Drug and Alcohol REVIEW

APSAD

Drug and Alcohol Review (May 2020), 39, 394–414
DOI: 10.1111/dar.13051

REVIEW

How often are outcomes other than change in substance use measured? A systematic review of outcome measures in contemporary randomised controlled trials

JONE BJORNESTAD[1,2], JAMES R. McKAY[3,4], HENRIK BERG[5], CHRISTIAN MOLTU[2] & SVERRE NESVÅG[3,6]

[1]Department of Social Studies, Faculty of Social Sciences, University of Stavanger, Stavanger, Norway, [2]Department of Psychiatry, District General Hospital of Førde, Førde, Norway, [3]Centre for Alcohol and Drug Research, Stavanger University Hospital, Stavanger, Norway, [4]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA, [5]Centre for the Study of the Sciences and the Humanities, University of Bergen, Bergen, Norway, and [6]Department of Public Health, Faculty of Health Sciences, Stavanger, Norway

Abstract

Issues. *Recovery is a theoretical construct and empirical object of inquiry. The aim was to review whether outcome measures used in randomised controlled trials of drug treatment reflect a comprehensive conceptualisation of recovery.* ***Approach.*** *Systematic review using the following databases: Cochrane Database of Systematic Reviews, Cochrane Controlled Register of Trials, Database of Abstracts of Reviews of Effect, Web of Science, MEDLINE, Embase and PsycINFO. Search returned 6556 original articles and 504 met the following inclusion criteria: randomised controlled trial in English-language peer-reviewed journal; sample meets criteria for drug dependence or drug use disorder; reports non-substance use treatment outcomes. Review protocol registration: PROSPERO (CRD42018090064).* ***Key Findings.*** *3.8% of the included studies had a follow up of 2 years or more. Withdrawal/craving was present in 31.1% of short-term versus 0% of long-term studies. Social functioning in 8% of short-term versus 36.8% of long-term studies. Role functioning (0.9 vs. 26.3%), risk behaviour (15.6 vs. 36.8%) and criminality (3.8 vs. 21.1%) followed a similar pattern. Housing was not examined short-term and unregularly long-term (2.0%). 'Use of health-care facilities', clinical psychological, behavioural factors were frequently reported. Physiological or somatic health (15.2 vs. 10.5%), motivation (14.2 vs. 15.8%) and quality of life (7.1 vs. 0%) were less frequently reported.* ***Conclusion.*** *The short time interval of the follow up and lack of information on relevant factors in recovery prevents the development of evidence-based approaches to improve these factors. Particularly, measures of social and role functioning should be added to reflect an adequate conceptualisation of recovery.* [Bjornestad J, McKay JR, Berg H, Moltu C, Nesvåg S. How often are outcomes other than change in substance use measured? A systematic review of outcome measures in contemporary randomised controlled trials. *Drug Alcohol Rev* 2020;39:394–414]

Key words: substance use, substance use disorder, outcome measure, randomised controlled trial, systematic review.

Introduction

There is little consensus on the conceptualisation of long-term recovery in the drug use disorder (DUD) use literature. Recovery operationalisations influence treatment research, inform clinical practice and determine the efficacy or effectiveness of treatments and interventions. Thus, these operationalisations need to be valid to understand what is and what is not high-quality care. In severe mental illness, the operationalisation of

Jone Bjornestad PhD, Associate Professor and Researcher, James R. McKay PhD, Professor, Henrik Berg PhD, Researcher, Christian Moltu PhD, Professor, Sverre Nesvåg PhD, Principal Investigator, Centre for Alcohol and Drug Research and Professor II, Department of Public Health. Correspondence to: Associate Professor Jone Bjornestad, Department of Social Studies, Faculty of Social Sciences, University of Stavanger, PO Box 8600 FORUS, 4036 Stavanger, Norway. Tel: +47 97141599, E-mail: jone.r.bjornestad@uis.com

[Correction added on 22 April 2020, after first online publication: the 5th author's affiliation has been corrected.]

Received 13 June 2019; accepted for publication 11 February 2020.

recovery is more developed than in DUD [1]. Concrete operationalisation suggestions have been made (e.g. personal and clinical recovery), including functional and social aspects central to recovery in severe mental illness [2–4]. While specific factors, such as reduction in criminality, are more prominent in DUD recovery than in recovery from severe mental illness, general core factors, including an increase in community and social functioning, are common to these conditions [5–9]. The same applies for the reduction in core symptoms, for example substance use and severe psychiatric symptoms, as essential for achieving stable long-term recovery [10–12]. In this systematic review, we propose that conceptualisations of recovery from severe mental illness are applicable in DUD. Second, we systematically review to what extent substance use outcome measures used in randomised controlled trials (RCT) of drug treatment reflect a comprehensive understanding of recovery.

Clinical recovery traditionally refers to mental illness or DUD as distinct disorders displaying core symptoms. Clinical recovery is achieved when the core symptoms subside below diagnostic thresholds. Furthermore, the criteria for clinical recovery are based on researcher-derived thresholds and predefined objectives, including symptoms and functioning. Recovery also has a temporal criterion intended to indicate the stability of the recovery [4,13,14]. While subject to ongoing debate, a minimum duration of 2 years has been proposed. Two years allows for the possibility of new habits and behaviours to take hold, a relapse may have occurred or not, the maintenance of a drug-free social network has begun to consolidate, etc. [15–17]. There is more widespread agreement on symptom criteria for changes in drug use (i.e. use to abstinence or moderation) in the DUD literature [18,19]. However, consensus is lacking regarding criteria for functional and social recovery. Because of the extensive identity changes that are often considered necessary to handle a drug-free life, or even drug moderation, some have set a 5-year temporal criterion for DUD recovery [20–23].

The *personal recovery* tradition arose as a reaction to researcher-derived recovery criteria. Personal recovery is conceptualised beyond core symptom reduction as: '…a process of restoring a meaningful sense of belonging to one's community and positive sense of identity apart from one's condition while rebuilding a life despite or within the limitations imposed by that condition' [24,25]. Synthesising the research on personal recovery into an empirically based concept, Leamy *et al.* [26] outlined the Connectedness, Hope & Optimism, Identity, Meaning and Empowerment framework, in which five main long-term processes characterise recovery: (i) connectedness; (ii) hope and optimism; (iii) identity; (iv) meaning in life; and (v) empowerment. Empirical research suggests that these processes are relevant for DUD recovery [6,12,21,22].

The *relational recovery* tradition critiques the clinical and personal recovery approaches for not incorporating the interpersonal embeddedness of recovery [27]. This framework sees interpersonal contexts as permeating individualistic or subjective concepts like connectedness and self-agency [28], and advocates against conceptualising recovery as separate from the social and relational reality that partly defines the potentialities for each individual. These issues are just as relevant for DUD as for serious mental illness [29,30].

Though there are differences between these three approaches, the perspectives of clinical, personal and relational recovery share common ground [30]. Consistent with empirical findings, symptom reduction is seen as a necessary but insufficient requirement to maintain recovery over time. Although clinical recovery is unique in its definition of a concrete temporal criterion [15,16], recovery is universally described as a non-linear and cumbersome long-term growth process, with the threat of relapse often present. It is also acknowledged that a good outcome sometimes requires a long-term care effort [11–13,31]. Empirical support for these findings is solid and consistent across different clinical disciplines and research traditions [10,17,22,32–35]. On this basis, it is proposed that treatment outcome research in DUD should reflect these findings when assessing recovery.

The aim of this review was to systematically review and identify non-substance use (non-SU) treatment outcome measures used in RCTs on illicit drug use over the last 10 years, and to assess the degree to which they reflect any of the above-mentioned perspectives of recovery. RCTs were chosen because this methodology is generally considered the most valuable for both evaluating treatment efficacy and effectiveness and developing treatment guidelines.

Methods

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [36] to ensure comprehensive and transparent reporting of procedures and results. The protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) in March 2018 (registration number: CRD42018090064) (Appendix 1).

Search strategy

Two independent researchers (JB and SN) conducted a search of the literature using the following electronic

databases: Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effect, Web of Science, MEDLINE, Embase and PsycINFO. Variations and combinations of terms targeting five main concepts were used in the search: RCTs, substance abuse, substances, therapeutic approaches and recovery success. Subject headings belonging to the individual databases (e.g. MeSH subject terms) and free-text terms (see Appendix 2 for model search) were also used. The search queries were reviewed by an information scientist. In addition, a hand search was performed using reference lists from reviews and meta-analyses identified in the main search. In cases of doubt, the full-text paper was read to determine eligibility. Papers published between January 2008 and January 2019 were included. The last search was conducted on 11 January 2019.

Eligibility criteria

The included articles met the following criteria:

- Empirical study published in English-language peer-reviewed journal.
- Study sample meets the criteria for dependence syndrome (International Classification of Diseases, 10th revision) or moderate–severe DUD (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition).
- Randomised controlled trial.
- Reports non-SU treatment outcomes in addition to changes in substance use (e.g. social functioning, employment/school status, criminality, psychological symptoms).
- Empirical study from the past decade (2008–), as the recovery field has gained a more solid theoretical and empirical foundation during this time [1,4,11,26,27,35–37].

Exclusion criteria

Articles were excluded if the study sample was only or predominantly comprised of individuals with alcohol dependence, or if the study did not include non-substance use outcomes.

Data collection

All potential studies were exported into a reference citation manager (EndNote) before removing duplicates. Two independent reviewers (authors JB and SN) separately performed the screening of titles and abstracts, full-text analysis and selection of non-SU treatment outcome measures. Outcome categories (as presented in Tables 1–3) were developed during 13 consensus meetings (≈60 min each, JB and SN) and existing taxonomies as given below. Disagreements were resolved through discussion until consensus was reached. A third reviewer (JRM) was available to resolve disagreements and provide critical evaluation.

Analytic methods and data extraction procedure

A narrative descriptive synthesis was performed for the included articles. The qualitative synthesis was used to determine the taxonomy of non-SU outcomes. We used the suggested taxonomies of Dodd *et al.* [71], Bray *et al.* [72] and Shorter *et al* [73] as our basis for the synthesis. Dodd *et al.* was chosen as their standardisation includes flexible categories, applicable for general dimensions that emerge across conditions, such as functioning [71]. Bray *et al.* [72] and Shorter *et al.* [73] were used to adapt the categorisation specifically to DUD. Where we could not find normative taxonomies covering outcomes satisfactory, or we assessed factors as particularly relevant and specific for DUD (e.g. criminality), we used the study authors' outcome operationalisations as a compass for developing categories. In this context the following data extraction procedure was used: first, non-SU treatment outcome measures across different domains (e.g. work, community functioning, social functioning, health behaviour) were identified. Second, the properties of each outcome measure were analysed and categorised based on similarity (e.g. hepatitis C and HIV related to risky sexual behaviour were both organised under the 'Risk behaviour' tab in Table 1).

Contemporary recovery perspectives address issues of functioning (e.g. community and social), incorporate various perspectives on outcome (e.g. service user and researcher perspectives) and are explicit that a long-term perspective is crucial particularly with regards to functional recovery [11,12,29,30]. Since research on recovery has been growing over the past 10 years, this became a central rationale for the time limitation in our search—to test whether the DUD field had incorporated this shift in focus, from symptom relief (typically some measure of change in substance use), to more explicitly addressing function and social factors as important outcome measures.

For the same reasons, the second part of the synthesis was a pre-planned sub-analysis to identify long-term studies using non-substance use outcomes. Here, cut-off was set to studies with a follow up of at least 2 years, following Lieberman's criteria of stable recovery [2]. Also, the temporal criterion was set to 2 years, as this is suggested as the temporal requirement for recovery in

Table 1. *Non-substance use outcomes used in the contemporary drug use disorder trial literature*

	% (*n*)						
	Follow up, weeks					Time period	
	<13	13–26	27–52	53–103	Min. 2-years follow up	2008–2013	2014–2018
Total, *n* = 504							
No. studies	42.1 (212)	29.6 (149)	21.8 (110)	2.8 (14)	3.8 (19)	51 (257)	49 (247)
Outcomes							
Adverse effects							
Sub-categories: withdrawal/cravings	31.1 (66)	13.4 (20)	6.4 (7)	0	0	21.0 (54)	15.8 (39)
Psychological/behavioural factors							
Clinical							
Sub-categories: Psychological symptoms, psychiatric diagnosis, adverse events, cognitive functioning	41.0 (87)	38.3 (57)	33.6 (37)	64.3 (9)	31.6 (6)	35.0 (90)	42.9 (106)
Motivation							
Sub-categories: Self-efficacy, coping, readiness to change, treatment satisfaction, satisfaction/reward	14.2 (30)	10.7 (16	12.7 (14)	21.4 (3)	15.8 (3)	15.2 (39)	10.9 (27)
Physiological/clinical (somatic)							
Sub-categories: Physical health, seizure, pain experience, adverse effects, tolerability, insomnia, sleep	19.8 (42)	16.8 (25)	19.1 (21)	21.4 (3)	10.5 (2)	17.9 (46)	23.1 (57)
Housing	0	2.0 (3)	3.6 (4)	21.4 (3)	5.3 (1)	2.3 (6)	2.0 (5)
Functioning							
Social functioning							
Sub-categories: Community functioning, social relations and support, friend relationships, parenting, psychosocial, personal finance, aggression	8.0 (17)	12.1 (18)	15.5 (17)	14.3 (2)	36.8 (7)	11.7 (30)	12.6 (31)
Use of health-care facilities							
Sub-categories: Hospital days, treatment adherence/compliance, medication adherence (including therapeutic naltrexone levels), use of community-based organisation, retention, attrition, days to readmission, drop out, alliance	26.9 (57)	43.6 (65)	37.3 (41)	42.9 (6)	52.6 (10)	42.8 (110)	27.9 (69)
Role functioning							
Sub-categories: Taxed job, school participation, disability days	0.9 (2)	6.7 (10)	8.2 (9)	14.3 (2)	26.3 (5)	7.0 (18)	4.1 (10)
Global functioning							
Sub-categories: Composite of several functioning domains; employment, medical, psychiatric, legal, family/social, drugs, alcohol, etc.(typically based on composite scores from ASI, GAF, GSI)	6.1 (13)	12.8 (19)	12.7 (14)	21.4 (3)	10.5 (2)	9.7 (25)	10.5 (26
Quality of life							
Sub-categories: Quality of life measures, health related quality of life	7.1 (15)	12.8 (19)	12.7 (14)	28.6 (4)	0	9.4 (24)	10.5 (26)
Criminality	3.8 (8)	7.4 (11)	13.6 (15)	14.3 (2)	21.1 (4)	7.4 (19)	8.5 (21)
Risk behaviour							
Sub-categories: Risk behaviour (e.g. engaging in behaviours which increase the risk of hepatitis C, HIV or overdose)	15.6 (33)	24.2 (36)	22.7 (25)	14.3 (2)	36.8 (7)	25.7 (66)	15.4 (38)
Studies including only one non-substance use outcome in addition to change in substance use						41.2 (106)	55.1 (136)

ASI, Addiction Severity Index; GAF, Global Assessment of Functioning; GSI, Global Severity Index.

the clinical recovery literature [15,17]. Acknowledging the debate in this area, and some researchers advocating a temporal criterion up to 5 years [20–23], our 2-year criteria can primarily be viewed as a practical tool and as a minimum criterion to identify long-term studies. Finally, descriptive statistics were generated, aimed at summarising and quantifying significant treatment effects across studies.

Results

Search results

The electronic search returned 6556 articles. After duplicates were removed, 4545 articles remained. A hand search of reference lists from reviews and meta-analyses returned a further 21 articles. Full-text evaluation was conducted for 761 articles, of which 504 met the inclusion criteria and were included in the final synthesis. Details of the search results are summarised in Figure 1. Since the number of screened and included articles was extensive, it was necessary to develop superordinate categories (e.g. social functioning). Seven non-SU outcome categories and seven sub-categories were developed.

Non-SU outcome measures

Details of the included non-SU outcomes are summarised in Table 1 (see Appendix 3 for substance use measures used in the included articles). The five most frequently included outcomes were: clinical factors (from the category psychological/behavioural factors) (n = 196); use of healthcare facilities (from the category functioning) (n = 179); risk behaviour (n = 104); physiological/clinical (somatic) (n = 103); and withdrawal/cravings (from the category adverse effects) (n = 93). The five least frequently included outcomes were: housing (n = 11); role functioning (from the category functioning) (n = 28); criminality (n = 40); global functioning—mostly community-related functioning (from the category functioning) (n = 51); and quality of life (from the category functioning) (n = 51). In comparison, all studies had at least one DUD measure, which was also almost always reported as an outcome. Substance use outcome measures were spread across 22 different subcategories (e.g. days of drug use last month, substance use problems past 90 days, illicit opiate use).

Follow-up duration

From the included 504 research studies, 42.1% had less than 13 weeks of follow up, 29.6% had between 13 and 26 weeks, 21.8% had between 27 and 52 weeks, 2.8% had between 53 and 103 weeks and 3.8% had at least 2 years of follow up. The longest follow up was 416 weeks.

Relation between length of follow up and non-SU outcomes included

The most evident differences in non-SU outcome inclusions emerged between studies with less than 13 weeks of follow up and studies with at least 2 years of follow up (see Table 1). A measure of withdrawal/craving was present in 31.1% of the short-term versus 0% of the long-term studies. A reverse pattern was demonstrated with measures of social functioning, which were present in 8% of the short-term studies versus 36.8% of the long-term studies. Measures of role functioning (0.9 vs. 26.3%), risk behaviour (15.6 vs. 36.8%) and criminality (3.8 vs. 21.1%) followed a similar pattern. Housing was not examined in studies with short-term follow up, and only examined in one with long-term follow up (5.3%). 'Use of health-care facilities' was frequently reported across follow-up duration categories. Here, however, the greatest difference was again between follow ups of less than 13 weeks and greater than 1 year (26.9 vs. 52.6%). Clinical psychological and behavioural factors were generally frequently reported (41.0 vs. 31.6%). Physiological or clinical (somatic) health (15.2 vs. 10.5%), motivation (14.2 vs. 15.8%) and quality of life (7.1 vs. 0%) showed similar patterns, but with substantially lower percentages. More studies with only one outcome in addition to change in substance use were found between 2014 and 2019 (55.1%) than 2008 and 2013 (41.2%).

Long-term interventions and reported effects on DUD and non-SU outcomes

Table 2 displays details on studies with follow ups of between 1 and 2 years, and Table 3 presents details on studies with at least 2 years of follow up. Reported treatment effects are also presented. Slightly over two-thirds (69.7%) of the studies evaluated what may be termed complex interventions, which were primarily treatment programs with multiple components or several treatments/treatment elements merged together. Conversely, slightly less than one-third (30.3%) of the studies evaluated more narrowly focused interventions, usually single treatments such as cognitive behavioural therapy or targeted HIV-prevention programs. Ten percent of the studies showed a positive effect on DUD outcomes but no effect on non-SU outcomes. Conversely, 6.7% had a

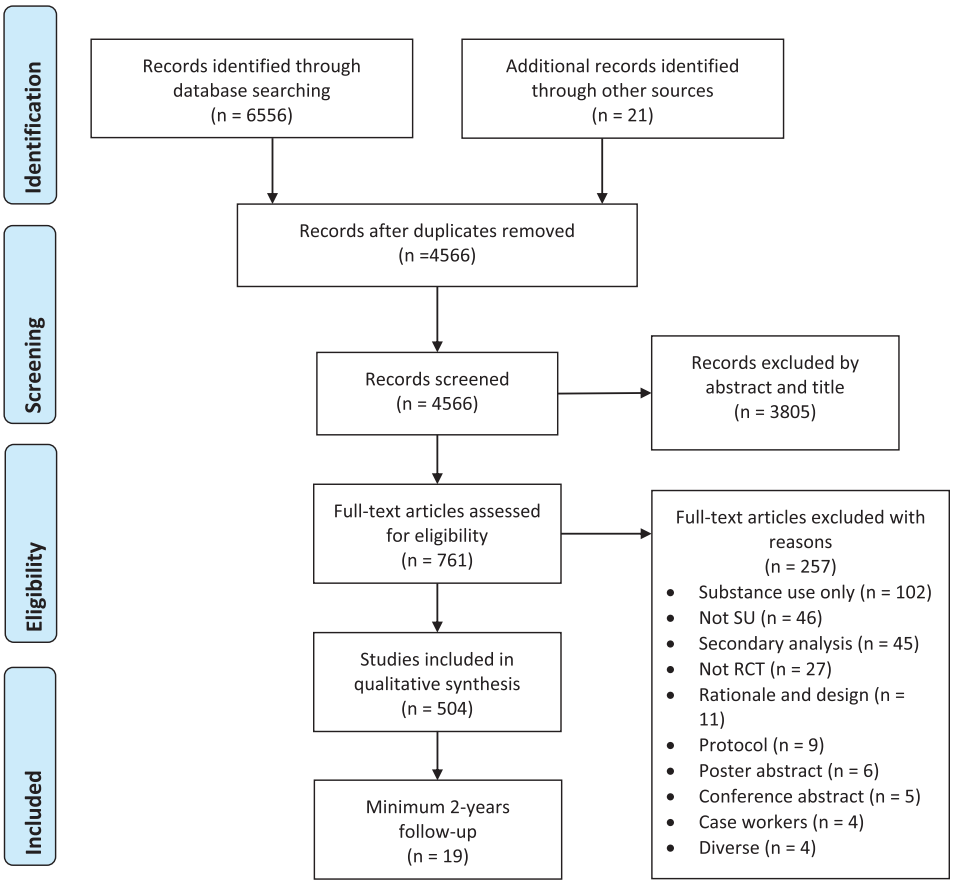

Figure 1. *Flow diagram of the reviewing process according to PRISMA. RCT, randomised controlled trial; SU, substance use.*

positive effect on non-SU outcomes but no effect on DUD outcomes. In total, 57.6% of the studies showed a significant positive effect on at least one of the non-SU outcomes examined during the intervention period and/or during follow up. Slightly more than half of the studies (54.6%) had at least one significant positive effect on DUD outcome, and 42.4% had a significant positive effect on at least one non-SU outcome and at least one DUD outcome, indicating a more general positive recovery effect.

Discussion

New agendas for contemporary recovery research

This review reveals that only a limited number of RCTs have been conducted using non-SU factors as treatment outcomes over time. Only 19 of the 504 included studies (3.8%) had follow ups of at least 2 years. Of these, 11 studies (2.2%) had follow ups of longer than 2 years. Given the suggested temporal criterion of a minimum of 2 years' follow up for recovery, this finding alone suggests that the substance use RCT treatment literature from the past decade only reflects the above-mentioned perspectives of clinical, personal and relational recovery to a very limited degree [1,26,27,31]. Focus on functional and social recovery are prominent in all these perspectives. Functional and social recovery are non-linear and cumbersome processes that usually require more time than that required to achieve abstinence [11–13,31]. The threat of relapse may continue for years following the achievement of abstinence [5–9]. Hence, contemporary substance use RCT research may omit important social recovery factors and processes, including loneliness, social alienation and the pursuit for citizenship [2,8,29,30]. When poorly handled, these factors are related to poor course development and relapse. Conversely, when overcome, they facilitate personal growth, perceived agency and social inclusion, possibly making the hard work of recovery attractive and seen as a realistic life solution over time [12]. Further, the ways in which people strengthen and maintain functional outcomes over time, such as increased school participation or more frequent social meetings [20,21,26], are difficult to understand, given the current evidence base. This requires a longitudinal study design and focused mediation analyses, which are usually beyond the scope of most RCTs.

Table 2. *Characteristics and reported findings of studies with 53–103 weeks follow up* (n = *14*)

Author, publication year, country	Sample, substance type	Follow up (weeks)	Experiment condition	Control condition	Non-SU outcomes	DUD outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
Litt *et al.* (2013), USA [38]	Recruited through newspaper advertisement. Adults. *n* = 215. Marijuana	61	Three conditions: (i) Intended to enhance self-efficacy (motivational enhancement plus cognitive-behavioural treatment plus contingency management reinforcing); (ii) Same condition plus reinforcing drug abstinence completion of treatment homework; (iii) Case management control condition. COMPLEX	See experiment condition	(i) Psychological/ behavioural factors (motivation); (ii) Psychological/ behavioural factors (clinical); (iii) Functioning (composite scores); (iv) Functioning (use of health-care facilities)	Days abstinent during 90-days time period	No clear treatment condition main effects. A subsample of patients with significant increase in self-efficacy did very well over time. This subsample was more likely to have been treated in the 'case management abstinence' condition. This treatment effect appears to have been accounted for by days of continuous abstinence accrued during treatment, and by pre–post increases in self-efficacy.	See non-SU findings
Ohana (2016), Israel [39]	Detoxification program enriched with intensive psychosocial interventions and aftercare. Adults. *n* = 121. Poly use	64	Dehydroepiandrosterone (DHEA biochemical substance) group. SPECIFIC	Placebo group	(i) Psychological/ behavioural factors (clinical); (ii) Functioning (use of health-care facilities); (iii) Quality of life	Drug use during study period	While in treatment, the experiment group reduced negative affect on the Positive and Negative Affect Scale. DHEA treatment	At 16-month follow up, re-use rates in the experiment condition were about one-third, compared with placebo

(Continues)

Table 2. (Continued)

Author, publication year, country	Sample, substance type	Follow up (weeks)	Experiment condition	Control condition	Non-SU outcomes	DUD outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
							resulted in an increase in DHEA sulphate 1 month following treatment, and level of DHEA-S predicted relapse in the follow-up assessment.	
Eack (2016), USA [40]	Psychiatric Institute and Clinic. Individuals with schizophrenia who were receiving outpatient services and had moderate or higher addiction severity scores. Adults. n = 31. Poly use	72	Cognitive enhancement therapy (n = 22). SPECIFIC	Usual care (n = 9)	Psychological/ behavioural factors (clinical)	Substance use problem severity past month	No significant between-group differences on non-SU outcomes	Substance use days were highly variable and followed nonlinear trajectories. Patients receiving cognitive enhancement therapy were significantly less likely to use alcohol but not cannabis.
Slesnick (2016), USA [41]	Mothers who sought family systems therapy outpatient treatment through a local substance use treatment facility. Adults. n = 183. Poly use	72	Women's Health Education. Office therapy or home-based family systems therapy (n = 123). COMPLEX	Women's Health Education, mothers only (n = 60)	Functioning (social functioning)	Substance problems past 90 days	No treatment effects were found for parent–child interactions	Women assigned to family systems therapy showed more rapid decline in alcohol, marijuana and cocaine use

(Continues)

Table 2. (Continued)

Author, publication year, country	Sample, substance type	Follow up (weeks)	Experiment condition	Control condition	Non-SU outcomes	DUD outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
Zhang (2018), USA [42]	Mothers who had at least one biological child in their care. Mothers were recruited from a community treatment centre for substance use in a large Midwestern city. Adults. n = 183. Poly use	72	Three conditions: (i) EBFT, home (n = 62); (ii) EBFT, office (n = 61); (iii) Women's Health Education (n = 60). COMPLEX	See experiment condition	(i) Criminality; (ii) Functioning (use of health-care facilities)	Substance use problems past 90 days	Mothers in the family therapy conditions had better psychological control. Children with mothers who showed decreased substance use and psychological control exhibited lower levels of problem behaviours compared to children with mothers showing increased substance use and psychological control	Mothers in the family therapy conditions were more likely to show reduced substance use
Milby *et al.* (2008), USA [43]	Birmingham Health Care. Agency serving homeless persons. Adults. n = 206. Cocaine	76	Full treatment model; contingency-managed housing, vocational training, and work therapy plus extensive behavioural day treatment. COMPLEX	Abbreviated treatment model: contingency-managed housing, vocational training, and work therapy	Functioning (use of health-care facilities)	(i) Prevalence of abstinence for the group at a point in time; (ii) Overall abstinence a period; (iii) Ability to initiate and sustain abstinence	No significant between-group differences on 'use of health services compliance'	Full treatment model produced better sustained abstinence outcomes

(Continues)

Table 2. (Continued)

Author, publication year, country	Sample, substance type	Follow up (weeks)	Experiment condition	Control condition	Non-SU outcomes	DUD outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
Dakof *et al.* (2010), USA [44]	Mothers recruited from the drug court. Adults. n = 62. Poly use	78	Engaging Moms drug court program. COMPLEX	Usual drug court care	Functioning (use of health-care facilities)	Number of days of substance use	Significant time effects for both intervention groups on mental health, parenting practices and family functioning. There were no significant differences between the experiment group and controls on non-SU outcomes	Significant time effects for both intervention groups' substance use outcomes. No significant differences between the experiment group and controls on DUD outcomes
McKay *et al.* (2010), USA [45]	Intensive outpatient programs at the time of entrance to treatment. Adults. n = 100. Cocaine.	78	Three conditions: (i) Cognitive–behavioural RP; (ii) CM; (iii) Combination of RP and CM COMPLEX	See experiment condition	(i) Psychological/ behavioural factors (clinical); (ii) Functioning (composite scores); (iii) Functioning (use of health-care facilities)	Abstinence from any cocaine use within each 3-month segment of follow up	No significant between-group differences on non-SUB variables	Significant effect favouring CM, but on cocaine use, not RP. Best outcome in the CM + RP condition
Milby *et al.* (2010), USA [46]	Birmingham Health Care. Agency serving homeless persons. Adults. n = 103. Cocaine	78	Abstinence-contingent housing, vocational training, and work plus cognitive behavioural day treatment. COMPLEX	Abstinence-contingent housing, vocational training and work	(i) Housing; (ii) Functioning (role)	Largest number of consecutive weeks abstinent over the first 52 weeks of study participation	No significant between-group effect differences on non-SU outcomes	Consecutive weeks of abstinence during treatment predicted long-term housing and employment stability

(Continues)

Table 2. (Continued)

Author, publication year, country	Sample, substance type	Follow up (weeks)	Experiment condition	Control condition	Non-SU outcomes	DUD outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
Tobin *et al.* (2011), USA [47]	Self-reported injection drug users. Adults. n = 227. Poly use	78	Peer-based, personal risk network-focused HIV prevention intervention. SPECIFIC	Five group-based sessions addressing injection drug use-related topics	Risk behaviour	Frequency of use of these paraphernalia in the past 6 months. All injection use	Increase in condom use among females in the experiment group	Reduced injection risk in the experiment group
González-Menéndez (2014), Spain [48]	Female inmates from state prison. Adults. n = 37. Poly use	78	ACT. SPECIFIC	CBT	(i) Psychological/ behavioural factors (motivation); (ii) Psychological/ behavioural factors (clinical); (iii) Functioning (composite scores)	Composite scores and information about days of drug or alcohol use in the intervention period	Percentages of mental disorders were reduced only in ACT participants. No other non-SU between-group differences	ACT was better than CBT in the maintaining of abstinence rates
Schottenfeld *et al.* (2008), Malysia [49]	Outpatient research clinic and detoxification program. Adults. n = 126. Heroin	95	Three conditions: (i) Manual-guided drug counselling and maintenance treatment with naltrexone; (ii) Manual-guided drug counselling and maintenance treatment with buprenorphine; (iii) Manual-guided drug counselling and maintenance treatment with placebo. COMPLEX	See experiment condition	(i) Risk behaviour; (ii) Physiological/ clinical (somatic)	(i) Days to first heroin use; (ii) Days to heroin relapse; (iii) Maximum consecutive days of abstinence from heroin	HIV risk behaviours decreased significantly from baseline for all three groups, primarily driven by substantial reductions in injection drug use. HIV risk behaviours did not differ significantly between conditions	Buprenorphine was associated with greater time to first heroin use than naltrexone
Messina *et al.* (2012), USA [50]	Female offenders participating in four drug	96	Gender-responsive programs using Helping Women Recover and	Standard mixed-gender treatment	(i) Psychological/ behavioural factors		Better in-treatment performance, more positive	Both groups reported non-significant reductions in drug

(Continues)

Table 2. (Continued)

Author, publication year, country	Sample, substance type	Follow up (weeks)	Experiment condition	Control condition	Non-SU outcomes	DUD outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
	court programs. Adults. n = 94. Poly use		Beyond Trauma. COMPLEX		(clinical); (ii) Psychological/ behavioural factors (motivation); (iii) Criminality		perceptions related to their treatment experience, and trends indicating reductions in PTSD symptomology in the experiment condition. Both groups improved in their self-reported psychological well-being and arrest records	use. No between-group differences
Saal (2016), Germany [51]	Four inpatient rehabilitation departments. Unemployed patients. Adults. n = 179. Poly use	96	Employment-focused case management. COMPLEX	Standard care	(i) Housing; (ii) Functioning (role); (iii) Functioning (use of health-care facilities); (iv) Quality of life	Substance use past month	No significant between-group differences on return to work rates, satisfaction with life, housing and precarious financial situation. Experiment condition was more effective on linking SUD patients with federal employment agency or job centre services	No significant between-group differences on drug abstinence

COMPLEX refers to complex intervention program. SPECIFIC refers to specific intervention program. ACT, acceptance and commitment therapy; CBT, cognitive behavioural therapy; CM, contingency management; DUD, drug use disorder; EBFT, ecologically based family therapy; non-SU, non-substance use; PTSD, post-traumatic stress disorder; RP, relapse prevention; SUD, substance use disorder.

Table 3. *Characteristics and reported findings of studies of at least of 2-years follow up* (n = *19)*

Author, publication year, country	Sample, substance type	Follow-up (w)	Experiment condition	Control condition	Non-SU outcomes	DUD-outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
Morgenstern *et al.* (2009), USA [52]	Women receiving temporary assistance for needy families. Adults. *n* = 302. Poly use	104	Intensive case management. COMPLEX	Screen-and-refer program	Functioning (role)	Monthly rates of absolute abstinence (daily screening)	Employment rate greater for the 'intensive case management' group	Abstinence rates were higher for the intensive case management group
Barrowclough *et al.* (2010), UK [53]	Secondary care. Persons with combined psychosis and SUD. Adults. *n* = 327. Poly use	104	Integrated motivational interviewing and CBT plus standard care. COMPLEX	Standard care	(i) Psychological/ behavioural factors (motivation); (ii) Psychological/ behavioural factors (clinical); (iii) Functioning (composite scores); (iv) Functioning (use of health-care facilities)	Number of days of substance use	Readiness to change use was higher in the experiment group at 12 months. This effect was not maintained at 24 months. No significant effects on any other non-SU outcomes	Amount of drugs used per substance use day was lower in the experiment group
Carpenedo *et al.* (2010), USA [54]	Large, urban, non-profit, free-standing methadone maintenance treatment centre. Adults. *n* = 130. Cocaine-abusing or dependent	104	Long-term voucher-based reinforcement therapy. SPECIFIC	Short-term voucher-based reinforcement therapy	Functioning (use of health-care facilities)	Number of days of cocaine use	No significant between-group differences in follow-up retention	Longer durations of self-reported continuous abstinence in the experiment group. This effect was not maintained during year 2
Oviedo-Joekes *et al.* (2010), Spain [55]	Granada Penitentiary Centre. Chronic-opioid-dependent people with severe drug-related health problems. Adults. *n* = 62. Poly use	104	Injected diacetylmorphine (diacetylmorphine [pharmaceutical grade heroin]. Three subgroups in relation to their HAT history: (i) Currently on HAT; (ii) Discontinued HAT; (iii) Never received HAT. SPECIFIC	Oral methadone	(i) Physiological/ clinical (somatic); (ii) Functioning (use of health-care facilities)	Number of days of substance use	Those currently in heroin-assisted treatment were the only group that had sustained at 2 years their marked improvement in physiological health after 9 months of treatment during the trial period	Illicit heroin use had a significant decrease in all three groups from baseline to 2 years post trial
Go (2015), Vietnam [56]	455 HIV-infected injectors and 355 of their HIV-negative injecting network members living in 32 sub-districts in Thai Nguyen Province. Adults. Poly use	104	Enhanced individual level post-test counselling and group support sessions vs. standard care HIV testing and counselling. This resulted in four arms: (i) Standard of care; (ii) Community-level intervention; (iii) Individual-level intervention; (iv) Community plus individual intervention. COMPLEX	See experiment condition	(i) Risk behaviour; (ii) Functioning (social functioning)	No outcomes presented for DUD outcomes	Fewer participants reported sharing injecting equipment and unprotected sex from baseline to 24 months in all arms. No other significant effects on non-SU outcomes and no significant differences between arms	No outcome analyses on that variable presented
Cather *et al.* (2018), USA [57]	Co-ordinated specialty care service program (NAVIGATE) for people with comorbid psychosis and SUD.	104	Received NAVIGATE (*n* = 223). COMPLEX	Usual care (*n* = 181)	Functioning (use of health-care facilities)	Substance use past 30 days	See DUD outcomes	No treatment group by time interaction effect on days of self-reported substance use over the two-year follow up. Participant exposure to the substance use

(Continues)

Table 3. (Continued)

Author, publication year, country	Sample, substance type	Follow-up (w)	Experiment condition	Control condition	Non-SU outcomes	DUD-outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
	Adults. *n* = 404. Poly use							component of the NAVIGATE program was low, suggesting that modifications to the program and training method for clinicians may be needed.
Hoffman *et al.* (2013), Russia [58]	Injection drugs users. Adults. *n* = 160. Poly use	104	Peer-educator network intervention as a strategy to reduce HIV acquisition and their drug and/or sexual networks. SPECIFIC	No peer-educator network intervention	Risk behaviour	No outcomes presented for DUD outcomes	No significant between-group differences on non-SU outcomes	No outcome analyses on that variable presented
Rotheram-Boris *et al.* (2009), USA [59]	HIV-infected marginally housed persons. Adults. *n* = 270. Poly use	108	Healthy Living Programme aimed at reducing risky sexual behaviour and substance use. COMPLEX	Passive control (did not receive intervention)	Risk behaviour	(i) Number of days marijuana used; (ii) Used hard drugs; (iii) Number of days hard drugs used	Reductions in unprotected risky sexual acts, number of sexual partners of HIV negative or unknown serostatus higher in the experiment group	Marijuana use and hard drug use lower in the experiment group
Schaeffer (2014), USA [60]	High-risk juvenile offenders. Adolescents (mean age = 15.8 years). *n* = 97. Poly use	120	Community Restitution Apprentice-Focused Training. COMPLEX	Education as usual	(i) Psychological/ behavioural factors (clinical); (ii) Functioning (role); (iii) Criminality	Substance problems past 30 days	Increased rate of youth employment and General Equivalency Diploma attendance in the experiment group. No effects for months employed, hours worked or hourly wage. Measures of mental health symptoms and criminal activity showed no favourable effects.	No between-group differences on substance use
DeFulio and Silverman (2011), USA [61]	Unemployed welfare recipients who persistently used cocaine while in methadone treatment. Adults. *n* = 51. Cocaine	130	Abstinence-contingent employment group that was required to provide cocaine and opiate-negative urine samples to work and maintain maximum rate of pay. COMPLEX	Work independently from drug use group	(i) Risk behaviour; (ii) Functioning (social functioning); (iii) Functioning (role); (iv) Criminality	Drug urine samples. Testing became random and progressively less frequent as abstinence was sustained. Control group had less strict rules.	No significant between-group differences on either non-SU outcomes during the employment year or follow up	The experiment group had significantly more cocaine-negative samples during the employment year. Differences were not maintained during follow up
Scott (2017), USA [62]	Women released from the Cook County Jail's Department of Women's Justice Services. Adults. *n* = 480. Poly use	144	Recovery management check-ups for two years. SPECIFIC	One assessment only	(i) Risk behaviour; (ii) Criminality; (iii) Functioning (use of health-care facilities)	Drug use during the study period	Higher rates of participation in substance use treatment, and recovery support services, less criminality and less HIV risk behaviours in the experiment group. These effects were strongest when participants were not on probation. No between-group differences when	The experiment condition effect favoured the control condition when women were not on probation

(*Continues*)

Table 3. (Continued)

Author, publication year, country	Sample, substance type	Follow-up (w)	Experiment condition	Control condition	Non-SU outcomes	DUD-outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
							participants were on probation	
Rawson *et al.* (2012), USA [63]	Outpatient psychosocial treatment project. Adults. n = 871. Methamphetamine	156	16-week Matrix model treatment. COMPLEX	Treatment as usual	(i) Psychological/ behavioural factors (clinical); (ii) Functioning (use of health-care facilities)	Negative urinalysis results at 12 months and 36 months after treatment	Retention rates and treatment engagement were higher in the experiment group	More negative urine analyses in the experiment group at 16 weeks
O'Connell *et al.* (2012), USA [64]	Recruited from Health Care for Homeless Veterans program. Adults. n = 259. Poly use	156	Three conditions: (i) ICM plus rent subsidy vouchers; (ii) ICM only; (iii) Treatment as usual. COMPLEX	See experiment condition	(i) Psychological/ behavioural factors (clinical); (ii) Functioning (social functioning); (iii) Housing; (iv) Functioning (role); (v) Functioning (use of health-care facilities); (6) Quality of life	Composite scores and information about days of drug or alcohol use in past 30 days, including days of intoxication	Compared with ICM alone, condition 1 was associated with more positive housing outcomes for Caucasians, veterans with co-occurring mental disorders, and veterans who were active substance users. Condition 1 was associated with more positive socio-clinical outcomes for African Americans. No differences were observed in housing or socio-clinical outcomes as a function of age	African Americans, who had significantly lower baseline scores on ASI psychiatric and alcohol subscales, reported fewer days of use of drugs than Caucasian individuals across all follow-up time periods. No between-group differences
van der Pol *et al.* (2018), Netherlands [65]	Sample for this study was enrolled in a Dutch randomised controlled trial conducted as part of a transnational trial (Germany, France, Belgium, Switzerland, and the Netherlands). Adolescents (13–18 years). n = 109. Cannabis and comorbid problem behaviour	156	MDFT. COMPLEX	Treatment as usual	(i) Functioning (social functioning); (ii) Criminality	Substance use past 90 days	No between group differences between the treatment groups for time to first offence from start of the treatment or changes in frequency or severity of offending over time. Treatment effect trend favouring MDFT was found for property offending in the subgroup of adolescents with high baseline severity of cannabis use.	No outcome analyses on that variable presented
Mueser *et al.* (2013), USA [66]	3 community mental health centres. Co-occurring substance use and severe psychiatric disorders. Adults. n = 108. Poly use	156	Long-term (9–18 months) program combining education with teaching communication and problem-solving skills. COMPLEX	Brief (2–3 months) family education program	(i) Psychological/ behavioural factors (motivation); (ii) Psychological/ behavioural factors (clinical); (iii) Functioning (social functioning); (iv) Functioning (composite	Days of drug use over the past 6 months	Experiment group patients had less severe overall psychiatric symptoms and psychotic symptoms and tended to improve more in functioning. Relatives in experiment condition improved more in mental health functioning and knowledge of co-occurring disorders. No	No consistent differences between the programs in substance use severity

(Continues)

Table 3. (Continued)

Author, publication year, country	Sample, substance type	Follow-up (w)	Experiment condition	Control condition	Non-SU outcomes	DUD-outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
					scores); (v) Functioning (use of health-care facilities)		between-group differences on family burden	
Suchman *et al.* (2011), USA [67]	Mothers from outpatient substance-abuse treatment and caring for children between birth and 3 years of age. Adults. *n* = 47. Poly use	162	Mothers and Toddlers Program. COMPLEX	Parent education program (individual case management and developmental guidance)	(i) Psychological/ behavioural factors (clinical); (ii) Functioning (social functioning)	Relapse to substance use	Better reflective functioning (link between parental representation and attachment patterns), representation quality, and caregiving behaviour in the experiment group. Less psychiatric distress in the experiment group. No group differences in overall quality of representations or children's interactions with their mothers	Mothers in both groups showed reduction in relapse rates
Wechsberg *et al.* (2010), USA [68]	Out-of-treatment African American women. Adults. *n* = 455. Poly use	208	Three conditions: (i) Female-focused intervention; (ii) Modified NIDA (standard) intervention; (iii) Delayed treatment control condition. SPECIFIC	See experiment condition	(i) Risk behaviour; (ii) Physiological/ clinical (somatic); (iii) Psychological/ behavioural factors (clinical); (iv) Functioning (social functioning); (v) Functioning (use of health-care facilities)	Number of days of substance use	No significant between-group differences on any non-SU measures over time. Results suggest effects that might be considered durable, resistant and eroding and a sleeper effect from high to low risk	No significant between-group differences over time on DUD outcomes
Dennis and Scott (2012), USA [69]	Treatment agency in Illinois. Adults. *n* = 446. Poly use	208	Quarterly recovery management check-ups. SPECIFIC	Quarterly assessment only	Functioning (use of health-care facilities)	(i) Substance frequency scale; (ii) Substance problems past month; (iii) Days of abstinence	Experiment group participants were more likely to return to treatment sooner, to return at all, to return more times, and to receive more total days of treatment. Significantly fewer quarters in need of treatment. Effects were larger (general) for those with earlier onset and higher crime/violence scores.	Fewer substance-related problems per month, and more total days of abstinence in the experiment group
Aklin (2014), USA [70]	Pregnant and postpartum women enrolled in methadone treatment. Adults. *n* = 40. Poly use	416	Therapeutic workplace. SPECIFIC	Usual care	(i) Risk behaviour; (ii) Functioning (role)	Substance use past 30 days	During year 4 when the business was open, therapeutic workplace participants reported more days employed, higher employment income, and less money spent on drugs. During	During year 4 therapeutic workplace participants provided significantly more cocaine- and opiate-negative urine samples than controls

(Continues)

Table 3. (Continued)

Author, publication year, country	Sample, substance type	Follow-up (w)	Experiment condition	Control condition	Non-SU outcomes	DUD-outcomes	Reported findings on non-SU outcomes	Reported findings on DUD outcomes
							the 3 years after the business closed, therapeutic workplace participants only reported higher income	

CBT, cognitive behavioural therapy; COMPLEX refers to complex intervention program; DUD, drug use disorder; HAT, heroin-assisted treatment; ICM, intensive case management; MDFT, multidimensional family therapy; NIDA, National Institute on Drug Abuse; non-SU, non-substance use; SPECIFIC refers to specific intervention program.

These limitations make it challenging for clinicians to work from an evidence base in their attempts to tailor phase-specific DUD treatment strategies for long-term recovery efforts.

In line with contemporary recovery research, the 3.8% of studies with a follow up of at least 2 years are more likely to report general health and recovery effects than studies with shorter follow ups. However, one limitation of these 19 studies is that they typically report the non-SU outcomes of psychological health (typically reduction in depression) and use of health-care facilities (typically treatment retention), but do not report on other non-SU outcomes. Only seven studies (1.4%) reported social functioning outcomes, five (1.0%) on role functioning, four (0.8%) on criminality, two (0.4%) on global functioning and zero studies on quality of life. The severely limited number of studies measuring these factors stands in contrast to the fact that they have consistently been associated with good and stable DUD outcomes in the recovery literature [10–12,31,74,75]. Moreover, conclusions that cut across different recovery traditions around what constitutes recovery—for example long-term increase in community and social functioning, along with reductions in or elimination of substance use [4,26,27]—are largely ignored. Likewise, the increasing trend of studies using only one non-SU outcome in addition to change in substance use (41.2% between 2008 and 2013 vs. 55.1% between 2014 and 2019) represents a step away from the longitudinal and multi-dimensional study approach required to investigate long-term recovery.

Limitations

The strengths of the study are evident in its protocol's public availability before the review was conducted (via PROSPERO); this ensured transparency and that the review was conducted according to PRISMA guidelines [36].

One limitation concerns the fact that no advanced statistical tests were used to assess the reliability and validity of the reported findings of the included studies. The scope of the paper was to evaluate outcome measures and not treatment efficacy, per se. Another limitation is that each individual study was not assessed for key sources of biases (e.g. sample characteristics). In addition, and in line with previous research, some studies were based on small samples, and most instruments were constructed and tested within Anglo-American cultures. This typically increases the risk of reporting bias, suggesting that the included studies represent selective research dissemination. However, it should be emphasised that the aim was to identify outcomes with a high level of use within the field and that the search

was conducted within several literature databases. The included studies did use samples with somewhat different characteristics (e.g. sex, age and level of symptomatology), which may violate the transitivity assumption and thus raises questions regarding the validity of direct comparisons across the included studies.

Suggested research directions

To improve the scientific knowledge base of treatment outcomes in DUD it will be advisable to incorporate functional and social outcomes into longitudinal research designs more consistently. These outcomes are already actively used by other initiatives, such as the Treatment Episode Dataset discharge data [76]. Empirical studies indicate that future research should focus on detailing the specific effects of social and community functioning in recovery. For example, we need to know more about which treatment interventions bring about sustained improvements in these areas, and which post-treatment factors mediate improvements in social and community functioning. In addition, a more valid temporal criterion that would enable professionals to more accurately identify vulnerable phases in recovery would be useful for tailoring treatment efforts towards expected fluctuations in relapse. A broad investigation should also aim to overcome specific limitations inherent in the RCT study designs, including sensitivity to contextual factors and comparison of single, common clinical metric across different study contexts. As suggested by Donovan *et al.* [77], applying within-study comparisons may be a more valid alternative to studying complex phenomena, such as recovery in DUD. Furthermore, systematic inclusion of service-user perspectives could prove a viable route to meet this aim [78]. By asking individuals with first-hand experience and those outside of the traditional scientific community for input in the research design, the risk of implementing measures with low ecological validity is considerably reduced [79]. In practical terms, the application of a mixed research design, combining exploration, hypothesis development and further large-scale testing (RCTs), could be a feasible solution.

Acknowledgements

A special thanks to the staff at the Medical Library of Stavanger University Hospital for assistance with the literature search.

Conflict of Interest

The authors have no conflicts of interest.

References

[1] Slade M, Amering M, Farkas M *et al.* Uses and abuses of recovery: implementing recovery-oriented practices in mental health systems. World Psychiatry 2014;13:12–20.
[2] Liberman RP, Kopelowicz A. Recovery from schizophrenia: a challenge for the 21st century. Int Rev Psychiatry 2002;14:245–55.
[3] Andreasen NC. Standardized remission criteria in schizophrenia. Acta Psychiatr Scand 2006;113:81.
[4] Davidson L, Schmutte T, Dinzeo T, Andres-Hyman R. Remission and recovery in schizophrenia: practitioner and patient perspectives. Schizophr Bull 2008;34:5–8.
[5] McLellan AT, Alterman AI, Metzger DS *et al.* Similarity of outcome predictors across opiate, cocaine, and alcohol treatments: role of treatment services. J Consult Clin Psychol 1994;62:1141–58.
[6] McLellan AT, Arndt IO, Metzger DS, Woody GE, O'Brlen CP. The effects of psychosocial services in substance abuse treatment. Addict Nurs Netw 1993;5:38–47.
[7] Woody GE, Luborsky L, McLellan AT *et al.* Psychotherapy for opiate addicts: does it help? Arch Gen Psychiatry 1983;40:639–45.
[8] O'Brien C, McLellan A. Myths about the treatment of addiction. Lancet 1996;347:237–40.
[9] Galea S, Nandi A, Vlahov D. The social epidemiology of substance use. Epidemiol Rev 2004;26:36–52.
[10] Tiffany ST, Friedman L, Greenfield SF, Hasin DS, Jackson R. Beyond drug use: a systematic consideration of other outcomes in evaluations of treatments for substance use disorders. Addiction 2012;107:709–18.
[11] Slade M, Leamy M, Bacon F *et al.* International differences in understanding recovery: systematic review. Epidemiol Psychiatr Sci 2012;21:353–64.
[12] McKay J. Making the hard work of recovery more attractive for those with substance use disorders. Addiction 2017;112:751–7.
[13] Davidson L. The recovery movement: implications for mental health care and enabling people to participate fully in life. Health Aff (Millwood) 2016;35:1091–7.
[14] American Psychiatric Association. Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition. Arlington, VA, American Psychiatric Association, 2013
[15] Hegelstad WT, Larsen TK, Auestad B *et al.* Long-term follow-up of the TIPS early detection in psychosis study: effects on 10-year outcome. Am J Psychiatry 2012;169:374–80.
[16] Liberman RP, Kopelowicz A, Ventura J, Gutkind D. Operational criteria and factors related to recovery from schizophrenia. Int Rev Psychiatry 2002;14:256–72.
[17] Jääskeläinen E, Juola P, Hirvonen N *et al.* A systematic review and meta-analysis of recovery in schizophrenia. Schizophr Bull 2013;39:1296–306.
[18] Calabria B, Degenhardt L, Briegleb C *et al.* Systematic review of prospective studies investigating 'remission' from amphetamine, cannabis, cocaine or opioid dependence. Addict Behav 2010;35:741–9.
[19] Fleury M-J, Djouini A, Huỳnh C *et al.* Remission from substance use disorders: a systematic review and meta-analysis. Drug Alcohol Depend 2016;168:293–306.
[20] El-Guebaly N. The meanings of recovery from addiction: evolution and promises. J Addict Med 2012;6:1–9.
[21] Hser Y-I, Evans E, Grella C, Ling W, Anglin D. Long-term course of opioid addiction. Harv Rev Psychiatry 2015;23:76–89.
[22] Buckingham SA, Frings D, Albery IP. Group membership and social identity in addiction recovery. Psychol Addict Behav 2013;27:1132–40.
[23] Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: applications to addictive behaviors. Am Psychol 1992;47:1102–14.
[24] Davidson L, Tondora J, O'Connell MJ, Kirk T Jr, Rockholz P, Evans AC. Creating a recovery-oriented system of behavioral health care:-moving from concept to reality. Psychiatr Rehabil J 2007;31:23–31.
[25] Anthony WA. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. Psychosoc Rehabil J 1993;16:11–23.
[26] Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011;199:445–52.
[27] Price-Robertson R, Obradovic A, Morgan B. Relational recovery: beyond individualism in the recovery approach. Adv Ment Health 2017;15:108–20.

[28] Bjornestad J, Bronnick K, Davidson L *et al*. The central role of self-agency in clinical recovery from first episode psychosis. Psychosis 2016; 9:140–8.
[29] Dingle GA, Stark C, Cruwys T, Best D. Breaking good: breaking ties with social groups may be good for recovery from substance misuse. Br J Soc Psychol 2015;54:236–54.
[30] Davidson L, Andres-Hyman R, Bedregal L, Tondora J, Frey J, Kirk TA Jr. From 'double trouble' to 'dual recovery': integrating models of recovery in addiction and mental health. J Dual Diagn 2008;4:273–90.
[31] Sobell L, Ellingstad TP, Sobell MB. Natural recovery from alcohol and drug problems: methodological review of the research with suggestions for future directions. Addiction 2000;95:749–64.
[32] Thoits PA. Mechanisms linking social ties and support to physical and mental health. J Health Soc Behav 2011;52:145–61.
[33] van Os J, Kenis G, Rutten BP. The environment and schizophrenia. Nature 2010;468:203–12.
[34] Davidson L, Stayner DA, Nickou C, Styron TH, Rowe M, Chinman ML. 'Simply to be let in': inclusion as a basis for recovery. Psychiatr Rehabil J 2001;24:375–88.
[35] Tew J, Ramon S, Slade M, Bird V, Melton J, Le Boutillier C. Social factors and recovery from mental health difficulties: a review of the evidence. Br J Soc Work 2012;42:443–60.
[36] Kaskutas LA, Borkman TJ, Laudet A *et al*. Elements that define recovery: the experiential perspective. J Stud Alcohol Drugs 2014;75: 999–1010.
[37] Hutton B, Salanti G, Caldwell DM *et al*. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations PRISMA extension for network meta-analysis. Ann Intern Med 2015;162:777–84.
[38] Litt MD, Kadden RM, Petry NM. Behavioral treatment for marijuana dependence: randomized trial of contingency management and self-efficacy enhancement. Addict Behav 2013;38:1764–75.
[39] Ohana D, Maayan R, Delayahu Y *et al*. Effect of dehydroepiandrosterone add-on therapy on mood, decision making and subsequent relapse of polydrug users. Addict Biol 2016;21:885–94.
[40] Eack S, Hogarty SS, Bangalore SS, Keshavan MS, Cornelius JR. Patterns of substance use during cognitive enhancement therapy: an 18-month randomized feasibility study. J Dual Diagn 2016;12:74–82.
[41] Slesnick N, Zhang J. Family systems therapy for substance-using mothers and their 8- to 16-year-old children. Psychol Addict Behav 2016;30: 619–29.
[42] Zhang J, Slesnick N, Feng X. Co-occurring trajectory of mothers' substance use and psychological control and children's behavior problems: the effects of a family systems intervention. Fam Process 2018;57:211–25.
[43] Milby JB, Schumacher JE, Vuchinich RE, Freedman MJ, Kertesz S, Wallace D. Toward cost-effective initial care for substance-abusing homeless. J Subst Abuse Treat 2008;34:180–91.
[44] Dakof GA, Cohen JB, Henderson CE *et al*. A randomized pilot study of the engaging moms program for family drug court. J Subst Abuse Treat 2010;38:263–74.
[45] McKay JR, Lynch KG, Coviello D *et al*. Randomized trial of continuing care enhancements for cocaine-dependent patients following initial engagement. J Consult Clin Psychol 2010;78:111–20.
[46] Milby JB, Schumacher JE, Wallace D, Vuchinich R, Mennemeyer ST, Kertesz SG. Effects of sustained abstinence among treated substance-abusing homeless persons on housing and employment. Am J Public Health 2010;100:913–8.
[47] Tobin KE, Kuramoto SJ, Davey-Rothwell MA, Latkin CA. The STEP into action study: a peer-based, personal risk network-focused HIV prevention intervention with injection drug users in Baltimore, Maryland. Addiction 2011;106:366–75.
[48] Gonzalez-Menendez A, Fernández MP, Rodríguez F, Lanza P. Long-term outcomes of acceptance and commitment therapy in drug-dependent female inmates: a randomized controlled trial. Int J Clin Health Psychol 2014;14:18–27.
[49] Schottenfeld RS, Chawarski MC, Mazlan M. Maintenance treatment with buprenorphine and naltrexone for heroin dependence in Malaysia: a randomised, double-blind, placebo-controlled trial. Lancet 2008;371: 2192–200.
[50] Messina N, Calhoun S, Warda U. Gender-responsive drug court treatment: a randomized controlled trial. Criminal Justice Behav 2012;39:1539–58.
[51] Saal S, Forschner L, Kemmann D, Zlatosch J, Kallert TW. Is employment-focused case management effective for patients with substance use disorders? Results from a controlled multi-site trial in Germany covering a 2-year period after inpatient rehabilitation. BMC Psychiatry 2016;16:279.
[52] Morgenstern J, Neighbors CJ, Kuerbis A *et al*. Improving 24-month abstinence and employment outcomes for substance-dependent women receiving temporary assistance for needy families with intensive case management. Am J Public Health 2009;99:328–33.
[53] Barrowclough C, Haddock G, Wykes T *et al*. Integrated motivational interviewing and cognitive behavioural therapy for people with psychosis and comorbid substance misuse: randomised controlled trial. BMJ 2010; 341:c6325.
[54] Carpenedo CM, Kirby KC, Dugosh KL, Rosenwasser BJ, Thompson DL. Extended voucher-based reinforcement therapy for long-term drug abstinence. Am J Health Behav 2010;34:776–87.
[55] Oviedo-Joekes E, March JC, Romero M, Perea-Milla E. The Andalusian trial on heroin-assisted treatment: a 2 year follow-up. Drug Alcohol Rev 2010;29:75–80.
[56] Go V, Frangakis C, Minh NL *et al*. Efficacy of a multi-level intervention to reduce injecting and sexual risk behaviors among HIV-infected people who inject drugs in Vietnam: a four-arm randomized controlled trial. PLoS One 2015;10:e0125909.
[57] Cather C, Brunette MF, Mueser KT *et al*. Impact of comprehensive treatment for first episode psychosis on substance use outcomes: a randomized controlled trial. Psychiatry Res 2018;268:303–11.
[58] Hoffman I, Latkin CA, Kukhareva PV *et al*. A peer-educator network HIV prevention intervention among injection drug users: results of a randomized controlled trial in St. Petersburg, Russia. AIDS Behav 2013;17: 2510–20.
[59] Rotheram-Boris MJ, Desmond K, Comulada WS, Arnold EM, Johnson M, Healthy Living Trial Group. Reducing risky sexual behavior and substance use among currently and formerly homeless adults living with HIV. Am J Public Health 2009;99:1100–7.
[60] Schaeffer CM, Henggeler SW, Ford JD, Mann M, Chang R, Chapman JE. RCT of a promising vocational/employment program for high-risk juvenile offenders. J Subst Abuse Treat 2014;46:134–43.
[61] DeFulio A, Silverman K. Employment-based abstinence reinforcement as a maintenance intervention for the treatment of cocaine dependence: post-intervention outcomes. Addiction 2011;106:960–7.
[62] Scott CK, Dennis ML, Lurigio AJ. The effects of specialized probation and recovery management checkups (RMCs) on treatment participation, substance use, HIV risk behaviors, and recidivism among female offenders: main findings of a 3-year experiment using subject by intervention interaction analysis. J Exp Criminol 2017;13:53–77.
[63] Rawson RA, Gonzales R, Greenwell L, Chalk M. Process-of-care measures as predictors of client outcome among a methamphetamine-dependent sample at 12-and 36-month follow-ups. J Psychoactive Drugs 2012;44:342–9.
[64] O'Connell MJ, Kasprow WJ, Rosenheck RA. Differential impact of supported housing on selected subgroups of homeless veterans with substance abuse histories. Psychiatr Serv 2012;63:1195–205.
[65] van der Pol TM, Hendriks V, Rigter H *et al*. Multidimensional family therapy in adolescents with a cannabis use disorder: long-term effects on delinquency in a randomized controlled trial. Child Adolesc Psychiatry Ment Health 2018;12:44.
[66] Mueser KT, Glynn SM, Cather C *et al*. A randomized controlled trial of family intervention for co-occurring substance use and severe psychiatric disorders. Schizophr Bull 2013;39:658–72.
[67] Suchman NE, Decoste C, McMahon TJ, Rounsaville B, Mayes L. The mothers and toddlers program, an attachment-based parenting intervention for substance-using women: results at 6-week follow-up in a randomized clinical pilot. Infant Mental Health J 2011;32:427–49.
[68] Wechsberg WM, Novak SP, Zule WA *et al*. Sustainability of intervention effects of an evidence-based HIV prevention intervention for African American women who smoke crack cocaine. Drug Alcohol Depend 2010;109:205–12.
[69] Dennis M, Scott C. Four-year outcomes from the early re-intervention (ERI) experiment using recovery management checkups (RMCs). Drug Alcohol Depend 2012;121:10–7.
[70] Aklin WM, Wong CJ, Hampton J *et al*. A therapeutic workplace for the long-term treatment of drug addiction and unemployment: eight-year outcomes of a social business intervention. J Subst Abuse Treat 2014;47:329–38.
[71] Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. J Clin Epidemiol 2018;96:84–92.

[72] Bray JW, Cowell AJ, Hinde JM. A systematic review and meta-analysis of health care utilization outcomes in alcohol screening and brief intervention trials. Med Care 2011;49:287–94.
[73] Shorter GW, Bray JW, Giles EL *et al*. The variability of outcomes used in efficacy and effectiveness trials of alcohol brief interventions: a systematic review. J Stud Alcohol Drugs 2019;80:286–98.
[74] Davidson L, Harding C, Spaniol LJ. Recovery from severe mental illnesses: research evidence and implications for practice, Vol. 1. Boston, MA: Center for Psychiatric Rehabilitation, Sargent College of Health and Rehabilitation Sciences, Boston University, 2005.
[75] Slade M, Hayward M. Recovery, psychosis and psychiatry: research is better than rhetoric. Acta Psychiatr Scand 2007;116:81–3.
[76] SAMSHA. Treatment Episode Dataset discharge data (TEDs). Available at: https://www.samhsa.gov/data/data-we-collect/teds-treatment-episode-data-set (accessed October 2019).
[77] Donovan DM, Bigelow GE, Brigham GS *et al*. Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use end-points in clinical trials. Addiction 2012;107:694–708.
[78] Wykes T. Great expectations for participatory research: what have we achieved in the last ten years? World Psychiatry 2014;13:24–7.
[79] Bjornestad J, Hegelstad WTV, Berg H *et al*. Social media and social functioning in psychosis: a systematic review. J Med Internet Res 2019; 21:e13957.

APPENDIX 1. DEVIATIONS FROM THE STUDY PROTOCOL

1. Change in research question: Our research question in the protocol registration (PROSPERO) was the following: *To review the existing literature (RCTs) on treatment efficacy, and to determine the treatment factors important for long-term drug reduction and functional recovery after substance abuse*. Our research question in the submitted review was the following: *To systematically review non-SU treatment outcome measures used in longitudinal randomised controlled trials over the last 10 years, and to assess the degree to which they reflect any of the above-mentioned perspectives of recovery. RCTs were chosen because this methodology is generally considered the most valuable for both evaluating treatment efficacy and developing treatment guidelines*. The reason for this change: Early in the review process (after PROSPERO registration) we did our first search, using our previous aim (broader criteria, including no time restriction etc.) as guidance. This search identified an insurmountable number of articles and needed to be narrowed. Also, we identified Tiffany *et al*. from 2012 (see reference list), explicitly addressing issues similar to the aim of our review. However, their findings were a result of an expert consensus meeting and not a systematic review of the literature. We knew that functioning and social functioning had previously been addressed more prominently in the earlier DUD literature (e.g. in the 1970s). At this point we discussed possible approaches that could provide meaningful and clinically relevant contextualisation for our review. Here, the idea of different recovery perspectives (personal, clinical, relational) emerged as a viable contextualisation, as modern recovery perspectives both address issues of functioning (e.g. community and social), incorporate various perspectives on outcome (e.g. service user and researcher perspectives) and are explicit that a long-term perspective is crucial particularly with regards to functional recovery. Since research on recovery has been in particular growth over the past 10 years, this became a central reason for the time limitation in our search—to test whether the DUD field had incorporated this shift in focus, from symptom relief (typically some measure of change in substance use), to more explicitly addressing function and social factors as important outcome measures.
2. Extended the study inclusion period to January 2019.
3. Some minor changes in search setup, including eligibility criteria, title search (see model search).
4. Removed the kappa coefficient to assess the level of agreement of the two independent reviewers for the selection of included and excluded measures. Due to the heterogeneity of the data material it was assessed that the best approach to reach valid outcome categories was a continuous collaboration (consensus meetings on a weekly basis).
5. Exclusion criteria: Studies that measured change in substance use only.

APPENDIX 2. MEDLINE SEARCH FOR REPLICATION

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>.

1. substance-related disorders/ or amphetamine-related disorders/ or cocaine-related disorders/ or heroin dependence/ or inhalant abuse/ or marijuana abuse/ or opioid-related disorders/ or morphine dependence/ or opium dependence/ or phencyclidine abuse/ or psychoses, substance-induced/ or substance abuse, intravenous/ or substance abuse, oral/ or narcotics/.
2. ((heroin or marijuana or marihuana or hashish or cannabis* or amphetamine* or opioid* or cocaine or opiate* or opium* or morphine* or ecstacy or metaamphetamine or polydrug* or polysubstance* or multidrug* or solvent* or inhalant* or narcotic*) adj2 (abus* or misus* or addict* or dependen* or 'use' or usage or disorder*)).hw,kf,ti,ab.
3. ((drug* or substance*) adj2 (abus* or misus* or addict* or dependen* or disorder*)).kf,ti,ab.
4. (sniff* or designer drug* or narcotism).hw,kf,ti,ab.

5. addiction.hw,kf,ti.
6. 1 or 2 or 3 or 4 or 5
7. therapeutics/ or drug therapy/ or rehabilitation/ or psychotherapy/.
8. (treatment* or intervention* or rehabilitation or inpatient* or outpatient* or hospitali?ed. patient* or residential or day hospital or partial hospital or continuing care or 'contin* of Care' or CBT or community reinforcement or motivational interviewing or motivational enhancement therapy or incentives or family therapy or couples therapy or methadone or suboxone or buprenorphine or therapeutic community or medication* or mentali?ation* or dialectic* or emotion* focused or 'action and commitment*' or psychodynamic* or psychoanaly* or behavio?r* modification* or behavio?r* therapy or 'drug adj2 therapy' or pharmacotherapy).hw,kf,ti.
9. (('12' or twelve) adj (step facilitation or step program*)).hw,kf,ti.
10. 7 or 8 or 9.
11. (recovery or autorecovery or remission* or autoremission* or abstinen* or abstain* or drug free or sobriety or (life adj2 satisfaction) or wellbeing or well being or self-quit* or self-change* or self-agen* or self-restrain* or change strateg* or life* change* or 'readiness to change' or 'stages of change' or 'quality of life').hw,kf,ti.
12. ((increas* adj2 function*) or (improv* adj2 function*)).hw,kf,ti,ab.
13. ((reduc* or modif* or decreas*) adj2 (abus* or misus* or addict* or dependen*)).hw,kf,ti,ab.
14. 'Quality of Life'/.
15. (vocation* or occupation* or job or jobs or work or employe* or employment or education* or educating or school).hw,kf,ti.
16. (social adj2 interact*).kf,ti.
17. life change events/.
18. (life style* adj2 change*).hw,kf,ti.
19. Interpersonal Relations/.
20. 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19.
21. randomized controlled trial.pt.
22. rct*.ti,ab.
23. (randomi?ed. and controlled and trial*).ti,ab.
24. 21 or 22 or 23 (585927).
25. exp. animals/ not humans.sh. (4474870).
26. 24 not 25 (573722).
27. 6 and 10 and 20 and 26 (1065).
28. limit 27 to english language (1052).
29. limit 28 to journal article (1050).
30. remove duplicates from 29 (1030).

Note: RCT filter based on Therapy, category specific/narrow here: https://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.Clinical_Queries_Filters

APPENDIX 3. DRUG USE SUB-CATEGORIES

- Days of drug use last month
- Substance use problem severity
- Substance use problem severity past month
- Monthly frequency of cannabis use
- Monthly frequency of alcohol use
- Monthly frequency of other drug use
- Substance problems past month
- Substance use problems past 90 days
- Number of days of substance use
- Number of days cocaine use
- Drug use during study period
- One-month abstinence
- Days abstinent during 90-days time period
- Illicit opiate use
- Days of heroin use
- Maximum days of consecutive heroin abstinence
- Drug cessation
- Dependence
- Change in substance use context
- Relapse
- Days to first relapse