to answer the specific question (Petticrew & Roberts, 2002; Hansen & Rieper, 2012). When an evidence typology is used, study context and external validity is often taken into account for the valuating of evidence. There is no further discussion or motivation other than the above as to why all quantitative studies are excluded. It is somewhat unexpected that survey studies – a design often recommended as appropriate for investigating user satisfaction with service when the use of evidence typology is suggested (Petticrew & Roberts, 2002; Hansen & Rieper, 2012) – are excluded here without further motivation.

The guidelines differ somewhat also with regard to which databases were
searched. The APA guideline basically used only MEDLINE and Cochrane, complemented with less formal literature searches conducted by APA staff and individual members of the Work Group (p. 7). Only English-language papers were used.

In addition to MEDLINE and Cochrane, CINAHL, EMBASE and PsycINFO were searched for the UK guideline. Also, additional informal search methods were used (p. 48). It is stated that although no language restrictions were applied at the searching stage, foreign-language papers were not requested or reviewed “unless they were of particular importance to a review question” (p. 48).

The chapter describing the procedures of developing the UK guideline states that for some review questions it was necessary to prioritize the evidence with respect to the UK context (that is, external validity) (p. 48) in terms of participant factors, provider factors and cultural factors. Finally, the approach to unpublished evidence is described.

The databases searched for the Australian guidelines include MEDLINE, the Cochrane Database of Systematic Reviews, ISI Web of Knowledge, PsycINFO and Evidence Based Medicine Reviews (p. 5). Also, other more informal search methods are described (p. 5).

The issue of internal versus external validity is not explicitly discussed in the review that the Australian guidelines are based on. However, in the introduction to the actual guidelines it is noted that evidence which is about an area of practice of little clinical importance can invite a lower strength of recommendation, even if it were methodologically sound.

Which interventions are listed in the studied guidelines?
In the APA guidelines, the section about alcohol use disorders is divided into Management of Intoxication and Withdrawal, Pharmacological Treatments and Psychosocial Treatments. The psychosocial treatments presented as effective for some patients with an alcohol use disorder include MET (Motivational Enhancement Therapy) [I], 3 behavioural therapies [I], CBT (Cognitive Behavioural Therapy) [II], TSF (Twelve Step Facilitation) [I], group therapies [II], marital and family therapies [I] and psychodynamic therapy [III]. Patients are also recommended to participate in self-help groups, such as AA (Alcoholics Anonymous) [I]. Brief interventions, case management and self-guided therapies are also mentioned, but not graded.

Generally, few of the references are graded high and considered providing strong evidence. No subdivisions regarding what types of groups of comparisons are used when judging the interventions are presented, i.e. whether comparisons are made between intervention and control, tau or different interventions.

In a section called General Treatment Principles, it is pointed out that individuals with substance use disorders are heterogeneous, which implies differences in treatment conditions. These differences involve types and number of substances used, individuals’ genetic vulnerability for developing a substance use disorder, readiness for change and motivation to enter into treatment, severity of the disorder, and other physical and psychiatric problems, but also differences in social and cultural context.

When assessing patients, it is recom-
mended to use “a spectrum that includes use, misuse, abuse and dependence” (p. 15), where the latter two represent formal diagnostic categories. From the comments on the reviewed and graded studies it can be discerned that it is difficult to draw general conclusions from the results. In this context it is pointed out that differences between groups of clients/patients and combinations of interventions and differences in intensity can result in different outcomes.

In the UK guideline the interventions reviewed are based both on trials involving comparisons between the actual intervention and minimal intervention control and on trials comparing the actual intervention with some other active intervention. In some cases comparisons are also, or instead, made between different types of CBTs, for example. Grading of the evidence is done separately for the different comparisons and followed by comments on the results.

The guideline includes 13 interventions: MI, TSF, CBT, behavioural therapies, contingency management, social network and environment-based therapies, couples therapy, counselling, short-term psychodynamic therapy, multi-modal treatment, self-help treatment, psycho-educational interventions and mindfulness meditation. Generally, the evidence is graded as moderate. There are three exceptions: the evidence is graded high for studies showing no differences between TSF and other active intervention as well as between CBT vs. other active intervention. The evidence is graded low for studies of differences between multi-modal treatment and other active interventions. According to the comments on the evidence it is very difficult to draw general conclusions and judgments from the results of the studies included. For example, several studies show completely different results at different times for follow-up and significant differences disappear at longer follow-ups (see, for example, p. 243 and p. 250). Furthermore, different outcome variables show different results at different times.

In the Australian guidelines, the following interventions are presented (some of them are only mentioned with a short comment, while most are graded and commented): CBT (general), behavioural self-control, coping skills training, cue exposure, behavioural couples therapy, MI, contingency management, solutions-focused therapy, mindfulness-based stress reduction, psychodynamic therapy, narrative therapy, relapse prevention strategies, self-help programmes, SMART recovery, aftercare and long-term follow-up, brief interventions.

For the majority of these interventions the evidence is presented according to the type of studies included; reviews, randomised studies or “other” studies (another target group or another methodology). Compared to the UK guideline, the Australian guidelines are less strict regarding what studies are included, both in respect of target group and methodology. Some common treatment interventions are mentioned and discussed without giving/presenting a graded recommendation.

The recommendation for CBT-based interventions and MI is A (directly based on category [I] evidence) supported by meta-analysis, reviews and RCTs. There are some differences in how specific the recommendations are. Relapse prevention is, for example, described more as a strategy
Table 2. Overview of MI/MET recommendations

| Country | Recommendation | Grading of evidence | References |
|---------|----------------|---------------------|------------|
| USA     | [I] A          | A, A-, E, F, G      | Project MATCH 1998 |
|         |                |                     | Project MATCH 1997 |
|         |                |                     | Dunn et al. 2001 |
|         |                |                     | Miller & Wilbourne 2002 |
|         |                |                     | Project MATCH 1993 |
| UK      | No recommendations are given (see note 2) | Moderate (MI vs control) Moderate (MI vs other) | Hester 2005; Rosenblum 2005b; Sellman 2001 |
|         |                |                     | Davidson 2007; MATCH 1997; Sellman 2001; Shakeshaft 2002 |
|         |                |                     | Sobell 2002; UKATT 2005 |
| Australia | A              | la                   | Hettema et al. 2005; Vasilaki et al. 2006; Project MATCH 1997–8, a,b; Cutler and Fishbein 2007; John et al. 2003; UKATT 2005a, b; Monti et al. 2007; Beckham & Beckham 2007 |

3 “Control” refers to minimal intervention control and “other” refers to other active intervention.

than an intervention, but is recommended to all patients with addiction. MI, on the other hand, is described as a first line or stand-alone intervention but also as a complement to other treatment modalities.

Comparisons between the three guidelines in how they grade MI and TSF
In order to illustrate similarities and differences between the three guidelines, two interventions are presented and discussed in more detail. The interventions, MI and TSF, are presented in two separate matrices with information about the included studies/trials underlying grading of the evidence and, in cases where it is presented, recommendation of the intervention for each guideline. MI and TSF have been chosen because they are two of the most established psychosocial treatment interventions and also clearly demarcated in all three guidelines.

MI and its derived manual-guided MET aim to result in rapid internally-motivated changes by exploring and resolving ambivalence towards behaviour using motivational methods and strategies to utilise the patient’s resources.

MET is the longer-term follow-up to an initial brief intervention strategy (APA p. 40) and is given the highest recommendation [I], i.e., “recommended with substantial clinical confidence”. This intervention is claimed to be effective even for patients who are not highly motivated to change, which gives it a practical advantage over other therapies. However, this statement in the APA guideline is not followed by a reference. The claim of MET effectiveness regardless of grade of motivation is presented in a general part of the guideline. There are no descriptions as to what client groups with alcohol problems might benefit from MET. It is simply claimed that there is support for the use of MET as a stand-alone treatment.

In the NICE guidelines the quality of the evidence regarding MET is evaluated as moderate. MI is recommended to be used as part of the initial assessment for all people who misuse alcohol. However, it is claimed that there is not sufficient ev-
idence to recommend motivational techniques as a stand-alone intervention. The clinical evidence for this claim is based on results from three RCTs where MI was compared with minimal intervention control and on six RCT studies where MI was compared with other active interventions. According to the UK guidelines, there is no clinical evidence for MI being more effective than other interventions. Further research on comparisons between MI and tau and control groups is demanded. Motivational techniques are suggested as best suited to function as a complement to other psychosocial interventions, where working with motivation and encouraging positive change are described as key elements.

The Australian guidelines recommend MI as “first line” or “stand-alone” treatment, alternatively as a complement to other treatment. This recommendation is graded A, which equals “directly based on category [I] evidence; evidence obtained from a systematic review or meta-analysis of randomized controlled trials (Ia)”. Besides a description of what MI is and the recommendation, no further information about the treatment is given. The Review of Evidence which the recommendations are based on describe two meta-analyses and five RCT studies, while a summation of the evidence describes the efficiency of MI as varying between different practitioners of MI, different settings and different target groups.

The comparison above indicates relatively large variations both in the way MI/MET is judged in the three guidelines and how evidence is valued. The APA guideline refers to Project MATCH as support for MET as stand-alone treatment. Project MATCH is included in the NICE guidelines as one of six studies where MI-based treatments are compared with other interventions. It is concluded that there is not strong enough support for MI/MET as stand-alone treatment (p. 345). These different interpretations could perhaps be related to the broader base of evidence underlying the UK recommendations. However, it is more difficult to understand why the RCT studies that were already published when searches for the other two guidelines were conducted (such as Sellman, 2001; Sobell, 2002; Shakeshaft, 2002) and that were used in the UK guidelines, are not included in either the American or the Australian guidelines. Those studies are both within the frames of the evidence hierarchy (as RCT studies) and are searchable in the database used in all three guidelines (MEDLINE).

The fact that the other two references used in the APA guideline to support their claim that MI/MET are effectual (Dunn et al., 2001; Miller & Willbourne, 2002) are not included in the UK guideline can, however, be related to the variations in evidence hierarchy between the guidelines. Dunn et al. (2001) and Miller & Willbourne (2002) are review studies and thus completely excluded from the UK guideline, which uses only RCT studies as evidence in its protocol for psychosocial interventions. But also the APA and Australian guidelines, which include reviews in their evidence hierarchy, use these in different ways. The Mesa Grande study by Miller & Willbourne (2002) is graded low in the APA evidence hierarchy, but is frequently used in the part dealing with recommendations for treatment of alcohol-related disorders. In the Australian guidelines
reviews are graded high in the evidence hierarchy; yet Mesa Grande (Miller & Willsbourne 2002) is not used at all as support for MI/MET, although used in other parts of the guidelines, for example as support for CBT.¹

**Twelve step facilitation (TSF)** is based on the 12-step/AA concept that alcohol is a “spiritual” medical disease. The intervention aims to actively encourage commitment to and participation in AA meetings. This intervention is studied and evaluated in, for example, project MATCH (Babor & Del Boca, 2003) and is frequently used internationally.

Gradings of and comments on TSF are presented in Table 3. While the US and UK guidelines only highlight TSF, the Australian guidelines make a distinction between AA, 12-step and TSF. The US guideline includes many more studies than the other two and has a large spread of graded evidence. However, the overall recommendation of TSF is the highest [I]. In the comments on the evidence it is revealed that there are no randomised studies measuring treatment effects of AA. Other studies, however, show that AA, as well as 12-step and TSF, have positive outcomes, and several studies show that greater AA participation is associated with greater rates of abstinence from alcohol as well as with better drinking outcomes. It is therefore concluded that patients should be encouraged to attend at least several AA meetings. Further, it is argued that patients would benefit from homogenous AA groups in terms of gender and occupational status, etc., and that evidence from small-scale trials suggests that patients with greater severity of drinking problems benefit from a more affective rather than a

| Country | Recommendation | Grading of evidence | References |
|---------|----------------|---------------------|------------|
| USA     | [I]            | A                   | Carroll et al. 1998  
Project MATCH 1997, 1998; Crits-Christoph et al. 1999  
Timko et al. 2000  
Humphreys et al. 1999; Moos et al. 1999; Quimette et al. 1998; Cross et al. 1990  
Vaillant 1983  
McCready 1998; McKey et al. 1991; Moos et al. 1990; Hoffman & Miller 1992  
Humphreys et al. 2004; Miller & Hester 1986; Mercer & Woody 1999; Gilbert 1991; Nowinski et al. 1992; Alcoholics Anonymous 2002; Emrick & Tonigan 2004; McCready & Irvine 1989; Rychtarik et al. 2000 |
| UK      | No recommenda-  | High (TSF vs other) | Easton 2007; Falsstewart 2005; Falsstewart 2006; MATCH 1997; Waltzer 2009  
Timko 2007; Waltzer 2009 |
|         | tions are given (see note 2) | Moderate (compares different TSF) |
| Australia | B | II (AA only) | Tonigan & Connors 2008; Bond et al. 2003; Litt et al. 2007; Vaillant 2005; Timko et al. 2006; Timko & DeBenedetti 2007  
Weiss et al. 2005  
Moos & Timko 2008; Moos & Moos 2005,6, a, b  
Quimette et al. 1997; Project MATCH 1997; Ries et al. 2008  
Ferri et al. 2006 |
|         | A | I (AA + treatment) | |
|         |     | TSF (No rec.) | |
cognitive focus. When it comes to aftercare it is concluded that interventions with elements of TSF are more effective than those with CBT for patients without psychiatric disorders. For patients with psychiatric disorders the effects of TSF and CBT are equally good.

The UK guidelines divide the studies into those comparing TSF with other active interventions and those comparing different types of TSF. Grading and comments are presented separately. Studies comparing TSF with other active interventions show no significant differences and the quality of this evidence is graded as high. Comparisons between different types of TSF show that Directive TSF is more effective in maintaining abstinence at 12-month follow-up, but no differences are apparent in reducing heavy drinking episodes. Further, intensive TSF is significantly more effective than standard TSF in maintaining abstinence at 12-month follow-up. The quality of this evidence is graded as moderate. According to the UK guideline, there is a lack of evidence for the effectiveness of TSF compared with treatment as usual or control (p. 328), with the conclusion that TSF is therefore best seen as a component of any effective psychosocial intervention.

In the Australian guidelines TSF is presented under “Self-help programs”, which include AA only, AA associated with treatment, and TSF. Grading of the evidence for AA only is [II]. This, it is argued, has to do with the fact that just a few RCT studies exist. The highlighted studies indicate that the intervention – graded as B (second strongest recommendation) – offers a range of benefits. When it comes to AA associated with treatment and aftercare, the evidence is based on two studies given the highest rating of A. According to the Australian guidelines there is insufficient evidence to recommend TSF, but in the text it is revealed that studies can be found that show TSF to be effective.

Comparison between the guidelines shows that there are large differences regarding which studies are included and graded, but also in the perception of AA as an intervention or not and what conclusions – in terms of ranking – are drawn based on available evidence. Focusing only on TSF, significant differences emerge; the US guideline gives the highest recommendation, while neither the UK nor the Australian guidelines valuate the evidence as sufficient to recommend the intervention.

Discussion
Differences in the guidelines regarding interpretations, limitations and grading illustrate the difficulties with the dilemmas of sensitivity (to include factors that are significant for how a psychological intervention is to be judged) and specificity (that irrelevant studies are cleared off) in the recommendations presented. These dilemmas will be discussed below.

All three guidelines contain an introduction which describes and discusses the multidimensional, complex and context-bound nature of treatment of alcohol disorders. The APA guideline stresses that individuals with a substance use disorder are heterogeneous and regarding a number of clinically important features. The Australian guidelines state that the way individuals identify and handle their alcohol problems varies, that much still remains to understand and explain this, and that
professional treatment “at best can contribute to a person’s self-awareness” (p. 1). For its part, the UK guideline argues that evidence-based treatments are often delivered within an overall treatment context that supports the specific treatment intervention. It is stressed that this “service context” is frequently essential for effective interventions.

When the evidence is valued, one might expect that the generalisation of a study in a certain context would be thoroughly discussed after the stressing of the multidimensional, complex and context-bound character of treatment for alcohol and substance abuse in the introductions of all three guidelines. It is probably important to recognise under what circumstances a certain psychosocial treatment has been found to be evidence-based effective. On the other hand, it is difficult to evaluate the importance of the described specific circumstances. For example, if a study considers MI as effective for young patients in emergency care, is it possible to generalise to other young people seeking care in another country with another health care organisation, different alcohol culture, etc.? This becomes a paradox; if the treatment is complex, is it sure that all the parameters described are the central ones? Are the ones described in the studies central for the client group in focus?

The utmost purpose of the guidelines and for the evidence movement per se is to facilitate evidence-based decisions in practice. But closer comparison of the three guidelines shows that the general recommendations differ in a way that makes the recommendations difficult to valuate in a specific situation. One example is the set of recommendations for the MI-based intervention MET. To some extent these differences can be related to the three guidelines’ methodological differences. The UK guideline uses only RCT studies as an evidence base whereas, the APA and the Australian guidelines use other studies, but with a lower grade of evidence.7

However, for a practitioner consulting the guidelines it may not be reasonable or even possible to evaluate the technical differences that lead to the general recommendation of MET as a stand-alone treatment for clients with alcohol problems found in the APA guidelines compared with the claim in the UK guideline that there is not sufficient evidence to recommend motivational techniques as stand-alone interventions. One may then argue that the guidelines are supposed to be used in a national context. In the UK Guideline it is explicitly stated that it is for alcohol services funded by or provided for the NHS (Appendix 1, p. 468) and although this is only implicit in the other two, it is clear that the recommendations are addressing a national context. However, there is no such limitation in the search protocols for the evidence base used. Although the evidence base is dominated by North American and British research, the only recognised restriction is that of language (only papers written in English are used in the APA guideline). There is no mention of a discussion or restriction of the national context in which the studies are performed.

The producing of such thoroughly performed guidelines is a time-consuming, difficult and costly process. At the same time, the conclusions and recommendations eventually become out of date, espe-
cially in areas were the state of the art is insufficient. Thus, guidelines need to be revised regularly. In the description of the developing process of the APA guideline, revisions at regular intervals are planned. A solution to updating without repeating the process of developing completely new guidelines is the Guideline Watch published by the APA Executive Committee on Practice Guidelines. Watches may be authored and reviewed by experts associated with the original guideline development effort and are approved for publication by the Committee. Thus, watches represent the opinions of the authors with the approval of the Committee. After the publishing of the APA Guidelines in 2006, one such Guideline Watch was published in April 2007 (Smith Connery & Kleber, 2007). This is a five-page document including references, of which about two concern treatment of alcohol dependence. From the fact that no further document has since been published up to April 2013, when data for this article was collected, one may conclude that the updating of evidence is not working very well. It is also somewhat surprising that parallel or previously produced guidelines receive little mention. The procedure for the review of evidence in the Australian guidelines is described as including searches for other relevant guidelines; yet we found no reference to the APA guideline, published well before the Australian review.

In the study at hand we have shown that prestigious organisations in different national contexts have reached divergent conclusions on evidence-based practice and the quality of the scientific studies underpinning these conclusions. This situation of competing and non-congruent positions on what can be considered “best practice” (which must be regarded as totally unacceptable for the more fundamental EBP project) has by and large remained uncommented by the scientific community. At the international level some of the more influential actors have tried to implement even more far-reaching standardisations intended to counteract the presence of competing and non-congruent positions on which interventions constitute “best practice” in a given field. An organisation such as the Cochrane Collaboration can to a considerable extent be seen as built upon a fundament composed of its protocols, handbooks and databases. All of these can be viewed as explicit expressions of an aspiration to reduce and control irrelevant influences and to increase transparency concerning the reviews that it publishes. But this ambition to create an autonomous machinery (Bergmark, 2009) for synthesising knowledge, which in itself must be considered legitimate, also has to deal with the more fundamental nature of the scientific impulse:

There is always more to know on any matter...knowledge is always incomplete, always to be questioned, broadened and extended, enriched and refined, supplemented and controverted... the transparency process is potentially unending because there are always new accounts or revelations that can be sought (Cottorell, 1999).

The ambition to establish objective and thoroughly scientifically secured evidence will inevitably be undermined by the fact that the exponential growth of published studies in almost any scientific field dem-
onstrates that scientific knowledge will always be subject to revisions and it can therefore to a certain extent be said to be uncertain.

Declaration of Interest None

Anders Bergmark, Professor
Department of Social Work
Stockholm University
E-mail: anders.bergmark@socarb.su.se

Lisa Skogens, PhD
Department of Social Work
Stockholm University
E-mail: lisa.skogens@socarb.su.se

Ninive von Greiff, PhD
Department of Social Work
Stockholm University
E-mail: ninive.von.greiff@socarb.su.se

NOTES

1 In the APA guidelines, RCT stands for Randomized Clinical Trial. RCT is used in the article as an acronym both for Randomised Clinical Trial and Randomised Controlled Trial.

2 The PICO guide consists of questions such as: Patients/population: Which patients or population of patients are we interested in? Are there subgroups that need to be considered? Interventions: Which intervention, treatment or approach should be used? Comparison: What is/are the main alternative/s to compare with the intervention? Outcome: What is really important for the patient? Which outcomes should be considered? (p. 44)

3 The bracketed Roman numeral indicates the level of clinical confidence (see Table 1) for each treatment.

4 Examples of minimal intervention control are information and referral only, feedback and waitlist (p. 241).

5 The last literature search for the APA guidelines was performed in February 2005. A date for the last literature search is not mentioned in the Australian guidelines, but the review base for the guidelines was published in 2009. In the British guidelines, the last reported updating of the literature search is from March 2010.

6 A reason for this might be that Mesa Grande is a so called box-score review known as having major methodological drawbacks (Finney, 2000).

7 Even with the same strict use of evidence and similar focus, the selected studies may differ between reviews. An example of this is a review on MI for substance abuse made by the Cochrane Collaboration (Smedslund et al., 2011), published the same year as the UK guidelines. The Cochrane review has the same strict use of evidence, i.e. only RCT studies are used. The review’s selection criteria (persons dependent on or abusing substances) include the selection group of the UK guidelines (treatment-seeking population). However, only three of the references used as evidence in the UK guidelines (MATCH 1997, UKATT 2005, Sellman 2001) are used in the Cochrane review. Two of the references (Hester 2005, Sobell 2002) are excluded because they do not concern face-to-face interventions. Two references (Rosenblum 2005, Shakeshaft 2002) that are used in the UK guidelines are not mentioned at all in the Cochrane review.

The authors conclude that MI can reduce the extent of substance abuse compared with no intervention. However, the authors consider the evidence to be mostly of low quality and that further research is very likely to have an important impact on their confidence in the estimate of effect and is likely to change the estimate (Smedslund et al., 2011).
REFERENCES

Babor, T. F., & Del Boca, F. K. (2003). Treatment matching in alcoholism. Cambridge: Cambridge University Press.

Bergmark, A. (2009.) Drug misuse – Psychosocial interventions. Addiction, 104, 676–677.

Bergmark, A., & Lundström, T. (2011). Guided or independent? Social workers, central bureaucracy and evidence-based practice. European Journal of Social Work, 14(3), 323–337.

Cottorell, R. (1999). Transparency, mass media, ideology and community. Cultural Values, 3, 414–426.

Finney, J. W. (2000). Limitations in using existing alcohol treatment trials to develop practice guidelines. Addiction, 95(19), 1491–1500.

Hansen, H. F., & Rieper, O. (2012). Institutionalization of second-order evidence-producing organizations. In O. Rieper, F. L. Leeuw, & T. Ling (Eds.), The evidence book. Concepts, generation, and use of evidence (pp. 27–49). New Jersey: Transaction Publishers.

Mann, T. (1996). Clinical guidelines: Using clinical guidelines to improve patient care within the NHS. London: NHS Executive.

McCready, B. S. (2012). Health-care reform provides an opportunity for evidence-based alcohol treatment in the USA: The National Institute for Health and Clinical Excellence (NICE) guideline as a model. Addiction, 108, 231–232.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies and typologies: Horses for courses. Journal of Epidemiology and Community Health, 57, 527–529.

Sackett, D. L., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (1997). Evidence-based medicine: How to practice and teach EBM. New York: Churchill Livingstone.

Shand, F., Gates, J., & Fawsett, J. (2003). The treatment of alcohol problems: A review of the evidence. Canberra: Australian Government Department of Health and Aging.

Smidslund, G., Berg, R. C., Hammerstrøm, K. T., Steiro, A. Leiknes, K. A., Dahl, H. M., & Karlsen, K. (2011). Motivational interviewing for substance abuse (Review). The Cochrane Collaboration, John Wiley & Sons, Ltd.

Smith Connery, H., & Kleber, H. D. (2007). Guideline watch: Practice guideline for the treatment of patients with substance use disorders, 2nd edition. DOI: 10.1176/appi.books.9780890423363.149073

Timmermans, S., & Mauk, A. (2005). The promises and pitfalls of evidence-based medicine. Health Affairs, 24, 18–28.

GUIDELINES

Alcohol-Use Disorders. The NICE Guideline on Diagnosis, Assessment and Management of Harmful Drinking and Alcohol Dependence (2011). National Clinical Practice Guideline 115. National Collaborating Centre for Mental Health commissioned by the National Institute for Health & Clinical Excellence. Published by The British Psychological Society and The Royal College of Psychiatrists. Great Britain: Stanley L. Hunt Ltd.

Guidelines for the Treatment of Alcohol Problems (2009). Prepared by Haber, P., Lintzeris, N., Proude, E., & Lopatko, O. for Australian Government, Department of Health. The University of Sidney, Publication Number: P3-5625.

Practice Guideline for the Treatment of Patients with Substance Use Disorders (2006). Second Edition. American Psychiatric Association. Work group: Kleber, H. D., Weiss, R. D., Anton Jr, R. F., George, T. P., Greenfield, S. F, Kosten, T. R., O’Brien, C. P., Rounsaville, B. J., Strain, E. C., Ziedonis, D. M., Hennessy, M. D., & Connery, H. S.