12-step Programs for Reducing Illicit Drug Use: Protocol for a Systematic Review

Mette Deding, Madina Saidj, Lars Brännström, Margaretha Järvinen, Anne-Marie Klint Jørgensen & Maja Kärrman Fredriksson

PROTOCOL

THE CAMPBELL COLLABORATION

Contact details of lead author:
Ms. Mette Deding
Head of SFI- Campbell
SFI / The Danish National Centre for Social Research
Herluf Trolles Gade 11
Copenhagen K
DK- 1052 **DENMARK**
Email: mcd@sfi.dk

Table of contents

TABLE OF CONTENTS		**2**
1	**BACKGROUND**	**3**
1.1	Description of the condition	3
1.2	Description of the intervention	4
1.3	How the intervention might work	7
1.4	Why it is important to do this review	8
2	**OBJECTIVE OF THE REVIEW**	**10**
3	**METHODS**	**11**
3.1	Criteria for considering studies for this review	11
3.2	Search methods for identification of studies	14
3.3	Data collection and analysis	17
3.4	Data synthesis	25
4	**REFERENCES**	**27**
4.1	References	27
5	**FIGURES**	**31**
6	**SOURCES OF SUPPORT**	**33**
6.1	Internal sources	33
6.2	External sources	33
7	**APPENDICES**	**34**
7.1	Study eligibility screening level one & two	34
7.2	Data extraction	35
7.3	Assessment of RISK OF BIAS in included studies	42

1 Background

1.1 DESCRIPTION OF THE CONDITION

Illicit drug production and use[1] remains a severe problem worldwide (United Nations Office on Drugs and Crime [UNODC], 2010). A central issue in reducing the worldwide drug problem is the demand for illicit drugs, and hence the need to identify effective methods for reducing their use.

Prescription and recreational drugs may be differentiated. In this review, we reserve the term 'drug use' to apply to the illegal, nonmedical use of drugs. Globally, the United Nations Office on Drugs and Crime (UNODC) estimates that between 155 and 250 million people (3.5 to 5.7 percent of the population aged 15-64) used illicit substances at least once in 2008 (UNODC, 2010). Illicit substances include opium/heroin (opiates), coca/cocaine, cannabis[2], and amphetamine-type stimulants (which includes MDMA/Ecstasy). Cannabis is the most commonly used illicit substance (with an estimated 129 – 190 million users worldwide), followed by amphetamine-group substances, cocaine and opiates. UNODC considers some types of drug use to be more problematic than others, and defines problem drug use as that which involves the injection of drugs or the long-duration/regular use of opioids, cocaine and/or amphetamines. For 2008, UNODC estimates that between 16 and 38 million people worldwide are problematic drug users (UNODC, 2010).

Drug use is problematic because it is linked to a range of health and social problems, including crime, prostitution, and homelessness (Office of National Drug Control Policy [ONDCP], 2000; Shelton, Taylor, Bonner, & van den Bree, 2009; Silbert, Pines, & Lynch, 1982). The European Monitoring Centre for Drugs and Drug Addiction estimates that drug-induced deaths account for approximate 4 percent of all deaths of Europeans aged 15-39 (European Monitoring Centre for Drugs and Drug Addiction [EMCDDA], 2010). A number of studies have attempted to estimate the social costs of drug use, in terms of both the direct cost and the indirect costs that would be saved if drug use were abolished (EMCDDA, 2010). For Finland these costs have been estimated at between EUR 200 million and EUR 300 million in 2007, and in Scotland at EUR 5.1 billion in 2006.

[1] The terms use, misuse, abuse and dependence will be used interchangeably throughout the protocol and refer to an addiction stage of drug usage.
[2] Cannabis is illegal in most, but not all countries. For example, use of cannabis in small amounts is tolerated in the Netherlands.

The high human, social and economic costs of drug use motivate the strong political interest in treatment for illicit drug use and in examining different treatment types. The main types of treatment are cognitive-behavioral therapies, motivational enhancement, contingency management, psychoanalysis, network therapy and – the object of this review –12-step programs (Galanter & Kleber, 2008).

1.2 DESCRIPTION OF THE INTERVENTION

12-step approaches to the treatment of drug use are widespread in many countries. The oldest and most widely attended 12-step groups are provided by Alcoholics Anonymous (AA), which began in 1935 and have more than 2 million members according to their own recent membership survey (Alcoholics Anonymous, 2012; Mäkelä et al., 1996). The principles of AA serve as a model for other 12-step programs, of which Narcotics Anonymous (NA) is the largest focussing on drug use (Narcotics Anonymous, 2012). Today NA has more than 58,000 weekly meetings in 131 countries (Narcotics Anonymous, 2010). NA accepts all individuals with drug addiction, regardless of the particular drug or combination of drugs used. Other 12-step groups restrict themselves to specific types of abuse, such as Cocaine Anonymous, Pills Anonymous and Marijuana Anonymous (Cocaine Anonymous, 2012; Pills Anonymous, 2012; Marijuana Anonymous, 2012). The stated objective of the 12-step approaches is complete abstinence from the use of drugs, whereas other treatments such as psychosocial interventions or opioid substitution may focus on reducing drug use (EMCDDA, 2010). 12-step treatment approaches assume that, as a result of biological and/or psychological vulnerability, individuals have lost control over their drug use. Treatment attempts to bring about the individual's acceptance of the disease model of addiction, (i.e. that addiction is a lifelong disease), of an "addict" identity, and of abstinence as a treatment goal. It also attempts to motivate involvement in 12-step activities (for example, attending meetings, obtaining a sponsor, working through the 12 steps) (Finney, Noyes, Coutts, & Moos, 1998). The core ideology of the approach is to offer individuals a new way of living which will support them in breaking the cycle of addiction and in maintaining abstinence (Mercer & Woody, 1999). The suggested prescription for abstinence, referred to as the "six pack", is: *don't use no matter what, go to meetings, ask for help, get a sponsor, join a group, and get active* (Laudet, 2008).

The 12-step self-help groups work to specific principles. The meetings typically adhere to a prescribed format including 12-step readings (The Preamble, How and Why, The 12 Steps) at the start of the meeting, and a reciting of the Serenity prayer at the end for individuals who wish to

do so (Laudet, 2008). The disease model of addiction is central to the 12-step philosophy and recovery is seen as the process of attaining and maintaining abstinence. Recovery is viewed as a lifelong process, and members thus regard themselves as "recovering" (Mercer & Woody, 1999).

The basic idea is that individuals work their way into recovery by going through "Twelve Steps" starting with the recognition of being addicted to drugs, also known as "hitting the bottom[3]", and ending with the capability of helping others out of their own addiction (Narcotics Anonymous, 2008). A central element in the 12-step self-help groups is that participation is voluntary and that recovering individuals, and not professional staff, guide the treatment. Another important component is sponsorship whereby a member who has made progress in the 12-step recovery program shares his or her experience on an individual and continuous basis with another member who is attempting to attain or maintain abstinence through the 12 steps (Straussner & Spiegel, 1996; Laudet, 2008). Sponsors share their own "experience, strength and hope" with the sponsees and accompany them in working the steps towards recovery. The idea is that sponsorship also helps oneself in maintaining abstinence, formulated as: "*the cardinal virtue of sponsorship is the momentary loss of self-centeredness*" (Jennings & Alcoholics Anonymous, 1990). Sponsorship is something that a member her/himself decides to become involved in. In addition, sponsors themselves have their own sponsors who help them in their own struggle for abstinence (Narcotics Anonymous, 2008).

The exact wording of the 12 steps differs slightly between groups – the 12 steps of NA are presented in Figure 1 (Narcotics Anonymous, 2008). The steps contain a strong spiritual emphasis and encourage members to look outside themselves for strength (to seek a higher power) and to embrace spiritual values and practices that are outlined in the 12 steps. "*A power greater than ourselves*" is mentioned in step 2 and "*God as we understand Him*" in step 3. In addition, the steps emphasize the importance of reconstructing relationships with people who have been harmed by the drug use (e.g., family members). Inherent in the 12 steps is that addiction is a disease and as such is beyond personal influence. However an individual can decide to change and oppress this disease. The strong spiritual emphasis is unique to the 12 steps and the texts that are used include wording that appears religious. On the other hand, the texts do not endorse a particular faith and the phrase "*God as we understand Him*" is open to different interpretations. The spiritual emphasis may mean that some drug users are opposed to

[3] The insistence on hitting bottom lies in the belief that few individuals will be sincerely motivated to commit to recovery unless they have "hit the bottom". This bottom can be wherever the individual allows it to be. The central idea is that individuals must come to a turning point where they accept that they have reached a stage of complete defeat to drugs (Alcoholics Anonymous, 2005).

the intervention for ideological reasons, and hence the claim made by some that 12-step only works for individuals with the right motivation and outlook on life (Fiorentine, 1999).

The “Twelve Steps” are accompanied by “Twelve Traditions” providing guidelines for the self-help groups (Narcotics Anonymous, 2008). The 12 traditions of Narcotics Anonymous are presented in Figure 2. Because the individual groups are autonomous, there can be differences between them, but the basic concepts are the same. Each group meets at a regular time and place, and is in principle open to all drug users – the only requirement being the wish to become “clean”. Attendance is decided individually, but beginners are encouraged to attend “90 meetings in 90 days”. Furthermore, participation is not time-limited and the time needed to do the steps is also decided individually (Straussner & Spiegel, 1996). Often, drug users who have succeeded in attaining abstinence continue to participate in the meetings for years, and continue working the steps day after day. Abstinence anniversaries or birthdays are considered major accomplishments and an important way to mark success in attaining abstinence.

In addition to the self-help groups, 12-step approaches are used in Twelve Step Facilitation (TSF) treatments (MATCH, 1997; Ries, Galanter & Tonigan, 2008). These are typically of limited duration and organised around a treatment facilitation centre manned with professionals, with treatment based on 12-step principles. In short, TSF is the integration of 12-step self-help groups with professional treatment. Usually, the individual will participate in 12-step meetings (NA or other) after completing TSF treatment, and a parameter of success for the TSF treatment is that the individual becomes motivated and ready for the self-help group. The best known TSF treatment is the “Minnesota model”, originating from three centres founded in Minnesota in the late 1950s (Cook, 1988). The Minnesota model is characterised by the use of the 12-step philosophy as a foundation for therapeutic change, where the treatment goal is total abstinence. TSF treatment can be delivered in both in-patient and out-patient settings; the duration of the treatment can vary, but is typically around 6 weeks.

A cardinal rule of both TSF and self-help groups is anonymity. In attempt to protect individuals from society’s stigmatization, the 12-step approach gives priority to preserving members’ anonymity, which inevitably poses a challenge to research.

In this review, the focus is on 12-step treatments for users of illicit drugs, and we will include both self-help groups working with the 12 steps (like NA) and TSF treatments. Since drug users may also be alcohol dependent (Kessler et al., 1997), we will include studies where alcohol misuse is present providing drug misuse is the dominant presentation. Thus, although alcohol

can be part of the substance use we do not consider 12-step treatments primarily dealing with and aimed at alcohol dependencies.

1.3 HOW THE INTERVENTION MIGHT WORK

The 12-step interventions included in this review aim at making the substance users refrain completely from using illicit drugs. This is different from other treatments where a reduction of use can also be a goal. Abstinence is achieved by the drug user through working his/her way through the 12 steps. A central issue that dates back to the AA tradition is the acknowledgement of the addiction, and the acceptance of support by a sponsor in a self-help group or by professionals in a TSF setting. One of the keys to success posed by the AA, and hence the NA, is "*the therapeutic value of addicts working with other addicts*" along with the cardinal idea that the 12-steps offer "*a design for living*", a way of learning to live, that teaches skills and helps individuals to navigate and reach recovery (Alcoholics Anonymous, 2005; Narcotics Anonymous, 2008). The steps are carefully organized in an order, starting with the basic skills and continuing to the more advanced changes that individuals should seek to integrate into their life. Having a sponsor and being a sponsor is an important part of 12-step self-help groups. Sponsorship is viewed as an important tool in the progress of recovery in that it helps to grasp the components of living, offers encouragement and support such as when relapse occurs, but it may also "kill" any complacency among the sponsors themselves and thereby help them sustain their self-monitoring (Hornbacher, 2010). Also, central to the NA program is the statement that spirituality mediates 12-step involvement and later abstinence. It is suggested that, by working the 12 steps, one will have a spiritual awakening and that continued practice of spiritual principles will lead to sustained abstinence (Narcotics Anonymous, 2008).

The benefits associated with involvement in 12-step programs and the mechanisms by which these benefits occur have been thoroughly explored over the past two decades. Relatively little is known, however, about which specific behaviors catalyze the therapeutic psychological mechanisms. Self-efficacy, or the confidence to remain abstinent, has been identified as a major component and a consistent predictor of subsequent improvement (Moos & Timko, 2008; MATCH, 1997). The importance of spirituality for later abstinence is currently unclear (Moos & Timko, 2008; Tonigan & Connors, 2008). Regarding sponsorships, Humphreys & Noke (1997) point out that this social network can be more effective in helping the substance user than the support from concerned family members. The reliance on positive reinforcement (e.g., by recognizing abstinence anniversaries) and behavioral modelling (e.g., by having a sponsor) have

also been proposed as an underlying mechanism of change (Morgenstern, Bux, & Labouvie, 2002; Witkiewitz & Marlatt, 2011).

Membership demographics for Narcotics Anonymous collected at the 2009 World Convention of NA in Barcelona, Spain reveal that more men than women are members (58% vs. 42%); only very few teenagers are members of NA (2%), whereas the most typical member is aged 41-50 years (34%); ethnicity is dominated by Caucasians (73%) and most members are employed (71%) (Narcotics Anonymous, 2010). Perhaps due to this profile, Fiorentine (1999) notes that the 12 steps have been argued to work best for Christian, white, middle-class males. According to Fiorentine (1999) studies fail to support this view. Fiorentine (1999) also accentuates that 12-step interventions may also be inappropriate for drug users with major psychiatric disorders, drug users in early stages of addiction, and drug users uncomfortable with the religious or spiritual emphasis.

1.4 WHY IT IS IMPORTANT TO DO THIS REVIEW

Illicit drug use has enormous costs for individuals and societies in terms of social, health and criminal problems. 12-step programs are one of the most widespread treatments for drug use internationally. The general belief among clinicians is that 12-step is an effective approach (Forman, Bovasso, & Woody, 2001) and in the US, for example, it is common procedure by courts to mandate 12-step treatment. Although a large number of studies have examined the use of the programs, no systematic knowledge of the effectiveness of the intervention is currently available. This is, in part, due to the strict anonymity policy and the insistence on fluid membership, especially in the self-help groups, which makes it difficult for researchers to track members.

A Cochrane review has focused on the effect of 12-step programs on alcohol dependence (Ferri, Amato, & Davoli, 2009), but did not include dependence on illicit drugs. Individuals dependent on drugs, particularly illicit drugs, are often forced into a lifestyle that differs considerably from individuals dependent on alcohol, due to the criminal aspect of the drug use. People who are drug dependent are by definition engaging in illegal activity when obtaining their substance of dependence and/or the capital needed for its procurement. Research also suggests that it may be more difficult for the drug user to benefit from the 12 steps compared to the individual who is dependent on alcohol (Laudet, 2008), possibly due to the particular impact of drugs on brain neurons.

With its broad applicability, minimal cost, and potential benefit, the 12-step approach holds great appeal for policy makers. Knowledge about the effectiveness of the method compared to other treatments, as well as knowledge about the effect of different program elements, will therefore be of considerable interest to policy makers.

2 Objective of the review

The objective of this review is to assess the effectiveness of 12-step programs to reduce illicit drug use.

3 Methods

3.1 CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW

3.1.1 Types of studies

Study designs eligible for inclusion are:

- RCTs – randomized controlled trials, including cluster randomized trials.
- QRCTs – quasi-randomized controlled trials (where participants are allocated by, for example, alternate allocation, birth date, date of the week or month, case number or alphabetical order).
- QESs – quasi-experimental studies with a control group (where pre-treatment group equivalence is demonstrated via matching, statistical controls, or where there is evidence of equivalence on key risk variables (see section 3.3.3), or where key risk variables are controlled for statistically).

We justify the inclusion of QRCTs and QESs because the open-door membership policy of 12-step programs generally, and for the self-help groups in particular, leads to difficulty in assessing the effectiveness of treatment using a formal RCT design.

3.1.2 Types of participants

The population will consist of individuals who use illicit drugs, regardless of age, gender or ethnic background. We will include individuals who use one or more types of illicit drug.

Many studies report on individuals with both alcohol and drug use. We will exclude studies where alcohol use dominates drug use, either in consumption or in severity as measured by, for example, the Addiction Severity Index (McLellan, Luborsky, Woody, & O'Brien, 1980). When the reported information on usage is insufficient for a judgment to be made, we will contact the study authors for clarification and will use this information to determine whether the study is

included. If the study authors do not reply to our inquiry in time[4] the study will be listed as "Awaiting classification".

We will include studies of individuals who are enrolled in 12-step treatment regardless of the way in which their problem is labelled, and we will regard the terms 'use', 'abuse', 'misuse' or 'dependence' of illicit drugs as equivalent.

3.1.3 Types of interventions

Included studies must evaluate 12-step interventions (in either the self-help or TSF format) delivered with the aim of stopping or reducing illicit drug use. The following core principles need to be present for a 12-step intervention to be included:

1. Addiction is viewed as an illness.
2. There is a theme of spirituality (for example, a belief in a higher power).
3. The individual discusses problems within a fellowship of peers trying to help and encourage one another.
4. General guidance is provided in the 'Twelve Steps'.

We will include studies evaluating interventions that use the 12-step principles regardless of setting (for example, in-patient or out-patient) or the duration of treatment. 12-step interventions that focus solely on treating alcohol dependency, such as AA programs, will be excluded even though the study participants may also be addicted to an illicit drug or drugs.

Eligible comparison conditions will be no intervention, a waitlist control condition, or any other intervention(s).

3.1.4 Types of outcomes

Given that 12-step treatment accepts abstinence as the only successful treatment outcome, the primary outcome for this review is abstinence or reduction of drug use as measured by:

- Biochemical tests (such as urine screens).
- Estimates on drug use (self-reported, or reported by others such as parents, caregivers or therapists).

[4] We will apply a time limit of 14 days from sending our inquiry. If the study authors notify us before this deadline that they will be able to provide the information in a short time, we will include it even though provided after the deadline. All inquiries and answers will be stored electronically.

- Psychometric scales (such as the Addiction Severity Index (ASI; McLellan, Luborsky, Woody, & O'Brien, 1980) or a similar validated scale).

Secondary outcomes are as follows:

- Criminal behavior (criminal convictions; self-reported or reported by authorities, files, or registers).
- Prostitution (self-reported or reported by authorities, files, or registers).
- Psychiatric symptoms, measured using the Symptom Checklist-90-R (SCL-90-R; Derogatis 1983) or a similar validated scale.
- Social functioning, measured using the Social Functioning Questionnaire (SFQ; Tyrer et al, 2005) or a similar validated scale.
- Employment status (self-reported).
- Homelessness (self-reported).
- Retention (measured as treatment completion rate).

Studies will only be included if they consider the primary outcome.

Studies must permit calculation of a numeric effect size for the primary outcome(s) to be eligible for inclusion in the meta-analysis. If a study reports insufficient data for the calculation of a numeric effect size, we will contact the study authors requesting that data. If the study authors do not reply to our inquiry in time[5] the study will be included in the review but excluded from the meta-analysis.

Outcomes will be considered for the following intervals:

- Short-term effects (less than6 months after enrolment into treatment).
- Medium-term effects (6 months to less than 12 months after enrolment into treatment).
- Long-term effects (12 months or more after enrolment into treatment).

[5] We will apply a time limit of 14 days from sending our inquiry. If the study authors notify us before this deadline that they will be able to provide the information in a short time, we will include it even though provided after the deadline. All inquiries and answers will be stored electronically.

3.2 SEARCH METHODS FOR IDENTIFICATION OF STUDIES

3.2.1 Electronic searches

Relevant studies will be identified through electronic searches of bibliographic databases, government policy databanks and internet search engines. No date or language restrictions will be applied to the searches. The following bibliographic databases will be searched:

Medline
PsycINFO
Cinahl
Embase
Cochrane Library (including CENTRAL)
Science Citation Index
Social Science Citation Index
SocINDEX
Social Care Online
ASSIA
Sociological Abstracts
Dissertation Abstracts

The following Nordic databases will be searched:

SweMed+
Artikelsök
DiVA (Digitale vetenskapeliga arkivet)
Bibliotek.dk, Bibsys, Libris

3.2.2 **Search terms**

An example of the search strategy for MEDLINE searched through the OVID interface is listed below. This strategy will be modified for the different databases. We will report details of the modifications in the completed review.

1 Narcotics anonymous.ab,kw,ti.
2 Cocaine Anonymous.ab,kw,ti.
3 Crystal Meth Anonymous.ab,kw,ti.
4 Pills Anonymous.ab,kw,ti.
5 Marijuana Anonymous.ab,kw,ti.
6 Heroin Anonymous.ab,kw,ti.
7 1 or 2 or 3 or 4 or 5 or 6
8 (Self-Help adj1 group*).ab,ti.
9 (Support* adj1 group*).ab,ti.
10 twelve-step*.ab,ti.
11 12-step*.ab,ti.
12 Self-Help Groups/
13 (Recover* adj1 group*).ab,ti.
14 (Minnesota adj3 (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*)).ab,ti.
15 TSF.ab,ti.
16 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
17 Amphetamine-Related Disorders/
18 Cocaine-Related Disorders/
19 Marijuana Abuse/
20 Opioid-Related Disorders/
21 Substance Abuse, Intravenous/
22 Substance Withdrawal Syndrome/
23 Heroin Dependence/
24 17 or 18 or 19 or 20 or 21 or 22 or 23
25 Opium/
26 Heroin/
27 exp Cannabinoids/ or Cannabis/
28 Marijuana Smoking/
29 exp cocaine/
30 Methamphetamine/
31 Amphetamine/
32 Designer Drugs/
33 Drug*.ab,ti.
34 Substance*.ab,ti.
35 Stimulan*.ab,ti.
36 Narcotic*.ab,ti.
37 Opium.ab,ti. or opiate*
38 Heroin.ab,ti.
39 Crack.ab,ti.

40 Cocaine*.ab,ti.
41 (Methamphetamine* or Amphetamin*).ab,ti.
42 Ecstasy.ab,ti.
43 Fantasy.ab,ti.
44 (Cannabis or Marijuana or Hashish).ab,ti.
45 (Designer adj1 drug*).ab,ti.
46 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45
47 (substance* or drug* or stimulan* or Narcotic*) adj3 (abus* or us* or misus* or addict* or depend*) .ab,ti
48 7 or 16
49 24 and 48
50 46 and 48
51 47 and 48
52 49 or 50 or 51
53 limit 52 to humans

3.2.3 **Searching other resources**

We will examine the reference lists from relevant reviews identified in the electronic searches and from included primary studies for new leads. In addition, international experts will be contacted in attempt to identify unpublished and on-going studies.

3.2.4 **Grey literature**

Additional searches will be conducted using *Google* and *Google Scholar* and the first 200 hits examined in each case. *OpenGrey* will be used to search for European grey literature (http://opengrey.eu/). All relevant documents will be copied and the exact URL and date of access recorded. The following government policy databanks will be searched:

National Institute on Drug Abuse (NIDA) (http://www.drugabuse.gov/)

The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (http://www.emcdda.europa.eu/index.cfm)

Substance abuse and Mental Health Services administration (SAMHSA), Office of Applied Studies (http://www.samhsa.gov/)

Substance abuse and Mental Health Data Archive (SAMHDA) (http://www.icpsr.umich.edu/icpsrweb/SAMHDA/index.jsp

Government of Canada (http://recherche-search.gc.ca/s_r?t3mpl1t34d=1&s5t34d=canada&l7c1l3=eng&S_08D4T.1ct57n=form&S_08D4T.s3rv5c3=basic)

USA Government (http://www.usa.gov/)

The Australian Government (http://publications.gov.au/)

The Ministry of Social Affairs and Integration (http://www.sm.dk/Publikationer/Sider/Start.aspx)

Government Offices of Sweden (http://www.regeringen.se/sb/d/107/a/136)

The Norwegian Ministry of Health and Care Services (http://www.regjeringen.no/nb/dep/hod.html?id=421)

EU (http://europa.eu/index_en.htm)

Social Care Online (http://www.scie-socialcareonline.org.uk/)

3.2.5 Hand searching

The journals *Addiction* and the *Journal of Substance Abuse Treatment* will be hand searched from 2012-2013. These are the two journals we consider most likely to include relevant primary studies.

3.3 DATA COLLECTION AND ANALYSIS

3.3.1 Selection of studies

Under the supervision of review authors, review team assistants will first independently screen titles and abstracts to exclude studies that are clearly irrelevant. Studies considered eligible by at

least one of the reviewers will be retrieved in full text. Each full text will then be appraised by two members of the review team who will decide whether the inclusion criteria are met. Any disagreements about eligibility will be resolved by the review authors. The main reason for exclusion will be documented for each study that is retrieved in full text and subsequently excluded. A study inclusion coding sheet (appendix 1) will be piloted by the review authors and modified if required.

The selection process will have the following steps and will be documented using appropriate software:

1.Review team assistants will, under supervision of review authors, independently select potentially-eligible studies for full-text retrieval on the basis of the inclusion criteria by considering the Title, Abstract, and Subject Terms for each document. A study will be retrieved in full text if reviewers disagree about its potential eligibility.

2.Pairs of reviewers from the review team will independently read documents in full and decide to include or exclude the document on the basis of the inclusion criteria. If reviewers disagree, review authors will have a decisive vote. Primary reasons for exclusion will be documented.

3.The complete list of included documents will be sent to a selected group of external international experts together with the inclusion criteria. These experts will be asked whether they know of any eligible studies that are missing.

4.Suggested studies from the external international experts that meet the inclusion criteria will be processed in accordance with step 2 above

5.The final set of included documents will be examined to identify any multiple publications from the same study or multiple studies in single publications. The aim is to select a set of unique studies. The problem of multiple publications from single studies will be approached by looking at the site and time frame of the evaluation, the intervention, the number of participants in treatment and control groups respectively.

The overall search and screening process will be illustrated in a flow-diagram.

3.3.2 **Coding and management**

Guided by the checklist of items in the Cochrane Handbook (Higgins & Green, 2008, Table 7.3.a), at least two independent coders will extract and electronically store data, focusing on populations, interventions, comparisons, outcomes, independence of evaluators and study design as basic coding categories. Differences in coding will be resolved by discussion by the review authors. When necessary the trial investigators will be contacted to request further

information[6]. A data extraction sheet and questions (appendix 2) will be piloted on several studies and revised if necessary.

3.3.3 **Assessment of risk of bias in included studies**

We will assess the methodological quality of studies using a risk of bias model developed by Prof. Barnaby Reeves in association with the Cochrane Non-Randomised Studies Methods Group.[7] This model is an extension of the Cochrane Collaboration's risk of bias tool; it covers risk of bias both in RCTs, and also in non-randomized studies that have a well-defined control group.

The point of departure for the risk of bias model is the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2008). The existing Cochrane risk of bias tool needs elaboration when assessing non-randomized studies because particular attention must be paid to selection bias and risk of confounding for the latter. It is also important to try to discriminate between non-randomized studies at varying risk of bias, resulting in the model requiring assessment on a 5-point scale for some items.

Risk of bias judgement items

The risk of bias model is based on 9 items (see appendix 7.3). For some items, risk is judged to be High, Low, or Uncertain; other items are judged on a 5-point scale where 1 corresponds to No/Low risk of bias and 5 correspond to Yes/High risk of bias. A score of 5 indicates that the risk of bias is sufficiently high that the findings will not be considered in the data synthesis (because they are more likely to mislead than inform).

The 9 risk of bias items concern **sequence generation** (relevant for selection bias), **allocation concealment** (relevant for selection bias), **confounders** (relevant for selection bias in non-randomized studies), **blinding** (relevant for performance, detection and attrition bias), **incomplete outcome data** (relevant for attrition bias), **selective outcome reporting** (relevant for reporting bias), **other potential threats to validity** (relevant for performance,

[6] We will apply a time limit of 14 days from sending our inquiry. If the study authors notify us before this deadline that they will be able to provide the information in a short time, we will include it even though provided after the deadline. All inquiries and answers will be stored electronically.
[7] This risk of bias model was introduced by Prof. Reeves at a workshop on risk of bias in non-randomized studies at SFI Campbell, February 2011. The model is a further development of work carried out in the Cochrane Non-Randomised Studies Method Group (NRSMG).

detection and other sources of bias), **a priori protocol,** and **a priory analysis plan** (relevant for reporting bias).

Confounders

An important part of the risk of bias assessment for a non-randomized study is how the confounding factors have been dealt with (see appendix 7.3). Selection bias is understood as systematic baseline differences between groups, and which can therefore compromise their comparability. Baseline differences can be observable (e.g., age and gender) and unobservable to the researcher (e.g., motivation and "ability"). There is no single non-randomized study design that resolves the selection problem in all circumstances. Different designs attempt to solve the problem under different assumptions and require different types of data, particularly in relation to factors that are unobservable. The "right" method depends on the assumptions about the nature of the process by which participants are selected into a program. As there is no universally correct way to construct counterfactuals, we will assess the extent to which the identifying assumptions (the assumption that makes it possible to identify the counterfactual) are explained and discussed by the study investigators.

In this review, the risk of bias from confounding is an additional item for each non-randomized study, and will be assessed for each outcome. Such an assessment requires a list of pre-specified potential confounders. For this review, we have identified the following confounding factors as the most relevant: age, gender, socio-economic status, mental health problems, and history of drug use. The motivation for focusing on these confounders is that they are the major risk factors related to drug use. Young people have a higher risk of use than older people (Labouvie, 1996), women have lower risk than men and have different drug use patterns (Brady & Back, 2008), and people with poor socio-economic status have higher risk (Spooner & Hetherington, 2004). The issue of drug users with mental health problems needs special focus, because the mental health problems can interfere with the effect of the drug treatment (Ross, 2008). Finally, the history of drug use is important for the likelihood of treatment success, e.g. duration of use and previous treatment (Greenfield & Hennessy, 2008).

Other confounders identified by the study authors will be listed in the risk of bias item and assessed in the same way as the most relevant confounders. We will also assess how each study deals with factors that are unobservable.

The risk of bias item will take into account the following:

- Proportion of confounders considered.
- Whether most important confounders were considered.
- Precision with which confounders were measured.
- Extent of imbalance between groups at baseline.
- Care with which adjustment was done.

The final judgment of this risk of bias item will be made on a scale from 1 to 5 (or unclear), where a score of 1 reflects low risk and a score of 5 reflects a high risk of bias in relation to confounding. For a judgement of low risk of bias in this item, all important confounders should be balanced at baseline or measured "well" and "carefully" controlled for in the analysis. The final judgment of the confounding item will be included in the overall risk of bias table.

Assessment

Two review authors will independently assess the risk of bias for each included study and report this in a table in the completed review. Disagreements will be resolved by discussion or by seeking advice from experts with content and statistical expertise.

3.3.4 **Measures of treatment effect**

Discrete data

For dichotomous outcomes we will calculate odds ratios or risk ratios, together with 95% confidence intervals and p-values.

Continuous data

For continuous outcomes, effects sizes will be calculated with 95% confidence intervals when means and standard deviations are available. If such information is unavailable, we will use the methods described by Lipsey & Wilson (2001) to calculate standardized mean differences SMDs from, for example, F ratios, t-values, chi-squared values and correlation coefficients. We will use Hedges *g* to estimate SMDs where scales have been used to measure the same outcomes in different ways (e.g., reduction of drug use). If there is a mix of studies with some reporting change scores and others reporting final values, we will contact the trial investigators and

request the final values[8]. If these are unobtainable, we will analyse change scores and final values separately (Higgins & Green, 2008, section 9.4.5.2).

We will only transform dichotomous effect sizes to SMD if appropriate, as may be the case with the outcome abstinence and reduction of drug use which can be measured with both binary and continuous data. If this is necessary, we will use the methods suggested by Sáchez-Meca, Marín-Martínes, & Chacón-Moscoso (2003) to allow dichotomous and continuous data to be pooled together.

When effect sizes cannot be pooled, study-level effects will be reported narratively. Statistical analyses will be conducted using RevMan, Excel and STATA software.

3.3.5 **Unit of analysis issues**

For each study, we will determine whether individuals were randomized in groups, whether individuals had undergone multiple interventions, whether results were reported at multiple time points, and whether there were multiple treatment groups.

Cluster randomized trials

In cluster randomized trials, errors in statistical analysis can occur when the unit of allocation differs from the unit of analysis. In cluster randomized trials, participants are randomized to groups in clusters, either when data from multiple participants in a setting are included (creating a cluster within the residential or community setting), or when participants are randomized by treatment locality or clinic. In such studies, standard errors may be biased if the unit-of-analysis is the individual. In cases where study authors have not applied appropriate statistical methods to control for clustering, we will attempt to estimate the intra-cluster correlation coefficient (see Donner, Piaggio, & Villar, 2001) and correct standard errors.

Multiple interventions per individual

In some of the studies, participants in the treatment group may receive the 12-step intervention plus some other treatment. These studies will be analysed separately from studies where the

[8] We will apply a time limit of 14 days from sending our inquiry. If the study authors notify us before this deadline that they will be able to provide the information in a short time, we will include it even though provided after the deadline. All inquiries and answers will be stored electronically.

treatment group receives 12-step alone. The discussion of those results will take into account the additional treatments received.

Multiple time points

When the results are measured at multiple time points, each outcome at each time point will be analysed in a separate meta-analysis. Comparable studies taking measures at a similar time point will be analysed together, grouped as follows: short-term (less than 6 months after enrolment), medium-term (6 months to less than 12 months after enrolment), and long-term (12 months or more after enrolment).

Studies with multiple treatment/control groups

For studies where there are multiple treatment/control groups, data from the same group will not be analysed twice, and so multiple contrasts from the same study will not be pooled in the same meta-analysis. Where there are multiple treatment conditions, we will select the one which matches the inclusion criteria.

Where there is more than one control group, we will conduct pair-wise comparisons, so that effects will be calculated on the basis the following: A vs. B, A vs. C, and B vs. C. When A is a defined intervention and C is a no-intervention or placebo-type of intervention, we will use the term "absolute effect". When A is a defined intervention and B also is a defined intervention, we will use the term "relative effect". When "usual care" is a sufficiently described, we will use the term "relative effect". If "usual care" is effectively "no intervention", then the term "absolute" will be used. In cases when "usual care" is not defined, additional information from the authors or from other reliable sources will be sought and used to make the judgement.

3.3.6 **Dealing with missing data and incomplete data**

Missing data will be assessed in the included studies. Reasons, numbers, and characteristics of those dropping out of each included study will be investigated and reported. Efforts will be made to contact the trial investigators for further information if necessary[9]. Where possible, we will

[9] We will apply a time limit of 14 days from sending our inquiry. If the study authors notify us before this deadline that they will be able to provide the information in a short time, we will include it even though provided after the deadline. All inquiries and answers will be stored electronically.

conduct meta-analyses using data from all allocated participants, and will report when that is not the case. For studies where missing data are not available, a sensitivity analysis will be used to assess the extent to which the results might be biased by missing data.

Although we will seek any important but unreported data from the trial investigators, it is reasonable to assume that this approach will not always be successful. We will therefore consider utilizing in addition the imputation methods outlined in White & Higgins (2009). The imputation method selected will depend on the specific nature of the missing data.

3.3.7 **Assessment of heterogeneity**

Statistical heterogeneity within the included studies will be examined through the use of the χ^2-test, where a low p-value ($p<0.1$) will be taken as indicating significant heterogeneity of treatment effects. The I^2 statistic will also be used to determine the percentage of variability that is due to heterogeneity rather than sampling error or chance (Higgins & Green, 2008). According to Higgins & Thompson (2002), percentages over 75% are considered to indicate a high degree of heterogeneity, and we will use this percentage as cut-off. In addition, we will report the between-studies variance component (τ^2).

We will also consider issues such as design quality, publication bias, and differences in participant's characteristics as possible reasons for any heterogeneity and will conduct sensitivity analyses accordingly where data permit. If heterogeneity is identified, we will investigate possible sources using subgroup analyses, meta-regression and sensitivity analyses.

3.3.8 **Assessment of publication biases**

We will inspect funnel plots for information about possible publication bias if we find sufficient studies (Higgins & Green, 2008). We are aware, however, that asymmetric funnel plots are not necessarily caused by publication bias (and publication bias does not necessarily cause asymmetry in a funnel plot). If asymmetry is present, likely reasons will be discussed. If a sufficient number of studies are found, we will, in addition, conduct a regression-based test (Eggers test or Peters test) to assess the possibility of small study effects.

3.4 DATA SYNTHESIS

Analysis of absolute effects will involve comparing the 12-step intervention to no treatment conditions and to untreated waitlist controls. Analysis of relative effects comparing the 12-step intervention with other interventions will be conducted separately. Data from all follow-up durations will be recorded and we will conduct separate analyses for short-term, medium-term and long-term outcomes. Studies that have been coded with a very high risk of bias (score of 5 in any item judged on a 5-point scale) will not be included in the data synthesis.

Meta-analysis will be conducted where effect sizes are available or can be calculated, and where studies are similar in terms of the study design, the comparison condition, and the outcome measured. Random effects meta-analysis will be used. We will report the 95% confidence intervals and provide a graphical display (forest plot) of effect sizes. When meta-analysis is inappropriate, a narrative description of the individual study results will be provided.

3.4.1 Moderator analysis and investigation of heterogeneity

If heterogeneity is judged to be large, we will investigate the following factors in attempt to explain the observed heterogeneity: intervention characteristics (e.g., self-help vs. TSF, treatment duration, treatment intensity), and summaries of participant characteristics measured at the study level (e.g., gender, mental health problems, and history of drug use).

If the number of included studies is sufficient and the spread of the study means of the covariates and study sizes are appropriate (Borenstein, Hedges, Higgins, & Rothstein, 2009; Simmonds & Higgins, 2007; White, 2009), we will perform moderator analyses (meta-regression) to explore how observed variables are related to heterogeneity using a mixed model. We will estimate the (new) residual component to be used in a weighted least squares analysis conditional on this variance component estimate. The residual variance component will be estimated using the method-of-moments estimator (Hartung, Knapp, & Sinha, 2008; Konstantopoulos, 2006). We will report the 95% confidence intervals for regression parameters. Conclusions from meta-regression analysis will be cautiously drawn and will not be based on significance tests.

Otherwise, single factor subgroup analysis will be performed. The assessment of any difference between subgroups will be based on 95% confidence intervals. No conclusions from subgroup

analyses will be drawn and interpretation of relationships will be cautious, as they are based on subdivision of studies and indirect comparisons.

3.4.2 **Sensitivity analysis**

Sensitivity analysis will be used to evaluate whether the effect sizes are robust across study designs and components of methodological quality. For methodological quality, we will consider sensitivity analysis for each major component of the risk of bias model. Sensitivity analysis will further be used to examine the rigour of conclusions in relation to the quality of data (outcome measures based on different time intervals and different data sources).

4 References

4.1 REFERENCES

Alcoholics Anonymous. (2012). Retrieved February 28, 2012), from www.aa.org.

Alcoholics Anonymous. (2005). *Twelve Steps and Twelve Traditions*. Alcoholics Anonymous. Electronic PDF version.

Borenstein, M., Hedges, L. V. , Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to metaanalysis*. Chichester: Wiley.

Brady, K. T., & Back, S. E. (2008). Women and addiction. In Galanter & Kleber (eds.): *The American psychiatric publishing textbook of substance abuse treatment*. Washington, DC, American Psychiatric Publishing.

Cocaine Anonymous. (2012). Retrieved February 28, 2012, from www.ca.org.

Cook, C. (1988). The Minnesota model in the management of drug and alcohol dependency: Miracle, method or myth? Part I. The philosophy and the programme. *British Journal of Addiction*, 83, 625-634.

Derogatis, L.R. (1983). SCL-90: Administration, Scoring and Procedures Manual-I for the Revised Version and other Instruments of the Psychopathology Rating Scale Series. Baltimore, MD: Johns Hopkins University School of Medicine, Clinical Psychometrics Research Unit.

Donner, A., Piaggio, G., & Villar, J. (2001). Statistical methods for the meta-analysis of cluster randomized trials. *Statistical Methods in Medical Research, 10*, 325-338.

Dunst, C. J., Trivette, C. M., & Deal, A. G. (1994). Supporting and strengthening families: Methods strategies and practices. Cambridge: Brookline Books.

European Monitoring Centre for Drugs and Drug Addiction EMCDDA (2010). *Annual report 2009: The state of the drugs problem in Europe*. Luxembourg: Publications office of the European Union.

Ferri, M., Amato, L., & Davoli, M. (2009). Alcoholics Anonymous and other 12-step programmes for alcohol dependence. *The Cochrane Library, 3*.

Finney, J. W., Noyes, C. A., Coutts, A. I., & Moos, R. H. (1998), Evaluating Susbtance Abuse Treatment Process Models: I. Changes on Proximal Outcome Variables during 12-Step and Cognitive-Behavioral Treatment. *Journal of Studies on Alcohol, 59*, 371-380.

Fiorentine, R. (1999). After drug treatment: Are 12-step programs effective in maintaining abstinence. *American Journal of Drug and Alcohol Abuse, 25(1),* 93-116.

Forman, R. F., Bovasso, G., & Woody, G. (2001). Staff beliefs about addiction treatment. *Journal of Substance Abuse Treatment, 21,* 1-19.

Galanter, M., & Kleber, H. D. (eds.) (2008). *The American psychiatric publishing textbook of substance abuse treatment*. Washington, DC: American Psychiatric Publishing.

Greenfield, S. F., & Hennessy, G. (2008). Assessment of the patient. In Galanter & Kleber (eds.): *The American psychiatric publishing textbook of substance abuse treatment*. Washington, DC: American Psychiatric Publishing.

Hartung, J., Knapp, C., & Sinha, B. K. (2008). *Statistical meta-analysis with applications*. New York: Wiley.

Higgins, J. P. T., & Green, S. (eds.) (2008). *Cochrane handbook for systematic reviews of interventions version 5.0.2* [updated September 2009]. The Cochrane Collaboration (available from www.cochrane-handbook.org).

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a metaanalysis. *Statistics in Medicine, 21*, 1539-1558.

Hornbacher, M. (2010). Sane: mental illness, addiction, and the twelve steps. Center City, MN: Hazelden.

Humphreys, K., & Noke, J. M. (1997). The influence of posttreatment mutual help group participation on the friendship networks of substance abuse patients. *American Journal of Community Psychology, 25*, 1-16.

Jennings, J., & Alcoholics Anonymous (1990). *The Little Red Book*. Center City, MN: Hazelden.

Kessler, R. C., Crum, R. M., Warner, L. A., Nelson, C. B., Shulenberg, J., & Anthony, J. C. (1997). Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry, 54*, 313-321.

Konstantopoulos, S. (2006). Fixed and Mixed Effects Models in Meta-Analysis. *IZA Discussion Paper no.2198*.

Labouvie, E. (1996). Maturing out of substance use: Selection and self correction. *Journal of Drug Issues, 26(2)*, 457-476.

Laudet, A. B. (2008). The Impact of Alcoholics Anonymous on other substance abuse related Twelve Step programs. *Recent Developments in Alcoholism, 18*, 71-89.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks: Sage Publications.

Mäkelä, K., Arminen, I., Bloomfield, K., Eisenbach-Stangl, I., Hermersson Bergmark, K., Kurube, N., Mariolini, N., ... Zielinksi, A. (1996). *Alcoholics Anonymous as a mutual-help movement: A study in eight societies*. Wisconsin University Press, Madison, Wisconsin, USA.

Marijuana Anonymous. (2012). Retrieved February 28, 2012, from www.marijuana-anonymou.org.

MATCH (1997). Matching alcoholism to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol and Drugs, 58(1),* 7-29.

McLellan, A. T., Luborsky, L., Woody, G. E., & O'Brien, C. P. (1980). An improved diagnostic evaluation instrument for substance abuse patients. The addiction severity index. *Journal of Nervous and Mental Disorders, 168(1),* 26-33.

Mercer, D. E., & Woody, G. E. (1999). Contributions of the 12-Step Approach. In *An individual drug counselling approach to treat cocaine addiction*. National Institute on Drug Abuse, Division of Clinical and Services Research.

Moos, R. H., & Timko, C. (2008). *Outcome Research on 12-Step and Other Self-help Programs* In Galanter & Kleber (eds.): The American psychiatric publishing textbook of substance abuse treatment. Washington, DC:American Psychiatric Publishing.

Morgenstern, J., Bux, D., & Labouvie, E. (2002). Examining mechanisms of action in 12-step treatment: The role of 12-step cognitions. *Journal on Studies of Alcohol and Drugs, 63*, 665-671.

Narcotics Anonymous (2008). *Narcotics Anonymous –Basic text*. Sixth edition. Narcotics Anonymous World Services, Inc.

Narcotics Anonymous. (2010). Information about NA. Retrieved February 28, 2012), from http://www.na.org/admin/include/spaw2/uploads/pdf/PR/Information_about_NA.pdf.

Narcotics Anonymous. (2012). Retrieved February 28, 2012, from www.na.org.

Office of National Drug Control Policy ONDCP (2000). *Drug-related crime*. Office of National Drug Control Policy, Drug Policy Information Clearinghouse, fact sheet.

Pills Anonymous. (2012). Retrieved February 28, 2012, from www.pillsanomynous.org.

Ries, R. K., Galanter, M., & Tonigan, J. S. (2008). *Twelve-step facilitation*. In Galanter & Kleber (eds.): The American psychiatric publishing textbook of substance abuse treatment. Washington, DC: American Psychiatric Publishing.

Ross, S. (2008). *The mentally ill substance abuser*. In Galanter & Kleber (eds.): The American psychiatric publishing textbook of substance abuse treatment. Washington, DC: American Psychiatric Publishing.

Sáchez-Meca, J., Márin-Martínes, F., & Chacón-Moscoso, S. (2003). Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis. *Psychological Methods, 8(4),* 448-467.

Shelton, K. H., Taylor, P. J., Bonner, A., & van den Bree, M. (2009). Risk factors for homelessness: Evidence from a population-based study. *Psychiatric Services, 60*, 465-472.

Silbert, M. H., Pines, A. M., & Lynch, T. (1982). Substance abuse and prostitution. *Journal of Psychoactive Drugs, 14(3),* 193-7.

Simmonds, M. C., & Higgins, J. P. T. (2007). Covariate heterogeneity in meta-analysis: Criteria for deciding between meta-regression and individual patient data. *Statistics in Medicine, 26*, 2982-2999.

Spooner, C., & Hetherington, K. (2004). *Social determinants of drug use*. National Drug and Alcohol Research Centre, University of New South Wales, Technical Report Number 228.

Straussner, S., & Spiegel, B. (1996). An analysis of 12-step programs for substance abusers from a developmental perspective. *Clinical Social Work Journal, 24(3),* 299-309.

Tonigan, J. S., & Connors, G. J. (2008). *Psychological mechanisms in Alcoholics Anonymous*. In Galanter & Kleber (eds.): The American psychiatric publishing textbook of substance abuse treatment. Washington, DC: American Psychiatric Publishing.

Tyrer, P., Nur, U., Crawford, M., Karlsen, S., McLean, C., Rao, B. & Johnson, T. (2005). The Social Functioning Questionnaire: a rapid and robust measure of perceived functioning. Int J Soc Psychiatry, 51(3):265-75.

United Nations Office on Drugs and Crime UNODC (2010), *World Drug Report 2010*. United Nations Publication.

White, I. R., & Higgins, J. P. T. (2009). *Meta-analysis with missing data*. In Sterne (ed.): Meta-analysis in Stata: An updated collection from the Stata Journal. Stata Press.

White, I. R. (2009). *Multivariate random-effects meta-analysis*. In Sterne (ed.): Meta-analysis in Stata: An updated collection from the Stata Journal. Stata Press.

Witkiewitz, K., & Marlatt, A. G. (2011). Behavioral Therapy Across the Spectrum. *Alcohol Research & Health, 33*, 4.

5 Figures

Figure 1: The 12 steps of Narcotics Anonymous (Narcotics Anonymous, 2008)

1. We admitted that we were powerless over our addiction, that our lives had become unmanageable.
2. We came to believe that a power greater than ourselves could restore us to sanity.
3. We made a decision to turn our will and our lives over to the care of God *as we understand Him*.
4. We made a searching and fearless moral inventory of ourselves.
5. We admitted to God, to ourselves, and to another human being the exact nature of our wrongs.
6. We were entirely ready to have God remove all these defects of character.
7. We humbly asked him to remove our shortcomings.
8. We made a list of all persons we had harmed, and became willing to make amends to them all.
9. We made direct amends to such people wherever possible, except when to do so would injure them or others.
10. We continued to take personal inventory when we were wrong and promptly admitted it.
11. We sought through prayer and meditation to improve our conscious contact with God *as we understood Him*, praying only for knowledge of His will for us and the power to carry that out.
12. Having had a spiritual awakening as the result of these steps, we tried to carry this message to addicts, and to practice these principles in all our affairs.

Figure 2: The 12 traditions of Narcotics Anonymous (Narcotics Anonymous, 2008)

1. Our common welfare should come first; personal recovery depends on NA unity.
2. For our group purpose there is but one ultimate authority – a loving God as He may express Himself in our group conscience. Our leaders are but trusted servants; they do not govern.
3. The only requirement for membership is a desire to stop using.
4. Each group should be autonomous except in matters affecting other groups or NA as a whole.
5. Each group has but one primary purpose – to carry the message to the addict who still suffers.
6. An NA group ought never endorse, finance, or lend the NA name to any related facility or outside enterprise. Lest problems of money, property, or prestige divert us from our primary purpose.
7. Every NA group ought to be fully self-supporting, declining outside contributions.
8. Narcotics Anonymous should remain forever nonprofessional, but our service centres may employ special workers.
9. NA, as such, ought never be organised, but we may create service boards or committees directly responsible to those they serve.
10. Narcotics Anonymous has no opinion on outside issues: hence the NA name ought never be drawn into public controversy.
11. Our public relations policy is based on attraction rather than promotion; we need always maintain personal anonymity at the level of press, radio, and films.
12. Anonymity is the spiritual foundation of all our Traditions, ever reminding us to place principles before personalities.

6 Sources of support

6.1 INTERNAL SOURCES

SFI Campbell, Copenhagen, Denmark; Swedish Institute for Social Research, Stockholm University, Sweden; the National Board of Health and Welfare, Stockholm, Sweden.

Review team at SFI Campbell consists of research assistants Anne-Sofie Due Knudsen, Simon Helth Filges Stine Lian Olsen, and Pia Vang Hansen.

6.2 EXTERNAL SOURCES

7 Appendices

7.1 STUDY ELIGIBILITY SCREENING LEVEL ONE & TWO

Screening level one (on the basis of titles and abstracts)

Reference id. no.
Study id. no.
Reviewer's initials
Year of publication:
Author:

1. **Is the study about a 12-step intervention?**
   Yes
   No (if no stop here and exclude)
   Uncertain

2. **Are the participants in treatment for drug abuse?**
   Yes
   No (if no stop here and exclude)
   Uncertain

3. **Is the study a quantitative primary study?**
   Yes
   No (if no stop here and exclude)
   Uncertain

The report reference is excluded if one of the answers to question 1 to 3 are no.
If the answers are yes or uncertain the full report is retrieved for second level screening. All uncertain questions for 1-3 need to be posed again based on the full text. If information is not available or the report is unclear report authors will be contacted to clarify study eligibility.

Additional questions for second level screening

4. **Is the report a RCT study (with a control group that is no intervention, waitlist controls or other intervention)?**

Yes
No
Uncertain

5. **Is the report a quasi-randomised study (with a control group that is no intervention, waitlist controls or other intervention)?**
   Yes
   No
   Uncertain

6. **Is the report a non-experimental design (with a control group that is no intervention, waitlist controls or other intervention)?**
   Yes
   No
   Uncertain

7. **Is the study?**
   Included
   Excluded
   Uncertain (state reason)

7.2 DATA EXTRACTION

Study design

1. **How were comparison/control groups formed?**
   Random assignment
   Other (specify)

2. **If random assignment, specify design**
   Simple/systematic (individuals/families)
   Stratified/blocked (identify stratifying variables)
   Yoked pairs (created by timing of enrolment into the study)
   Matched pairs (identify matching variables)
   Cluster (group) randomized
   Other (specify)
   Can't tell

3. **If non-random assignment, specify how groups were formed**
   Quasi randomization
   Time difference
   Location difference
   By action of researcher
   By action of therapist
   By participant preferences
   Other (specify)
   Can't tell

4. **Who performed group assignment**?
   Research staff
   Clinical staff
   Other (specify)
   Can't tell

5. **How was random assignment performed?**
   Computer generated
   Random numbers table
   Coins or dice
   Other (describe)
   Can't tell

6. **How many separate sites were included in the study?**
   One
   Two
   Three
   Specify number

7. **Was random assignment performed in the same way in all sites**
   Yes
   No (explain)
   Can't tell

8. **How many intervention groups were there? (12-step counts as one)**
   One (12-step)
   Two (12-step plus what?)
   Three (12-step plus what?)

9. **How many intervention groups are relevant for this review?**
   One (12-step)
   More than one (explain)

10. **How many *different* control/comparison groups were there? (i.e., groups that received different treatments, not counting multiple sites)**
    One
    Two or more (explain)

11. **How many control/comparison groups are relevant for this review?**
    One
    More than one (explain)

12. **Study sample size**

N's	12-step1*	CONTROL1*	TOTAL	Pg. # & NOTES
Referred to study				
Consented				
Completed base line measures				
Randomly assigned				

Or non-randomly allocated				
Started treatment				
Completed treatment				
Completed 1st follow up (.... months)				
Completed 2nd follow up (.... months) **				

* Add columns for additional intervention and control/comparison groups.
** Add rows as required for additional follow-ups.

Participant/sample Characteristics:

13. Was participant inclusion criteria mentioned?
No
Yes (describe & cite pg.#)

14. Was participant exclusion criteria mentioned?
No
Yes (describe & cite pg.#)

15. Participant Characteristics

	12-step*	CONTROL*	TOTAL	Pg. # & NOTES
Gender (e.g. % male)				
Youth Ages				
Race/ethnicity				
Socioeconomic status				
Profession				
Family composition				
Other characteristics				

* Add columns for additional intervention and control/comparison groups.

16. Specify and describe type of drug use
Cannabis
Cocaine
Amphetamine
Opiates
Combination (specify, pg. #)
Combination of drugs and alcohol (specify, pg. #)
Other (specify, pg. #)

17. Were there any differences between intervention and comparison groups at baseline?
No
Yes (describe differences & cite pg.#)
Unclear

18. Was there any analysis of differences between completers and dropouts in the intervention group?
No
Yes (describe differences & cite pg.#)
Unclear

19. Was there any analysis of differences between completers and dropouts in the control group?
No
Yes (describe differences & cite pg.#)
Unclear

Settings

20. Location of interventions (check all that apply)
Urban
Suburban
Rural
Can't tell

21. Location details
City
Country
Can't tell

12-step characteristics

22. Was the 12-step?
NA
Other self-help group (specify)
TSF (specify)

23. Other characteristics of 12-steps (e.g. duration, frequency, intensity, received treatment)

24. Characteristics of the self-help group (members, meetings, etc.)

25. Characteristics of TSF treatment staff (education, demographics, etc.)

Services provided to control cases

26. Type of control group

Waitlist (describe)
Alternative service (describe)
No service

27. Describe services provided to control group

28. Characteristics of staff who provided services to control cases (education, demographics, etc.)

Outcome measures

29. When were data collected (year/months)?

Start: _____

Finish: _____

30. Who conducted interviews?

Research staff
Clinical staff
Both
No interviews

31. Were data collected in the same manner for 12-step and control groups?

Yes
No (what were the differences?)
Can't tell

Outcome measures

Instructions: Please enter outcome measures in the order in which they are described in the report. Note that a single outcome measure can be completed by multiple sources and at multiple points in time (data from specific sources and time-points will be entered later).

OUTCOME & MEASURE	RELIABILITY & VALIDITY	FORMAT	DIRECTION	SOURCE	MODE ADMIN	Pg. # & NOTES
	Info from: • Other samples • This sample • Unclear Info provided:	• Dichotomous • Continuous	High score or event is: • Positive • Negative • Can't tell	• Researcher • Clinician • Admin data • Other • Unclear	• Self-admin • Interview • Other	

*Repeat as needed

OUT COME DATA

DICHOTOMOUS OUTCOME DATA

OUTCOME	TIME POINT (record exact time taken from baseline	SOURCE	VALID Ns	CASES	NON-CASES	STATISTICS	Pg. # & NOTES
	• 1st measure after baseline • 1st follow-up • 2nd follow-up • 3rd follow-up • 4th follow-up • Other	Refer to table, text or other where the results are visible.	12-STEP	12-STEP	12-STEP	RR (risk ratio) OR (odds ratio) SE (standard error) 95% CI ES	
			Comparison	Comparison	Comparison		

*Repeat as needed

CONTINUOUS OUTCOME DATA

Enter change and gain scores under Statistics (Other)

OUTCOME	TIME POINT (record exact time taken from baseline	SOURCE	VALID Ns	MEANs	SDs	STATISTICS	Pg. # & NOTES
	• 1st measure after baseline • 1st follow-up • 2nd follow-up • 3rd follow-up • 4th follow-up • Other	Refer to table, text or other where the results are visible.	12-STEP	12-STEP	12-STEP	P t F Df ES Covariates Other	
			Comparison	Comparison	Comparison		

*Repeat as needed

7.3 ASSESSMENT OF RISK OF BIAS IN INCLUDED STUDIES

Risk of bias table

Item	**Judgement**[a]	**Description** (quote from paper, or describe key information)
1. Sequence generation		
2. Allocation concealment		
3. Confounding[b,c]		
4. Blinding?[b]		
5. Incomplete outcome data addressed?[b]		
6. Free of selective reporting?[b]		
7. Free of other bias?		
8. A priori protocol?[d]		
9. A priori analysis plan?[e]		

[a] Some items on low/high risk/unclear scale (double-line border), some on 5 point scale/unclear (single line border), some on yes/no/unclear scale (dashed border). For all items, record "unclear" if inadequate reporting prevents a judgement being made.

[b] For each outcome in the study.

[c] This item is only used for QESs. It is based on list of confounders considered important at the outset and defined in the protocol for the review (*assessment against worksheet*).

[d] Did the researchers write a protocol defining the study population, intervention and comparator, primary and other outcomes, data collection methods, etc. in advance of starting the study?

[e] Did the researchers have an analysis plan defining the primary and other outcomes, statistical methods, subgroup analyses, etc. in advance of starting the study?

Risk of bias tool

Studies for which RoB tool is intended

The risk of bias model is developed by Prof. Barnaby Reeves in association with the Cochrane Non-Randomised Studies Methods Group.[10] This model, an extension of the Cochrane Collaboration's risk of bias tool, covers both risk of bias in randomised controlled trials (RCTs and QRCTs), but also risk of bias in non-randomised studies (QES).

The point of departure for the risk of bias model is the Cochrane Handbook for Systematic Reviews of interventions (Higgins & Green, 2008). The existing Cochrane risk of bias tool needs elaboration when assessing non-randomised studies because, for non-randomised studies, particular attention should be paid to selection bias / risk of confounding. Additional item on confounding is used only for non-randomised studies (QESs) and is not used for randomised controlled trials (RCTs and QRCTs).

Assessment of risk of bias

Issues when using modified RoB tool to assess included non-randomised studies:

- Use existing principle: score judgment and provide information (preferably direct quote) to support judgment
- QESs.
- 5-point scale for some items (distinguish "unclear" from intermediate risk of bias).
- Keep in mind the general philosophy – assessment is not about whether researchers could have done better but about risk of bias; the assessment tool must be used in a standard way whatever the difficulty / circumstances of investigating the research question of interest and whatever the study design used.
- Anchors: "1/No/low risk" of bias should correspond to a high quality RCT. "5/high risk" of bias should correspond to a risk of bias that means the findings should not be considered (too risky, too much bias, more likely to mislead than inform)

1. Sequence generation
   - Low/high/unclear RoB item
   - Always high RoB (not random) for a non-randomised study
   - Might argue that this item redundant for QES since always high – but important to include in RoB table ('level playing field' argument)

2. Allocation concealment
   - Low/high/unclear RoB item
   - Potentially low RoB for a non-randomised study, e.g. quasi-randomised (so high RoB to sequence generation) but concealed (reviewer judges that the people making decisions about including participants didn't know how allocation was being done, e.g. odd/even date of birth/hospital number)

3. RoB from confounding (additional item for QES; assess for each outcome)
   - Assumes a pre-specified list of potential confounders defined in the protocol
   - Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item

[10] This risk of bias model was introduced by Prof. Reeves at a workshop on risk of bias in non-randomized studies at SFI Campbell, February 2011. The model is a further development of work carried out in the Cochrane Non-Randomised Studies Method Group (NRSMG).

- Judgment needs to factor in:
    - proportion of confounders (from pre-specified list) that were considered
    - whether most important confounders (from pre-specified list) were considered
    - resolution/precision with which confounders were measured
    - extent of imbalance between groups at baseline
    - care with which adjustment was done (typically a judgment about the statistical modeling carried out by authors)
- Low RoB requires that all important confounders are balanced at baseline (not primarily/not only a statistical judgment) OR measured 'well' and 'carefully' controlled for in the analysis.

Assess against pre-specified worksheet. Reviewers will make a RoB judgment about each factor first and then 'eyeball' these for the judgment RoB table.

4. RoB from lack of blinding (assess for each outcome, as per existing RoB tool)
- Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
- Judgment needs to factor in:
    - nature of outcome (subjective / objective; source of information)
    - who was / was not blinded and the risk that those who were not blinded could introduce performance or detection bias
    - see Ch.8

5. RoB from incomplete outcome data (assess for each outcome, as per existing RoB tool)
- Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
- Judgment needs to factor in:
    - reasons for missing data
    - whether amount of missing data balanced across groups, with similar reasons
    - see Ch.8

6. RoB from selective reporting (assess for each outcome, NB different to existing Ch.8 recommendation)
- Low(1) / 2 / 3 / 4 / high(5) /unclear RoB item
- Judgment needs to factor in:
    - existing RoB guidance on selective outcome reporting
    - see Ch.8
    - also, extent to which analyses (and potentially other choices) could have been manipulated to bias the findings reported, e.g. choice of method of model fitting, potential confounders considered / included
    - look for evidence that there was a protocol in advance of doing any analysis / obtaining the data (difficult unless explicitly reported); QES very different from RCTs. RCTs must have a protocol in advance of starting to recruit (for REC/IRB/other regulatory approval); QES need not (especially older studies)
    - Hence, separate yes/no items asking reviewers whether they think the researchers had a pre-specified protocol and analysis plan.

Confounding Worksheet

Assessment of how researchers dealt with confounding	
Method for *identifying* relevant confounders described by researchers: yes no If yes, describe the method used:	☐ ☐
Relevant confounders described: yes no List confounders described on next page	☐ ☐
Method used for controlling for confounding At design stage (e.g. matching, regression discontinuity, instrument variable): .... At analysis stage (e.g. stratification, multivariate regression, difference-indifference): .... Describe confounders controlled for below	

Confounders described by researchers

Tick (yes[0]/no[1] judgment) if confounder considered by the researchers [Cons'd?]

Score (1[good precision] to 5[poor precision]) precision with which confounder measured

Score (1[balanced] to 5[major imbalance]) imbalance between groups

Score (1[very careful] to 5[not at all careful]) care with which adjustment for confounder was carried out

Confounder	Considered	Precision	Imbalance	Adjustment
Gender	☐	☐	☐	☐
Age	☐	☐	☐	☐
Socio-economic status	☐	☐	☐	☐
Mental problems	☐	☐	☐	☐
History of drug misuse	☐	☐	☐	☐
Unobservables[11]	☐	Irrelevant	☐	☐
Other:	☐	☐	☐	☐
Other:	☐	☐	☐	☐

[11] See user guide for unobservables

User guide for unobservables

Selection bias is understood as systematic baseline differences between groups and can therefore compromise comparability between groups. Baseline differences can be observable (e.g. age and gender) and unobservable (to the researcher; e.g. motivation and 'ability'). There is no single non-randomised study design that always solves the selection problem. Different designs solve the selection problem under different assumptions and require different types of data. Especially how different designs deal with selection on unobservables varies. The "right" method depends on the model generating participation, i.e. assumptions about the nature of the process by which participants are selected into a programme.

As there is no universal correct way to construct counterfactuals we will assess the extent to which the identifying assumptions (the assumption that makes it possible to identify the counterfactual) are explained and discussed (preferably the authors should make an effort to justify their choice of method). We will look for evidence that authors using e.g. (this is NOT an exhaustable list):

Natural experiments:
Discuss whether they face a truly random allocation of participants and that there is no change of behavior in anticipation of e.g. policy rules.

Instrument variable (IV):
Explain and discuss the assumption that the instrument variable does not affect outcomes other than through their effect on participation.

Matching (including propensity scores):
Explain and discuss the assumption that there is no selection on unobservables, only selection on observables.

(Multivariate) Regression:
Explain and discuss the assumption that there is no selection on unobservables, only selection on observables. Further discuss the extent to which they compare comparable people.

Regression Discontinuity (RD):
Explain and discuss the assumption that there is a (strict!) RD treatment rule. It must not be changeable by the agent in an effort to obtain or avoid treatment. Continuity in the expected impact at the discontinuity is required.

Difference-in-difference (Treatment-control-before-after):
Explain and discuss the assumption that outcomes of participants and nonparticipants evolve over time in the same way.