Journal of Substance Abuse Treatment 144 (2023) 108928

Contents lists available at ScienceDirect

Journal of Substance Abuse Treatment

journal homepage: www.elsevier.com/locate/jsat

Systematic review of the efficacy, effectiveness, and cost-effectiveness of stepped-care interventions for the prevention and treatment of problematic substance use

Ashleigh K. Morse [a], Jayden Sercombe [a], Mina Askovic [a], Alana Fisher [a], Christina Marel [a], Mary-Lou Chatterton [b], Frances Kay-Lambkin [c], Emma Barrett [a], Matthew Sunderland [a], Logan Harvey [a], Natalie Peach [a], Maree Teesson [a], Katherine L. Mills [a,*]

[a] The Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, Level 6 Jane Foss Russell Building, Camperdown, NSW 2006, Australia
[b] Deakin Health Economics, Institute for Health Transformation, School of Health and Social Development, Faculty of Health, Deakin University, Geelong Waterfront Campus, Locked Bag 20001, Geelong, VIC 3220, Australia
[c] School of Medicine and Public Health, Faculty of Medicine and Health, The University of Newcastle Callaghan NSW 2308, Australia

ARTICLE INFO

Keywords:
Stepped-care
Adaptive-care
Matched-care
Substance use
Prevention
Treatment

ABSTRACT

Background: Stepped-care is a commonly recommended and implemented care model across health care domains, including substance use. Despite their presumed efficient allocation of treatment resources, a current and robust evidence synthesis is needed on the efficacy, effectiveness and cost-effectiveness of stepped-care for substance use.

Methods: This systematic review analyzed articles describing evaluations of stepped-care models that measured the use of acutely psychoactive substances (i.e., alcohol, cannabis, hallucinogens, inhalants, opioids, sedatives, hypnotics, anxiolytics, and stimulants) as a primary or secondary outcome, in participants over 18 years old. The analysis investigated model and participant characteristics associated with treatment outcomes.

Results: The study team conducted a search of five databases of literature (PsychINFO, MEDLINE, Embase, Cochrane Library and Scopus) published between January 1, 2010, and November 1, 2020. The search yielded 1051 unique articles, 19 of which were included in the analysis. The studies had considerable variability in sample sizes ($n = 18$–2310), time to follow-up (4.5 months to 3 years), and retention rates (35.1–100 %). Studies examined outcomes for either alcohol alone ($n = 9$), alcohol and other drug use ($n = 9$), or drug use alone ($n = 1$). Most studies ($n = 13$;) were rated as good quality. Three (15.8 %) were rated as fair and three (15.8 %) were rated as poor quality. The evidence regarding the efficacy, effectiveness and cost-effectiveness of stepped-care approaches is limited, but four of seven studies found that adaptive-care interventions delivered in the context of other systemic interventions produced greater benefit than control conditions in relation to at least one alcohol-related outcome. We have insufficient evidence to determine whether the modes or intensity of interventions included in the models, or decision rules used to step people up or down to differing levels of care, have an impact on outcome.

Conclusion: Heterogeneity between studies with regard to model and evaluation design limited the degree to which the analysis could draw robust conclusions. Sample recruitment and statistical power are particular challenges, and the field needs more innovative evaluation designs to assess the efficacy, effectiveness, and cost-effectiveness of stepped-care models.

1. Introduction

The global burden of disease caused by alcohol and other drug use is high, with 4.2 % and 1.3 % of all disability-adjusted life years (DALYs) attributable to alcohol and drug use, respectively (Graham et al., 2017; Sacks et al., 2015). Interventions designed to reduce this burden cross the spectrum of prevention, early intervention, and treatment; each of these has a range of different types and intensities of intervention

* Corresponding author.
E-mail address: katherine.mills@sydney.edu.au (K.L. Mills).

https://doi.org/10.1016/j.jsat.2022.108928
Received 15 November 2021; Received in revised form 15 September 2022; Accepted 30 October 2022
Available online 4 November 2022

(DiClemente et al., 2017; Mewton et al., 2018). Internationally, guidelines and policy documents routinely recommend the use of "stepped-care" approaches for selecting an appropriate level of care (LOC) for individuals in relation to substance use and a range of other health conditions (Australian Government Department of Health, 2016; Bair et al., 2015; Cross & Hickie, 2017; Kalarchian & Marcus, 2018).

Stepped-care models are designed to treat a range of symptom severity levels by offering a hierarchy of interventions, from least to most intensive, matched to an individual's needs based on the severity of their illness (Australian Government Department of Health, 2016; Bower & Gilbody, 2005). They may incorporate two or more intervention modalities of inherently different intensities, or the same modality at different intensity levels. These models may take an adaptive- or matched-care approach. In adaptive-care models, clients are first offered the least intensive (and typically least expensive) LOC that is appropriate based on severity of illness, and may then progressively 'step-up' or 'step-down' to more, or less, intensive LOCs contingent on their response to the previous tier (Kay-Lambkin et al., 2010). In matched-care, although a variety of steps are available, clients are assessed and allocated to the most appropriate LOC at intake, and remain at that LOC for the duration of treatment (Van Straten et al., 2006).

Systematic reviews and meta-analyses of stepped-care approaches for the prevention and treatment of mental health disorders such as depression and anxiety report mixed findings in relation to superiority of adaptive-care models relative to treatment as usual (TAU) (Ho et al., 2016), and a lack of evidence on matched-care and cost-effectiveness (Reeves et al., 2019; van Straten et al., 2015). For example, Ho et al. (2016) concluded that stepped-care interventions outperformed TAU in relation to preventing and treating anxiety but not depression, while van Straten et al. (2015) found that stepped-care had a moderate effect on depression relative to TAU. Similarly, studies evaluating the use of stepped-care for other chronic medical condition including chronic pain (Peterson et al., 2018) and comorbid mental and somatic disorder management (Maehder et al., 2018) report mixed or positive results from stepped-care relative to TAU or other comparators.

Only one systematic review has been conducted in relation to stepped-care for substance use since 2010. Jaehne et al. (2012) synthesized the findings of seven randomised controlled trials (RCTs) published up to 2010 that examined the efficacy and cost-effectiveness of adaptive-care approaches to psychosocial treatments for alcohol and tobacco use. This review found limited evidence to support the effectiveness of stepped care models with only one RCT demonstrating the efficacy and cost-effectiveness of a stepped care approach. However, due to the small number of studies conducted at that time, and the methodological limitations of those studies (e.g., small sample size, high variance in the intensity of base and step-up interventions), the authors were constrained in the questions they could examine and were unable to draw firm conclusions.

Since 2010, research examining stepped-care (both adaptive and matched) approaches to substance use has grown substantially, incorporating a broad range of interventions (e.g., brief interventions, motivational interviewing, telephone counselling and pharmacotherapy) for a variety of substances (e.g., alcohol, cocaine, cannabis and methamphetamine). This growth allows us the opportunity to examine key outstanding questions. In this systematic review, we aim to synthesize this literature to examine the:

i. Efficacy, effectiveness, and cost-effectiveness of stepped-care models for the prevention and treatment of problematic substance use; and
ii. Characteristics of stepped-care models (e.g., the types of interventions included in the model) associated with improvements in outcomes.

2. Method

The research team prospectively registered the protocol for this review with the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD42020218543).

2.1. Search strategy

Full details of the search strategy are available in the supplementary material. We conducted searches on five electronic databases of published literature: PsycINFO, MEDLINE, Embase, Cochrane Library and Scopus. A combination of free-text keywords and Medical Subject Headings (MeSH) were adapted to the conventions of each database. Search terms were specific to three concepts:

1. substance use (e.g., "substance use", "substance dependence");
2. care model (e.g., "stepped care", "matched care" "adaptive treatment"); and
3. study design (e.g., "randomized controlled trial", "cost-effectiveness").

Following full-text screening, the team manually searched the reference lists of included articles (snowballing), and those citing them (reverse snowballing), to identify other eligible articles.

2.2. Inclusion and exclusion criteria

Inclusion criteria were as follows: (i) written in English, (ii) published between 1 January 2010 and 1 November 2020, (iii) conducted with human participants, (iv) conducted with participants aged over 18, (v) reported on outcomes (primary or secondary) related to alcohol, cannabis, hallucinogens, inhalants, opioids, sedatives, hypnotics, anxiolytics, and stimulants, (vi) substance use outcomes not limited to caffeine and/or nicotine, (vii) intervention comprised a stepped model (i.e., composed of two or more treatment modalities and/or the same treatment modality at different intensity levels), and (viii) evaluation measured the efficacy, effectiveness, and/or cost-effectiveness of the stepped-care model. The time period examined was chosen to build on existing reviews of the literature that have focused predominately on research conducted prior to 2010 and provide an examination of the contemporary landscape of stepped-care research which has grown considerably, and incorporates a broader range of interventions, and substances. This review excludes articles that utilized cross-sectional, cohort or case studies/reports designs or did not describe the results of a study. We also applied post-hoc exclusion criteria, including the exclusion of studies that reported on secondary analysis of a primary study that was already included and systematic reviews that reported only on studies published before 2010.

2.3. Data collection

Study staff imported search results into Covidence for screening and data extraction. After removal of duplicates, three reviewers (AM, JS and MA; two reviewers per article) independently screened titles and abstracts and conducted full-text review on all articles retrieved. The group resolved conflicts via consensus. Four reviewers (AM, JS, MA and MLC) undertook data extraction. Two reviewers independently extracted data for all papers, with the exception of economic evaluation data (extracted by MLC). A third reviewer resolved conflicts (AM, JS or MA). Extracted data included study characteristics (year, design, setting, number of participants, nature of study control group/s); participant characteristics (age, sex, target population, participant inclusion and exclusion criteria); care model characteristics (LOCs, components of care/interventions, decision rules and timepoints); and primary and secondary outcomes measured. Components of care/interventions were operationalized according to study authors' labelling of those interventions.

2.4. Evidence synthesis and quality assessment

Although a meta-analysis would be optimal, due to the considerable heterogeneity in study designs, a narrative synthesis approach was used to answer the research questions. Quality was assessed using Korakakis et al's. (2018) modified version of the Downs and Black Quality Index (1998). The index is comprised of 27 items (e.g., "were study subjects randomised to intervention group). Two investigators independently rated each study; a third reviewer resolved conflicts (AM, JS and MA). This review categorized quality ratings as: poor (scores ≤14), fair (15–19), good (20–25) and excellent (26–28).

We completed quality assessment for the economic evaluation using the Drummond checklist (Drummond et al., 2015) containing 10 criteria (e.g., "were the cost and consequences valued credibly?") with 33 sub-questions (e.g., "were market values employed for changes involving resources gained or depleted) answered by 'yes' (scored 1), 'no' (scored 0), and 'can't tell' (scored 0.5) if there was insufficient information (Gonzalez-Perez, 2002). The team categorized studies as poor (scores ≤5) fair (6–8), or good (≥9) quality.

3. Results

3.1. Studies

3.1.1. Study selection

The search yielded 1051 unique articles after we removed duplicates (Fig. 1). Of these, the review excluded 934 in title and abstract screening, leaving 117 articles for full-text review. Twenty-three articles met the a priori inclusion criteria (see Fig. 1). Three articles (Borsari et al., 2016; Yurasek et al., 2015; Yurasek et al., 2017) were excluded post hoc as they presented findings of secondary analyses of pooled data from two primary studies (Borsari et al., 2012; Borsari et al., 2014). This review excluded post hoc a systematic review that reported findings of stepped-care studies for alcohol and nicotine use up to 2010 (Jaehne et al., 2012). Nineteen publications remained for inclusion in this review.

3.1.2. Study characteristics

Table 1 provides a summary of the characteristics of the studies included and Table 2 provides a more detailed information regarding the characteristics of all included studies. Baseline sample sizes ranged from 18 (Kay-Lambkin et al., 2010) to 2310 (Merkx et al., 2013). Time to

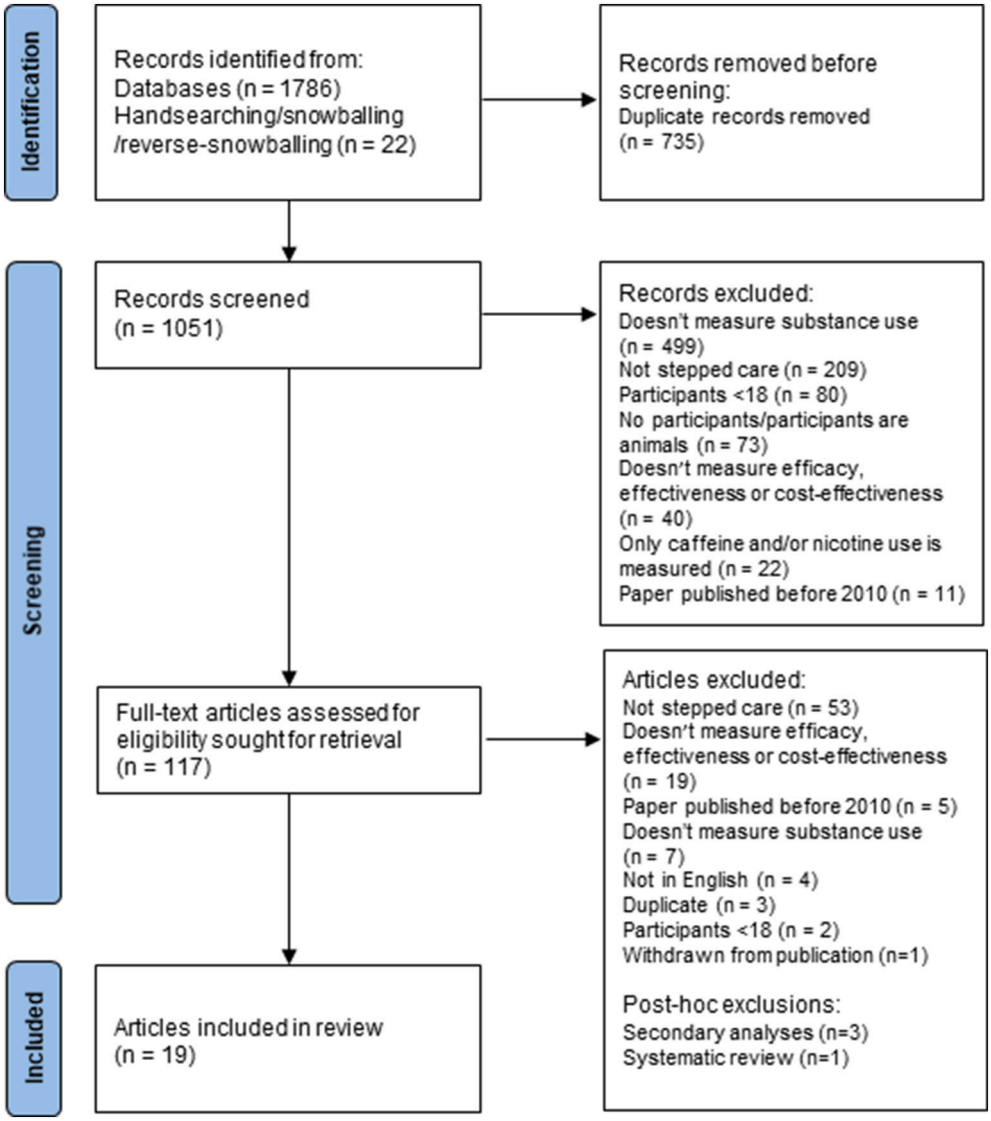

Fig. 1. PRISMA diagram.

Table 1
Characteristics of included studies.

Variable	Summary descriptives
Age	Mean age ranged from 18.63 to 63 years
Gender	% female ranged from 1 % - 93 %
Settings	Outpatient services (4 studies)
	Veteran Health Association HIV clinic (3 studies)
	Inpatient and outpatient services (2 studies)
	University (2 studies)
	Primary care (2 studies)
	Drug court (1 study)
	Inpatient withdrawal treatment (1 study)
	Prison (1 study)
	Hospital (1 study)
	Workplace (1 study)
	Not specified (1 study)
Country	United States (11 studies)
	Netherlands (2 studies)
	Australia (1 study)
	UK (1 study)
	Germany (1 study)
	Canada (1 study)
	Norway (1 study)
	Spain (1 study)
Substances of focus	Alcohol use (9 studies)
	Alcohol and other drug use (6 studies)
	Alcohol and cocaine use (2 studies)
	Methamphetamine use (1 study)
	Alcohol, other drugs and misuse of prescription drugs (1 study)
Intervention modalities	Motivational interviewing/enhancement therapy (15 studies)
	Brief intervention (10 studies)
	Pharmacotherapy (7 studies)
	Cognitive behavioral therapy (5 studies)
	Referral to specialist substance use services (5 studies)
	Other counselling (3 studies)
Outcomes	Self-reported alcohol use and related harms (ADU, ASI, ASSIST, AUDIT-C, DPI, MATE, TLFB, YAACQ; 18 studies)
	Self-reported drug use and related harms (ASI, ASSIST, OTI, TLFB; 4 studies)
	Alcohol biomarker (PEth concentration; 3 studies)
	Drug test (saliva toxicology, urine toxicology; 7 studies)
	Healthcare utilisation and costs (2 studies)
	General physical and mental health and quality of life (SF-12, EQ-5D; 1 study;)

ADU = Alcohol and Drug Use Measure.
ASI = Addiction Severity Index.
ASSIST = Alcohol, Smoking, and Substance Involvement Screening Test.
AUDIT-C = Alcohol Use Disorders Identification Test – Consumption.
DPI = Drinking Problems Index.
MATE = Measurements in the Addictions for Triage and Evaluations.
TLFB = Timeline Followback.
YAACQ = The Young Adult Alcohol Consequences Questionnaire.
OTI = Opiate Treatment Index.
Peth = phosphatidylethanol.
SF-12 = 12-item Short Form Survey.
EQ-5D = EuroQol – 5 D.

follow-up ranged from 4.5-months (Marlowe et al., 2012) to 3-years (Gómez-Recasens et al., 2018) post-baseline. Most studies followed up with participants up to a maximum of 6- ($n = 3$) or 12-months ($n = 7$) post-baseline. Participant retention rates ranged from 35.1 % (Laporte et al., 2018) to 100 % (Marlowe et al., 2012) at final follow-up. With regard to type of substance, nine studies examined outcomes related to alcohol use only, one examined other drug use outcomes only, and nine examined alcohol and other drug use outcomes.

3.2. Efficacy and effectiveness of stepped-care models

Table 3 summarizes the findings of all included studies. The authors identified three studies as studies of efficacy (Borsari et al., 2012; Borsari et al., 2014; Buchholz et al., 2020) and 16 studies as studies of effectiveness. However, as all were conducted in real world settings, we present their findings together as one group of studies. The 19 studies evaluated adaptive- ($n = 11$), matched- (n = 7) or adaptive/matched combination (n = 1) care models. In adaptive-care models, clients are first offered the least intensive LOC that is appropriate based on severity of illness, and may then progressively "step-up" or "step-down" to more, or less, intensive LOCs contingent on their response to the previous tier (Kay-Lambkin et al., 2010). In matched-care, although a variety of steps are available, clients are assessed and allocated to the most appropriate LOC at intake, and remain at that LOC for the duration of treatment (Van Straten et al., 2006).

3.2.1. Adaptive-care models

With the exception of Kay-Lambkin et al. (2010), all studies of adaptive-care models utilized RCT designs, typically with TAU or minimal intervention control conditions. One study included a matched-care comparison group (Marlowe et al., 2012). Six studies (Edelman et al., 2019a, 2019b; Edelman et al., 2020; McKay et al., 2010; McKay et al., 2015; Satre et al., 2019) evaluated the effectiveness of an adaptive-care model in the context of an existing health care program. Participants in the remaining four studies (Borsari et al., 2012; Borsari et al., 2014; Kay-Lambkin et al., 2010; Watson et al., 2013) did not systematically receive additional care to the intervention being tested, however, they were not precluded from accessing additional services.

The adaptive-care models varied with regard to the number of levels of care (LOC; range: 2–4) and the intensity of the interventions offered, as well as the decision rules that determined when participants were stepped up to a more intensive LOC. Models included either two (Borsari et al., 2012; Borsari et al., 2014; Satre et al., 2019), three (Edelman et al., 2019a, 2019b; Edelman et al., 2020; McKay et al., 2010; Watson et al., 2013), or four (Kay-Lambkin et al., 2010) LOCs. Most models used a low-intensity base intervention that was offered to all participants in the adaptive-care group. These typically involved a single-session brief intervention (BI) ranging from 15- to 60-minute delivered in person (Borsari et al., 2012; Edelman et al., 2019b; Edelman et al., 2020; Kay-Lambkin et al., 2010; Satre et al., 2019; Watson et al., 2013) or via telephone (Borsari et al., 2014), with some offering one or two brief telephone follow-ups (Edelman et al., 2019b; Edelman et al., 2020; McKay et al., 2010; Satre et al., 2019) and referral to online resources (Edelman et al., 2019b; Edelman et al., 2020). The most intensive base intervention involved provision of eight sessions of addiction physician management (Edelman et al., 2019a).

The frequency with which participants were assessed to determine whether they would step up to the next LOC ranged from weekly (Marlowe et al., 2012; McKay et al., 2010; McKay et al., 2015) to once 6-months post-baseline (Satre et al., 2019). Decision rules used to determine whether participants were stepped up to a more intensive LOC were based primarily on some measure of problematic/risky/heavy substance use (Borsari et al., 2012; Borsari et al., 2014; Edelman et al., 2019a; Edelman et al., 2020; Satre et al., 2019; Watson et al., 2013). Other measures included any alcohol use (Edelman et al., 2019b; Marlowe et al., 2012), a multi-component substance use risk assessment (McKay et al., 2010), and treatment engagement (Marlowe et al., 2012; McKay et al., 2015). One study (Kay-Lambkin et al., 2010) used a combination of substance use assessment, clinician judgement, and participant preference to determine whether a participant's care was stepped up.

3.2.1.1. Adaptive models in the absence of other systematic care. Four evaluations of adaptive-care models evaluated the model in the absence of additional care. Borsari et al. (2012, 2014) recruited students at a US university who were mandated to treatment following an alcohol-related incident (e.g., public intoxication, vandalism, hospital admission for intoxication). The study offered all participants a BI at LOC1 (consisting of a 15-minute advice session and education booklet), and the study later randomised those still reporting risky drinking at 6-weeks

Table 2
Methodological characteristics of studies evaluating stepped-care interventions (adaptive and matched) for substance use.

Study	Study design	Setting	Participant characteristics		Care model description	Decision rule(s)	Control group(s)
			Study population	Baseline sample (n)			
Adaptive-care studies							
Borsari et al., 2012	RCT	University	US undergraduate students who violated campus alcohol policy	598	LOC1) Brief advice session (15-min) and education booklet LOC2) Brief motivational intervention (60–90 min) and personalised feedback on alcohol use	Participants who reported continued risky alcohol use (≥4 heavy drinking episodes and/ or a score of ≥5 on the YAACQ) at 6 weeks progressed to LOC2	Received LOC1 and then assessment only
Borsari et al., 2014	RCT	University	As in Borsari et al. (2012)	57	As in Borsari et al. (2012), except the brief motivational intervention (LOC2) was delivered via telephone	As in Borsari et al. (2012)	As in Borsari et al. (2012)
Edelman et al., 2019a	RCT	Veterans Health Association HIV clinic	US Veterans with HIV and an alcohol use disorder	128	LOC1) Addiction physician management including pharmacotherapy LOC2) Addiction physician management including pharmacotherapy plus 4 sessions of motivational enhancement therapy LOC3) Referral to a higher level of specialty services (e.g., intensive outpatient or residential treatment)	Patients reporting heavy drinking on at least one day in the past 2 weeks at week 4 advanced from LOC1 to LOC2. At week 12, patients reporting heavy drinking at least once in the past 2 weeks on LOC2 were escalated to LOC3	TAU included regular treatment at the HIV clinic and referral to substance use treatment services
Edelman et al., 2019b	RCT	Veterans Health Association HIV clinic	US Veterans with HIV and liver disease	95	LOC1) Brief negotiated interview (15–20 min) with a telephone booster 2 weeks later and referral to web-based resources LOC2) Motivational enhancement therapy with content tailored to people with HIV. Additional web-based resources for self-help LOC3) Addiction physician management including pharmacotherapy	Patients reporting any alcohol use in the past 2 weeks at week 4 were stepped up from LOC1 to LOC2. At week 12, those who were advanced to LOC2 and who reported any alcohol use in the past 2 weeks were advanced to LOC3	TAU differed at each HIV clinic. Patients received a handout with healthy drinking advice. The Veterans Administration also prompted clinics to conduct annual AUDIT-C screening and brief interventions or referral to addiction treatment where necessary.
Edelman et al., 2020	RCT	Veterans Health Association HIV clinic	US Veterans with HIV with at-risk levels of alcohol use	93	As in Edelman et al. (2019b)	Patients reporting at-risk drinking in the past 2 weeks at week 4 were stepped up to LOC2. At week 12, patients still engaging in high-risk drinking in the past 2 weeks on LOC2 were escalated to LOC3	As in Edelman et al. (2019b)
Kay-Lambkin et al., 2010	Non-randomised	Not specified	Australian methamphetamine users with comorbid depression	18	LOC1) Brief integrated intervention including feedback, self-help and case formulation LOC2) 4 sessions of therapy (CBT combined with motivational enhancement) LOC3) An additional 4 sessions of therapy LOC4) An additional 4 sessions of therapy	Decisions to step treatment up or down at weeks 5, 10, 15 and 20 were based on the following factors: 1. Client preference 2. Clinical judgement 3. Assessment results	Received LOC1, then instead of receiving stepped-care, the control group were given an integrated CBT and MI treatment focusing on methamphetamine use and depression
Marlowe et al., 2012	RCT	Drug Court	US admissions to Drug Court who met criteria for substance use disorder	125	LOC1) Baseline matching to bi-weekly court hearings or as-needed court hearings LOC2) Increased frequency of court hearings and/or intensive	Those with too many unexcused absences from counselling sessions or failures to provide a valid urine specimen had their schedule of court hearings increased.	Baseline matching (LOC1) was conducted in the same way as the stepped-care group (matched to bi-weekly hearings if high-risk or as-needed court hearings if low-risk). The control group received

(*continued on next page*)

Table 2 (*continued*)

Study	*Study design*	*Setting*	*Participant characteristics*		*Care model description*	*Decision rule(s)*	
			Study population	*Baseline sample (n)*			*Control group(s)*
					clinical case-management program	Participants who provided ≥2 drug-positive urine specimens were referred to the intensive clinical case-management program	the same monthly assessment as stepped-care, but were not moved to LOC2 based on the results
McKay et al., 2010	RCT	Intensive outpatient treatment	US outpatients who met criteria for alcohol dependence	252	Stepped-care (telephone monitoring with counselling): LOC1) One face-to-face counselling session, followed by a phone call risk assessment (5–10 min) LOC2) More frequent telephone calls provided over several weeks LOC3) Face-to-face motivational-interviewing evaluation sessions followed by CBT sessions	If a participant was categorized as high-risk via phone risk assessment at LOC1, the stepped-care component was triggered. If their level of risk didn't drop at LOC2, LOC3 was entered	TAU: involved 4 months of intensive outpatient programs (mainly group-based treatment) and any step-down standard outpatient care that the patients received Telephone monitoring: A second comparison condition provided participants with LOC1 only
McKay et al., 2015	RCT (sequential randomization)	Intensive outpatient treatment	US outpatients who met criteria for alcohol or cocaine dependence	500	LOC1) Intensive outpatient program (IOP) for alcohol or cocaine dependence (9 h of group therapy per week) LOC2) Randomised to: a) Motivational interviewing phone calls (engagement; MI-IOP) aimed to facilitate re-engagement in the IOP. b) Motivational interviewing phone calls (patient choice; MI-PC) where patients were given a choice of enrolling in IOP, or individual CBT, telephone-based stepped-care, or medication management. LOC3) Motivational interviewing phone calls (patient choice as described in LOC2b) Note: this study included telephone-based stepped-care nested within the overall stepped-care model. As very few patients opted for this treatment ($n = 6$), it was not described in this table)	Participants who failed to engage in LOC1 (IOP) at week 2 or who dropped out received LOC2. Those not engaged at weeks 2 and week 8 received LOC3	There was a control condition (no further outreach) for those not engaged at weeks 2 and week 8 (participants who reached LOC3). Rather than receiving the motivational interviewing phone calls (patient choice), control participants received no further outreach
Satre et al., 2019	RCT	Primary care	US patients with HIV in primary care reporting unhealthy alcohol use	624	Motivational interviewing stepped-care (MI stepped-care): LOC1) One 45-min in-person motivational-interviewing session followed by two 20-min telephone sessions LOC2) Two additional 20-min motivational interviewing telephone sessions Emailed feedback stepped-care (EF stepped-care): LOC1) Personalised email feedback regarding the	Non-responders (participants who reported unhealthy alcohol use in the 30 days prior to the 6-month interview) were offered LOC2 of each intervention	TAU included routine HIV primary care treatment and annual screening for harmful alcohol use. Patients screening positive for harmful use were advised by general practitioners to decrease or stop drinking. Those with alcohol-related problems were referred to addiction treatment services. TAU did not include screening or intervention for use of drugs other than alcohol

(*continued on next page*)

Table 2 (*continued*)

Study	*Study design*	*Setting*	*Participant characteristics*		*Care model description*	*Decision rule(s)*	
			Study population	*Baseline sample (n)*			*Control group(s)*
					risks of alcohol/drug use LOC2) An additional message regarding substance use risks		
Watson et al., 2013	RCT	Primary care	Adults ≥55 in the UK scoring ≥8 on the AUDIT	529	LOC1) Session of motivational interviewing-based counselling with a nurse (20-min) LOC2) Motivational enhancement therapy delivered over three weekly sessions (40-min) in primary care LOC3) Referral to a specialist alcohol treatment services	Telephone assessment of the alcohol use in the past 4 weeks (AUDIT-C). The participant was stepped up if they were still using alcohol at hazardous levels	Received the baseline screening and a nurse-led structured advice session (5 mins). The session was informed from the participant's screening results and included personalised feedback and advice on reducing their alcohol use. Participants also received a booklet containing information about risks associated with high alcohol use and local referral information
Matched-care studies							
Buchholz et al., 2020	RCT	Inpatient withdrawal treatment	German patients with a diagnosis of alcohol dependence in inpatient withdrawal treatment	250	LOC 1) Brief outpatient advice LOC 2) Outpatient treatment LOC 3) Day/residential treatment LOC 4) Inpatient or outpatient long-term care	The MATE was administered during an assessment interview to calculate participant assignment to one of four levels of care	Patients in the control group followed TAU for inpatient withdrawal (medical detoxification and psychosocial treatment)
Laporte et al., 2018	Single group uncontrolled	Outpatient services	Canadian patients with borderline personality disorder	681	LOC 1) Short-term clinic, patients receive individual therapy and group therapy (4 months) LOC 2) Extended care clinic, patients receive group therapy, individual therapy, and pharmacological management (24 months)	Initial assessments were completed by an experienced rater who directed participants to LOC 1 or LOC 2. 12 % of patients allocated to LOC 1 requested more intensive treatment and were moved to LOC 2.	No control group
Merkx et al., 2011	Naturalistic prospective	Outpatient services	Dutch outpatients with alcohol use disorder	890	LOC 1) Brief outpatient treatment, 4–6 individual or group sessions based on principles of motivational interviewing (3 months) LOC 2) Standard outpatient treatment, 10–12 individual or group sessions based on principles of motivational interviewing and risk analysis (over 6 months)	Patients were matched to each LOC based on a manual-guided assessment considering treatment history, social stability, psychiatric impairment, and addiction severity	No control group
Merkx et al., 2013	Naturalistic prospective	Inpatient and outpatient services	Dutch outpatients with alcohol use disorder	2310	LOC 1) Brief outpatient treatment comprised of 4–6 sessions of manual-guided CBT (over 3 months) LOC 2) Standard outpatient treatment comprised 10–12 sessions of manual-guided CBT and motivational interviewing (over 6 months) LOC 3) Intensive outpatient/inpatient treatment comprised of 18 group sessions (over 6–12 weeks). Treatment included a varied combination of	As in Merkx et al. (2011). However, if the intake interviewer did not agree with the recommended allocation, they could overrule it.	No control group

(*continued on next page*)

Table 2 (*continued*)

Study	*Study design*	*Setting*	*Participant characteristics*		*Care model description*	*Decision rule(s)*	
			Study population	*Baseline sample (n)*			*Control group(s)*
					interventions including CBT, vocational therapy and pharmacotherapy		
Prendergast et al., 2017	RCT	Prison	Inmates at a US jail scheduled for release in the following month	732	LOC 1) Personalised feedback and education LOC 2) Brief intervention (15–20 min) in jail comprising of personalised feedback and education, using principles of motivational interviewing LOC 3) Referral to community treatment following release and offer to participate in brief treatment (eight sessions)	Risk level was assessed by the ASSIST. Low-risk participants received LOC 1. Moderate-risk participants received LOC 1 and 2. High-risk participants received LOC 1, 2 and 3	Control group participants received personalised feedback on their risk assessment results. They also received literature on reducing substance use and HIV risk, as well as a list of local treatment services
Stallvik et al., 2015	Naturalistic prospective	Inpatient and outpatient services	Norwegian inpatients or outpatients diagnosed with substance dependence or abuse	261	LOC 1) Outpatient treatment: group and individual therapy, and pharmacotherapy if needed (10–12 sessions) LOC 2) Residential/ inpatient treatment: individual and group sessions, exercise, legal assistance, housing assistance and vocational training (3–18 months)	Patients were matched to LOC as per TAU assignment. The ASAM Criteria interview was also conducted to calculate a ASAM recommendation	No control group
Woodruff et al., 2014	RCT	Hospital	US trauma/ emergency department patients who have used illegal drugs	700	LOC 1) Brief intervention delivered on-site including feedback and motivational enhancement LOC 2) 6 individual sessions (motivational interviewing and CBT) with a substance use counsellor over the phone LOC 3) Provision of a list of local agencies that could provide further assessment and support	Participants were classified into risk categories using the DAST-10. At-risk participants received LOC 1. High-risk participants received LOC 1 and 2. Severe-risk participants received LOC 1, 2 and 3	Control participants received a driving safety intervention (15–20 min), scaled to their risk level (2 levels). To match the treatment condition, the control was based on principles of motivational interviewing
Adaptive- / matched-care combined studies							
Gómez-Recasens et al., 2018	Single group	Workplace	Industrial workers in Spain (mainly blue collar)	1103	LOC1) Company-wide substance use education (e.g., brochures, training). In addition, medical examination interviews and drug tests were conducted to monitor employee substance use LOC2) Secondary prevention including brief intervention (10–15 min), personalised advice using principles of motivational interviewing or referral to specialist services	Participants were stepped up if they were consuming risky amounts of alcohol, or consuming drugs. Risky use was identified by semi-structured interview or positive drug test	No control group

RCT = Randomised controlled trial.
TAU = treatment as usual.
LOC = Level of Care.
ADU = Alcohol and Drug Use Measure.
YAACQ = The Young Adult Alcohol Consequences Questionnaire.
DRNF = Drinking Norms Rating Form.
BCEAS = Brief Comprehensive Effects of Alcohol Scale).
TLFB = Timeline Followback.
MATE = Measurements in the Addictions for Triage and Evaluations.
ASI = Addiction Severity Index.
OTI = Opiate Treatment Index.

ASSIST = Alcohol, Smoking, and Substance Involvement Screening Test.
American Society of Addiction Medicine = ASAM.
AUDIT-C = Alcohol Use Disorders Identification Test – Consumption.
DPI = Drinking Problems Index.
SF-12 = 12-item Short Form Survey.
DAST-10 = Drug Abuse Screening Test.
phosphatidylethanol = PEth (alcohol biomarker).

to either LOC2 (consisting of a 60–90-minute motivational intervention and personalised feedback delivered in-person or over the phone), or to no additional care (assessment-only control). This sequential randomization design allowed them to compare students who were offered additional treatment for risky alcohol use with those who were not. They found no significant change within- or between-groups in the number of heavy drinking episodes, drinks consumed, or estimated blood alcohol content (BAC) in peak or typical drinking episodes at 9-month follow-up, for either the in-person (2012) or phone-based (2014) intervention. The authors did, however, find in both studies that the adaptive-care group was significantly less likely to report alcohol-related problems (IRR = 0.84 & 0.40, respectively) or another alcohol-related incident (3.1 % vs 9.1 % & 11.1 % v 33.3.%, respectively), compared to assessment-only controls at 9-month follow-up.

Watson et al. (2013) opportunistically screened older (aged >55 years) primary care patients for risky alcohol use and randomised those who screened positive to adaptive-care or a minimal control intervention consisting of feedback and advice. They offered a 20-minute counselling session based on motivational interviewing (MI) in LOC1, three motivational enhancement therapy (MET) sessions delivered by a substance use counsellor in LOC2, and referral to specialist substance use services in LOC3. The authors did not analyze within-group changes and did not find any significant between-group differences on any outcomes at 6- or 12-month follow-up.

Kay-Lambkin et al. (2010) compared an integrated adaptive-care model for depression and methamphetamine use that assessed participants' need for additional care at 5-week intervals, to a full-care intervention that allocated participants to all four steps regardless of assessment outcomes. Although the authors were unable to perform inferential tests due to their small sample size, they reported that adaptive-care appeared to be as effective as the full-care intervention in reducing methamphetamine use.

3.2.1.2. Summary. Two of the four studies examined within-group differences, and both found no differences between baseline and the 9-month follow-up on a range of alcohol-related outcomes (Borsari et al., 2012; Borsari et al., 2014). Overall, an absence of between-group differences was evident in three of the four studies on outcomes at 6-, 9- and 12-month follow-up (Borsari et al., 2012; Borsari et al., 2014; Watson et al., 2013). The exception being that the adaptive-care group in the studies by Borsari et al. (2012, 2014) were significantly less likely to report alcohol-related problems or another alcohol-related incident compared to assessment-only controls at 9-month follow-up. Only one study focused on outcomes related to other drug use. Although promising, inferential testing was not conducted to determine either within- or between-group differences.

3.2.1.3. Adaptive models in the context of other systematic interventions. Four of the seven studies that evaluated adaptive-care in the context of existing care recruited patients from HIV treatment clinics and aimed to reduce alcohol use (Edelman et al., 2019a, 2019b; Edelman et al., 2020; Satre et al., 2019). Edelman et al. conducted three studies with US veterans with HIV who met criteria for i) an alcohol use disorder (AUD; assessed using the Mini-Structured Clinician interview for DSM – Mini-SCID) and were not already seeking treatment for alcohol use (2019a); ii) moderate (any alcohol use in the past 30 days) but not at-risk use (defined as self-reported consumption of 14 or more drinks per week or 4 or more per occasion for men aged 65 years or younger, or 7 or more drinks per week or 3 or more drinks per occasion for women or men older than 65 years old), plus liver disease (2019b); or iii) at-risk use (as defined in Edelman et al., 2019b) but not meeting criteria for an AUD (assessed using the Mini-SCID) (Edelman et al., 2020).

Edelman et al. varied the intensity of their care models with the severity of participants' alcohol use. Those designed for moderate (2019b) and risky (2020) alcohol use were less intensive, offering a 15-minute interview with telephone booster interview (LOC1) and MET (LOC2) before offering Addiction Physician Management (APM) in LOC3. In contrast, in their model for those with an AUD (2019a) APM was introduced at LOC1. LOC2 involved the addition of 4 sessions of MET, and LOC3 involved referral to specialist services. In all three studies, adaptive-care participants continued to receive treatment as usual (TAU) for HIV and the study compared them to TAU-only controls. TAU included annual screening for problematic alcohol use, with feedback and advice for patients who screened positive. In addition, in Edelman et al.'s (2019a) study for participants with AUD, those assigned to TAU also received facilitated referral to alcohol treatment services.

Satre et al. (2019) used a similar study design, recruiting participants with HIV who reported unhealthy alcohol use, defined as any days consuming ≥3 drinks in a day for women and ≥ 4 drinks in a day for men in the prior 12 months. The authors compared outcomes for participants randomised to one of two adaptive-care + TAU groups to TAU-only controls. TAU included annual screening with feedback and advice to reduce alcohol use for unhealthy drinkers. In contrast to Edelman et al., Satre et al. evaluated two low-intensity models: MI and email feedback.

All four studies (Edelman et al., 2019a, 2019b, 2020; Satre et al., 2019) reported that, regardless of group allocation, all participants reduced at least one alcohol use outcome at 12-month follow-up. However, only Satre et al. (2019) formally analyzed within-group changes over time, and found that all three groups significantly reduced alcohol use and alcohol-related problems at 12-month follow-up (see Table 3), but did not find significant changes in other drug use outcomes. They did not find significant differences between the groups in alcohol or other drug use at 6- or 12-month follow-up.

Of these four studies, only Edelman et al.'s study of participants who met criteria for an AUD (Edelman et al., 2019a) found that randomization to adaptive-care was associated with significantly lower alcohol use compared to TAU. Edelman et al. (2019a) found that relative to TAU the adaptive-care group reported fewer drinking days (23 % v 40 %) and heavy drinking days (31 % v 64 %), and drinks per drinking day (41. V 6.3) at 52-week follow-up. They did not, however, find a significant difference between adaptive-care and TAU in number of drinks per week at 52-week follow-up, and did not find any significant between-group differences in either of their other two studies (2019b, 2020), nor did Satre et al. (2019).

McKay et al. (2010, 2015) conducted two studies of adaptive-care interventions for patients with alcohol or cocaine dependence already engaged in an intensive outpatient treatment program consisting of nine hours of group therapy each week over 3–4 months. All participants had completed two (2015) or three (2010) weeks of outpatient treatment and continued to receive TAU throughout the trial. McKay et al. (2010) compared adaptive-care in the form of telephone monitoring + MI-based counselling to telephone monitoring only and TAU-alone groups. The telephone monitoring intervention involved substance use assessment and feedback, but no counselling. The study found significant group by time interactions for mean percentage of days with alcohol use and heavy alcohol use. Those in the telephone monitoring +

Table 3
Efficacy, effectiveness and cost effectiveness of stepped-care interventions for substance use.

Study	*Number and proportion of sample followed up (where available)*	*Substance use and cost-related outcomes (measure)*	*Analysis and main findings*	*Quality assessment*
Adaptive-care studies				
Borsari et al., 2012	LOC1 baseline: 598 6-week follow-up: 582 Risky drinkers randomised to LOC2 or control: 405 Follow-ups after LOC2: 3-month: 471 6-month: 468 9-month: 473	• Days of heavy episodic drinking (ADU) • Peak BAC (ADU) • Alcohol-related problems (YAACQ) • Recidivism (receiving another alcohol infraction)	Between-group differences: • Days of heavy episodic drinking and mean peak BAC at follow-up did not differ significantly between those randomised to receive LOC2 of the stepped-care intervention and those who were randomised to receive assessment-only over time. • Significantly greater reductions in YAACQ scores were found among those randomised to receive step 2 compared to those randomised to receive assessment-only (IRR = 0.84, 95 % CI: 0.74, 0.94, d = 0.23). • Recidivism rates over the course of the 9-month follow-up were significantly lower among those randomised to receive LOC2 (3.1 %) compared to the assessment-only control group (9.1 %).	Good (20/28)
Borsari et al., 2014	Baseline: 57 3-month: 46 (80.7 %) 6-month follow-up: 52 (91.2 %) 9 months follow-up: 48 (84.2 %)	• Drinks per week (ADU) Days of heavy episodic drinking (ADU) • Frequency of drinking (past 30 days, ADU) • Typical BAC (ADU) • Alcohol-related problems (YAACQ) • Recidivism (receiving another alcohol infraction)	Between-group differences: • Drinks per week, heavy episodic drinking, frequency of drinking and BAC did not differ significantly between those in the stepped-care group compared to those in the control group over the course of the 9-month follow-up. • Those randomised to receive LOC2 had lower scores on alcohol-related problems (IRR = 0.40, 95 % CI: 0.39, 0.45) and recidivism rates (likelihood ratio $\chi 2$ (1) = 4.06) over the course of the 9-month follow-up compared to assessment-only control participants.	Fair (15/28)
Edelman et al., 2019a	Baseline: 128 4-week follow-up (step-up assessment): 113 (88 %) 12-week follow-up (step-up assessment): 107 (84 %) 24-week follow-up: 98 (77 %) 52-week follow-up: 74 (58 %)	Past 30 day (TLFB): • Drinks per week • Proportion of participants with no heavy drinking days • Proportion of participants reporting alcohol abstinence • Drinks per drinking day Percentage of days abstinent Alcohol biomarker: • PEth concentration	Within-group differences: • Both groups decreased in the number of drinks per week over time[a]: - Stepped-care: baseline: M = 31.3 (SD = 23.5), 52 weeks: M = 7.8 (SD = 14.3). - TAU: baseline: M = 32.8 (SD = 27.3), 52 weeks: M = 15.9 (SD = 18.9). Between-group differences: • There were no significant differences between groups in number of drinks per week at 24-week or 52-week follow-up. • At 52 weeks (but not at 24 weeks), compared to the control group, the stepped-care group had: - a significantly larger proportion of participants with no heavy drinking days (69.4 % v 35.9 %; adjusted OR = 4.90, 95 % CI: 1.51, 15.84). - a significantly higher percentage of abstinent days (77 % v 60 %; adjusted treatment effect = 15 %, 95 % CI: 1 %, 30 %]. - significantly fewer drinks per drinking day (4.1 v 6.3; adjusted treatment effect = −2.22, 95 % CI: −3.79, −0·66). • No other comparisons were significant.	Good (22/28)
Edelman et al., 2019b	Baseline: 95 4-week follow-up (step-up assessment): 87 (92 %) 12-week follow-up (step-up assessment): 82 (86 %) 24-week follow-up: 81 (85 %) 52-week follow-up: 76 (80 %)	Past 30 day (TLFB): • Proportion of participants reporting alcohol abstinence • Drinks per week • Proportion of participants with no heavy drinking days • Drinks per drinking day Percentage of days abstinent Alcohol biomarker: • PEth concentration	Within-group differences: • Rates of abstinence for both groups increased over time[a]: - Stepped-care: baseline: 0 %, 24 weeks: 38.1 %, 52 weeks: 32.5 %. - TAU: baseline: 0 %, 24 weeks: 23.1 %, 52 weeks: 36.1 %. Between-group differences: • No between group comparisons were significant at 24 or 52 weeks.	Good (22/28)
Edelman et al., 2020	Baseline: 93 4-week follow-up (step-up assessment): 85 (91 %) 12-week follow-up (step-up assessment): 76 (82 %) 24-week follow-up: 79 (85 %) 52-week follow-up: 60 (65 %)	Past 30 day (TLFB): • Drinks per week • Proportion of participants with no heavy drinking days • Drinks per drinking day • Proportion of days abstinent	Within-group differences: • Both groups showed evidence of decreased number of drinks per week over time[a]: - Stepped-care: baseline: M = 17.0 (SD = 10.9), 24 weeks: M = 8.8 (SD = 8.6), 52 weeks: M = 10.6 (SD = 12.3).	Good (22/28)

(*continued on next page*)

Table 3 (*continued*)

Study	*Number and proportion of sample followed up (where available)*	*Substance use and cost-related outcomes (measure)*	*Analysis and main findings*	*Quality assessment*
		Alcohol biomarker: • PEth concentration	- Control: baseline: M = 22.8 (SD = 28.8), 24 weeks: M = 10.6 (SD = 14.8), 52 weeks: M = 11.0 (SD = 19.6). Between-group differences: • No between group comparisons were significant at 24 or 52 weeks.	
Kay-Lambkin et al., 2010	Baseline: 18 5-week follow-up: 11 (61.1 %) 10-week follow-up: 8 (44.4 %) 15-week follow-up: 8 (44.4 %) 20-week follow-up: 8 (44.4 %)	Methamphetamine use (OTI)	Within-group differences[a]: • Both groups showed evidence of decreased OTI scores over time[a]: - Stepped-care: baseline: M = 1.14, 20 weeks: M = 0.64 - Control: baseline: M = 1.08: 20 weeks M = 0.25. Between-group differences[a]: • The stepped-care group halved their methamphetamine use while the control group demonstrated an 80 % reduction.	Poor (7/28)
Marlowe et al., 2012	Baseline: 130 Follow-ups were conducted weekly for 18 weeks, final study sample: 125 (96.2 %)	Non-responsiveness (≥2 positive drug screens) Urine drug test results	Within-group differences: • The likelihood of being abstinent increased over successive weeks for participants in both groups (OR = 1.07, 95 % CI: 1.04, 1.10). Between-group differences: • Participants in the stepped-care group had similar rates of non-responsiveness as those in the control group (21 % v 25 %). • Participants in the stepped-care condition were drug-negative for an average of 68 % of weeks (SD = 0.39), which was significantly higher than an average of 49 % of weeks (SD = 0.43) in control group participants (d = 0.46). • The odds of being drug abstinent were 2.34 times higher in the stepped-care group then the control group across the 18 weeks of follow-ups (95 % CI: 1.22, 4.64).	Good (22/28)
McKay et al., 2010	Baseline: 252 3 months: 224 (88.9 %) 6 months: 217 (86.1 %) 9 months: 204 (81 %) 12 months: 198 (78.57 %) 15 months: 195 (77.4 %) 18 months: 191 (75.79 %)	Past 30 day (TLFB): • Percent days of alcohol use • Percent days of heavy alcohol use • Any alcohol use • Any heavy alcohol use Cocaine use (urine toxicology)	Between-group differences: • There was a significant group by time interaction for percent days of alcohol use: - Compared to participants randomised to TAU, those randomised to the stepped-care group had lower percent days of alcohol use during months 10–12 (d = 0.42), months 13–15 (d = 0.65), and months 16–18 (d = 0.50). - The stepped-care group had fewer days of alcohol use compared to the telephone monitoring group during months 4–6 (d = 0.43). - The telephone monitoring group had fewer days of alcohol use than the TAU group during months 10–12 (d = 0.41) and months 13–15 (d = 0.39). • There was a significant group by time interaction for percent days of heavy alcohol use: - Those in the stepped-care group had a lower percent days of heavy alcohol use compared to those in the TAU group in months 13–15 (d = 0.59), months 16–18 (d = 0.46), and fewer days of heavy alcohol use than the telephone monitoring group during months 4–6 (d = 0.43). • There was a significant group main effect for any alcohol use: - Participants randomised to receive TAU had significantly higher rates of alcohol use across the trial period compared to those randomised to receive stepped-care (OR = 1.88, 95 % CI: 1.13, 3.14). - The telephone monitoring group did not have significantly different rates of any alcohol use than the stepped-care group or the TAU group. There was no significant main effect for any heavy alcohol use or cocaine-positive samples.	Good (22/28)
McKay et al., 2015	Baseline: 500 Follow-up: 189 (37.8 %)	Past 30 day alcohol use (TLFB): • Percent days of alcohol use • Percent days of heavy alcohol use • Any alcohol use	Between-group differences: • For patients not engaged at 2 weeks, those in MI-IOP were less likely than those in MI-PC to have months with any drinking days (OR = 0.40, 95 % CI: 0.23,	Fair (15/28)

(*continued on next page*)

Table 3 (*continued*)

Study	*Number and proportion of sample followed up (where available)*	*Substance use and cost-related outcomes (measure)*	*Analysis and main findings*	*Quality assessment*
		• Any heavy alcohol use Past 30 day cocaine use (TLFB): • Percent days cocaine use • Any cocaine use Cocaine use (urine toxicology)	0.68) and without heavy drinking days (OR = 0.33, 95 % CI: 0.19, 0.58). • Frequencies of drinking days (B = −1.08, 95 % CI: −1.87, −0.29, d = −0.39) and heavy drinking days (B = −1.09, 95 % CI: −1.80, −0.39, d = −0.40) were also lower in MI-IOP than in MI-PC. • There were no differences in those who dropped out at weeks 3–8. • There were no advantages of MI-PC over MI-IOP or no further outreach for cocaine use outcomes.	
Satre et al., 2019	Baseline: 614 6-month follow-up: 582 (94.79 %) 12-month follow-up: 583 (94.95 %)	Past 30 day: • Any unhealthy alcohol use (≥ 4/≥ 5 drinks per day for women/men) • Any substance use/prescription drug misuse (not including marijuana; past 30 days) Alcohol-related problems (ASI)	Within-group differences: • Rates of any unhealthy alcohol use significantly reduced over time in all three groups: - MI: baseline: 50.3 %, 6 months: 24.7 %, 12 months. - EF: baseline: 44.6 %, 6 months: 23.1 %, 12 months. - Control: baseline: 49.8 %, 6 months: 22.4 %, 12 months: 27.1 % • ASI scores reduced significantly over time in all three groups: - MI: baseline: M = 0.21 (SD = 0.15), 6 months: M = 0.17 (SD = 0.15), 12 months: M = 0.15 (SD = 0.12). - EF: baseline: M = 0.19 (SD = 0.13), 6 months: M = 0.16 (SD = 0.12), 12 months: M = 0.14 (SD = 0.10). - Control: baseline: M = 0.21 (SD = 0.15), M = 0.17 (SD = 0.12), 12 months: M = 0.15 (SD = 0.11). Between-group differences: • There were no significant differences between groups in proportion of participants with any unhealthy alcohol use or alcohol-related problems at either follow-up. • The MI stepped-care group had lower rates of substance use/prescription drug misuse (not including marijuana) at 6 months (12.4 %) compared to the EF stepped-care group (22.1 %) and the control group (23 %), but not at 12 months.	Good (23/28)
Watson et al., 2013	Baseline: 529 6-month follow-up: 474 (89.6 %) 12-month follow-up: 462 (87.3 %)	Alcohol use outcomes: • Average drinks per day (derived from AUDIT-C) • AUDIT-C score • Alcohol-related problems (DPI) Economic outcomes: • Perception of health (SF-12 and EQ-5D) • Health and social care and justice services utilisation	Within-group differences: • The mean average drinks per day reduced over time in both groups from baseline to 6 months, and to 12 months. - Stepped-care: baseline: M = 3.38 (SD = 2.24), 6 months: M = 2.45 (SD = 1.85), 12 months: M = 2.56 (SD = 2.09). - Control: baseline: M = 3.41 (SD = 2.19), 6 months: M = 2.81 (SD = 2.03), 12 months: M = 2.49 (SD = 1.93) Between-group differences: • There were no significant differences between the stepped-care and control groups on any alcohol use or economic outcomes at 6- or 12-month follow-up. • The probability of stepped-care being cost-effective was 81 %–86 % at 6 months, and 93.5 %–93.8 % at 12 months.	Study rating: Good (23/28) Economic evaluation rating: Good (10/10)
Matched-care studies				
Buchholz et al., 2020	Baseline: 250 6-month follow-up: 165 (66.0 %)	Past 30 day: • Heavy alcohol use (past 30 days; ≥6 drinks for men or ≥ 4 for women) • Frequency of heavy drinking • Percent days of heavy drinking Cost outcomes: • Indirect health care costs • Direct health care costs • Total health care costs	Between-group differences: • There were no significant differences between the matched-care and control groups with regard to heavy drinking, indirect or total costs. Moderator analysis (level of concordance between the treatment recommended vs received): Irrespective of whether randomised to the matched-care or control groups: • Undermatched patients were more likely to report more heavy drinking days at 6 months (IRR = 2.09, 95 % CI: 1.26, 3.48). • Overmatched patients and matched patients did not have significantly different outcomes for heavy drinking days.	Good (22/28)

(*continued on next page*)

Table 3 (*continued*)

Study	*Number and proportion of sample followed up (where available)*	*Substance use and cost-related outcomes (measure)*	*Analysis and main findings*	*Quality assessment*
			• Overmatched patients had significantly greater total healthcare costs compared to matched patients (IRR = 1.79, 95 % CI: 1.05, 3.07) and direct healthcare costs (IRR = 1.79, 95 % CI: 1.07, 2.99). • Undermatched patients had significantly lower indirect healthcare costs than matched patients (IRR = 0.21, 95 % CI: 0.07, 0.63). • Participants in the matched care group did not significantly differ to those in the control group on alcohol use or cost outcomes.	
Laporte et al., 2018	Baseline (short-term): 97 Baseline (extended care): 584 12-week follow-up (short-term clinic): 79 (84 %) 6-, 12-, 18- or 24-month follow up (extended-care; depending on length of treatment): 239 (40.9 %)	Alcohol use to intoxication (ASI) Drug use to intoxication (ASI)	Within-group differences: • Significant reductions in alcohol use to intoxication (40 % at baseline, 28 % at follow-up) and drug use to intoxication care (58 % at baseline, 42 % at follow-up) were observed in participants allocated to receive extended care. • The percentage of participants who used alcohol and/or drugs to intoxication did not significantly diminish in the short-term clinic group.	Poor (13/28)
Merkx et al., 2011	Baseline (treatment allocation): 890 9 to 12 months post-intake follow-up: 427 (48 %)	Past 30 day: • Current drinking status (abstinent, controlled drinker, or currently drinking excessively) • Change in the number of excessive (≥5 drinks per drinking day) and nonexcessive drinking days from intake to follow-up	Between-group differences: • Undertreated (attended <75 % of recommended treatment), matched (attended 75–110 % of recommended treatment) and overtreated (attended >110 % of recommended treatment) patients had similar likelihoods of being abstinent, controlled drinkers or drinking excessively. • There were no significant differences in the change in the number of excessive or non-excessive drinking days between undertreated, matched and overtreated patients.	Fair (17/28)
Merkx et al., 2013	Baseline (treatment allocation): 2310 9 to 12 months post-intake follow-up: 1253 (54.2 %)	Past 30 day: • Current drinking status (abstinent, controlled drinker, or currently drinking excessively) • Change in the number of abstinent days from intake to follow-up • Change in the number of heavy drinking days (≥5 drinks per drinking day) from intake to follow-up	Between-group differences: • The percentage of patients that were currently abstinent or controlled drinkers was significantly higher in overtreated patients (attended >110 % of recommended treatment; 55.3 %), compared to matched (attended 75–110 % of recommended treatment; 43.9 %) or undertreated patients (attended <75 % of recommended treatment; 38.3 %). • There was no significant difference in the percentage of patients that were currently abstinent or controlled drinkers among patients who were matched or undertreated. • Change in the number of abstinent days and heavy drinking days was greatest among overtreated patients (10.5 & -12.5) compared to matched (7.6 & -10.5) and undertreated patients (5.5 & -8.7).	Fair (18/28)
Prendergast et al., 2017	Baseline: 732 Records data: 732 (100 %) 12-month follow-up (self-report): 521 (71.2 %)	Alcohol and other drug use (ASSIST): • Alcohol use (past 3 months) • Drug use (past 3 months) • Change in risk level from baseline to follow-up (four categories: *never used, no change in risk, improved risk, or worse risk*) • Length of time following release from jail to using drugs or alcohol	Between-group differences: • Past 3-month alcohol and drug use did not differ significantly between those in the stepped-care group compared to the control group at follow-up. • The stepped-care group did not significantly differ from the control group in the percentage of participants in each risk category at follow-up. • The stepped-care group did not significantly differ from the control group in the length of time following release to using alcohol or other drugs.	Good (20/28)
Stallvik et al., 2015	Baseline: 261 3-month follow-up: 152 (58.2 %)	Past 30 day: • Frequency of drug use (alcohol, opiates, sedatives, stimulants, cannabis, two or more per substances per day) • ASI subscale composite drug score • ASO subscale composite alcohol score	Within-group differences: • The matched group (TAU assignment to the same LOC as the ASAM recommendation) evidenced a significant reduction in days of alcohol use (baseline: M = 13.84, SD = 11.06; follow-up: M = 7.73, SD = 8.72) and cannabis use (baseline: M = 18.55, SD = 10.62; follow-up: M = 6.86, SD = 8.70), but no other drug classes. • The overmatched group (TAU assignment to a higher LOC than the ASAM recommendation) evidenced a significant decrease in days of alcohol use (baseline: M = 10.76, SD = 10.24; follow-up: M = 4.29, SD = 5.85). • No significant changes in alcohol or other drug use were observed in the undermatched group (TAU	Good (21/28)

(*continued on next page*)

Table 3 (*continued*)

Study	*Number and proportion of sample followed up (where available)*	*Substance use and cost-related outcomes (measure)*	*Analysis and main findings*	*Quality assessment*
			assignment to a lower LOC than the ASAM recommendation). • Significant decreases were observed in ASI composite alcohol scores in the matched and overmatched group, but not the undermatched participants. - Matched: baseline: M = 0.09, SD = 0.14; follow-up: M = 0.04, SD = 0.12 - Overmatched: baseline: M = 0.06, SD = 0.12; follow-up: M = 0.02, SD = 0.08. • Significant declines were observed in ASI composite drug scores in all groups: - Matched: baseline: *M* = 0.24, *SD* = 0.12; follow-up: *M* = 0.12, *SD* = 0.07. - Undermatched: baseline: *M* = 0.21, *SD* = 0.12; follow-up: *M* = 0.15, *SD* = 0.11. - Overmatched: baseline: *M* = 0.16, *SD* = 0.10; follow-up: *M* = 0.09, *SD* = 0.06. Between-group differences: • Controlling for baseline frequency of drug use, the matched group had significantly lower stimulant use at follow-up compared to the over-matched group (d = −1.06) • No other between-group comparisons on substance use were significant • There were no significant between group differences in changes from baseline to follow-up in ASI composite drug or alcohol scores.	
Woodruff et al., 2014	Baseline: 700 6-month follow-up: 292 (41.7 %)	Past 30 day (ASI-Lite): • alcohol and other drug abstinence at follow-up • alcohol use composite score • drug use composite score Biologically validated past 30-day drug abstinence at follow-up (hair sample)	Between-group differences: • There were no significant differences between groups, and no significant changes over time in ASI-Lite alcohol use scores or past 30-day drug abstinence (self-reported or biologically validated). • Complete case analysis revealed a significant group by time interaction for ASI drug use composite scores with smaller increases in the matched group (baseline: M = 0.059, follow-up: M = 0.068) compared to the control group (baseline: M = 0.055, follow-up: M = 0.095).	Good (22/28)
Adaptive- / matched-care combined				
Gómez-Recasens et al., 2018	Baseline: 1103 Year 1 follow-up: 990 (89.8 %) Year 2 follow-up: 700 (63.5 %) Year 3 follow-up: 625 (56.7 %)	• Risky alcohol use (assessed by semi-structured interview) • Risky drug use (positive drug test) • Total risky use (combination of risky drug and alcohol use).	Within-group differences: • A significant decrease in the prevalence of risky alcohol use was observed from baseline (14.6 %) to year 1 (10.6 %). Further reductions were observed through to year 2 (9.3 %) which were maintained at year 3 (10.7 %). • There was a significant difference in the prevalence of risky drug use from year 1 to year 2 (6.7 % to 6.9 %) but not at any other points. • The prevalence of total risky use significantly reduced from baseline (19.0 %) to year 1 (15.3 %) and was maintained to year 3 (year 2: 14.4 %; year 3: 14.9 %).	Poor (12/28)

RCT = Randomised controlled trial.
ADU = Alcohol and Drug Use Measure.
YAACQ = The Young Adult Alcohol Consequences Questionnaire.
DRNF = Drinking Norms Rating Form.
BCEAS = Brief Comprehensive Effects of Alcohol Scale).
TLFB = Timeline Followback.
MATE = Measurements in the Addictions for Triage and Evaluations.
ASI = Addiction Severity Index.
OTI = Opiate Treatment Index.
ASSIST = Alcohol, Smoking, and Substance Involvement Screening Test.
ASAM = American Society of Addiction Medicine.
AUDIT-C = Alcohol Use Disorders Identification Test – Consumption.
DPI = Drinking Problems Index.
SF-12 = 12-item Short Form Survey.
DAST-10 = Drug Abuse Screening Test.
PEth = phosphatidylethanol.
Quality-Adjusted Life Year = QALY.
[a] Results of inferential testing were not provided.

counselling group reported significantly fewer days of alcohol use and heavy alcohol use at 15- (d's = 0.65 & 0.59, respectively) and 18-month follow-ups (d's = 0.50 & 0.46), compared to TAU. However, the monitoring + counselling group only reported fewer alcohol use and heavy use days compared to the monitoring-only group at 6-month follow-up (d's = 0.43 for both).

McKay et al. (2015) evaluated a complex adaptive-care model to re-engage patients in intensive outpatient treatment who were disengaged 2- and/or 8-weeks after intake. Participants identified as disengaged and randomised to the control group at 2-weeks were stepped up to LOC2 and offered MI phone calls designed to re-engage them in outpatient treatment (two sessions of up to 1-h each). If they were disengaged at 8-weeks, they were stepped up to LOC3, which consisted of MI phone calls designed to help participants choose between re-enrolling in outpatient treatment, individual cognitive behavioral therapy (CBT), telephone-based stepped-care, or pharmacotherapy. Participants identified as disengaged and randomised to the intervention group at 2-weeks were offered LOC3 from the control condition (MI phone calls designed to help participants choose between treatment options). McKay et al. (2015) found that MI phone calls to re-engage participants in outpatient treatment significantly reduced both any drinking (OR = 0.40) and heavy drinking (OR = 0.33), but only for participants who were disengaged at 2-weeks and re-engaged thereafter. The study found no significant differences between-groups in any outcome among participants who were disengaged in weeks 3 to 12.

One study directly compared the effectiveness of adaptive- and matched-care for substance use with mandated drug court attendees (Marlowe et al., 2012). Marlowe et al. (2012) compared a matched-care intervention to an adaptive model that stepped participants up to more frequent monitoring if they failed to attend treatment sessions, or continued to return drug-positive urine screens. Overall, there was a 7 % increase in the rate of abstinence from baseline to 4-month follow-up. Participants randomised to adaptive-care screened drug-negative in significantly more weeks of the trial (68 %) compared to those randomised to matched-care (68 % v 49 %, d = 0.46).

3.2.1.4. Summary. There has been minimal inferential testing with regard to within-group change over time; however, in terms of between-group differences, four of seven studies found that adaptive-care interventions delivered in the context of other systemic interventions produced greater benefit than TAU or other control conditions in relation to at least one alcohol-related outcome over 12-months follow-up (Edelman et al., 2019a; McKay et al., 2010; McKay et al., 2015; Satre et al., 2019). The two studies that have examined outcomes related to other drug use reported mixed findings, with one study reporting a more favorable association between adaptive-care and drug use outcomes at 4-month follow-up (Marlowe et al., 2012) and the other finding no between-group differences at 6- or 12-month follow-up (Satre et al., 2019).

3.2.2. Matched-care models

Seven studies evaluated matched-care models, utilizing either single-group prospective (n = 4) or RCT (n = 3) designs. Three of the seven studies compared a matched-care model to a control intervention (Buchholz et al., 2020; Prendergast et al., 2017; Woodruff et al., 2014); one compared substance use outcomes in participants matched to a high LOC compared to a (comparatively) low LOC (Laporte et al., 2018); four assessed participant outcomes relative to the concordance between the LOC considered appropriate for their substance use severity (based on algorithms) and the LOC participants received (Buchholz et al., 2020; Merkx et al., 2011; Merkx et al., 2013; Stallvik et al., 2015).

In an uncontrolled trial of Canadian patients with borderline personality disorder, Laporte et al. (2018) compared the outcomes of patients who had received up to 4 months of individual and group therapy (LOC1) to those who received up to 24-months of individual and group therapy, as well as pharmacological management (LOC2; extended care). Only participants allocated to extended care significantly reduced their alcohol (40 % to 28 %) and other drug use (58 % to 42 %) at follow-up (which ranged from 12 weeks to 24 months, depending on the length of treatment).

Prendergast et al. (2017) examined the effectiveness of a matched-care model among prison inmates due for release consisting of three LOC: LOC1 consisted of personalised feedback and education; LOC2, personalised feedback and education and a 15–20-minute BI; and LOC3, referral to community treatment following release and an offer to participate in eight sessions of treatment. The study matched inmates based on level of risk as measured by the ASSIST. The control group received LOC1 only. Prendergast et al. (2017) did not examine changes in substance use over time, and the study found no significant between-group differences in substance use or the proportion of participants in each risk-change category (improved, no change, or worsened risk) at 12-month follow-up.

Woodruff et al. (2014) compared a matched-care model to a control driving safety intervention, among trauma/emergency department patients who had used illegal drugs. LOC1 consisted of a brief intervention delivered on-site including feedback and motivational enhancement; LOC2, 6 individual sessions (MI and CBT) with a substance use counsellor over the phone; and LOC3, provision of a list of local agencies that could provide further assessment and support. In Woodruff et al.'s (2014) intention-to-treat analysis, no significant main effects of or interaction between time or group occurred. Complete case analysis, however, found that addiction severity index (ASI) scores significantly *increased* from baseline to 6-month follow-up (M = 0.06 to 0.07), but that this increase was smaller in the matched-care group (M = 0.6 to 0.1).

Buchholz et al. (2020) evaluated a patient-treatment matching algorithm (the Measurements in the Addictions for Triage and Evaluations instrument; MATE) based on Dutch guidelines for aftercare following inpatient withdrawal treatment for alcohol use against standard aftercare procedures (control group). The study classified participants as "matched" if they were allocated to the LOC recommended by guidelines; under-matched if they were allocated to a lower LOC than recommended; and 'over-matched' if allocated to a higher LOC than recommended. Their model consisted of four LOCs: outpatient advice from a drug counsellor (LOC1); outpatient addiction specialist treatment (LOC2); 8+ weeks in day-residential treatment (LOC3); long-term inpatient or outpatient care (LOC4). Buchholz et al.'s (2020) found no significant differences between matched care and control groups, but reported that regardless of group allocation, the level of concordance between the LOC recommended by the MATE instrument and the actual treatment provided, mediated outcomes (Buchholz et al., 2020). Compared to matched patients, under-matched patients reported significantly more heavy drinking days (+3 days) at 6-month follow-up (IRR 2.09), and over-matched patients incurred significantly higher treatment costs (IRR = 1.79) but did not report significantly fewer drinking days.

Three uncontrolled prospective studies specifically examined outcomes in relation to concordance between the LOC recommended by treatment guidelines and the LOC allocated. Stallvik et al. (2015) evaluated the validity of patient-treatment matching guidelines, using American Society for Addiction Medicine criteria for treating alcohol and other drug use disorders. Their model had two LOCs, each delivered in a range of settings: LOC1 was outpatient treatment consisting of organized, regular substance use and/or mental health treatment delivered by clinicians; LOC2 was inpatient treatment consisting of residential services staffed 24-h by trained substance use and mental health clinicians. Stallvik et al. (2015) found that, at 3-month follow-up, matched patients reported significant reductions in days of alcohol (M = 13.84 to 7.73) and cannabis (M = 18.55 to 6.86) use, and a trend towards a significant reduction in stimulant use (p = 0.05). Over-matched patients reported a significant reduction in days of alcohol use at follow-

up (M = 10.76 to 4.29), and under-matched patients did not report any significant changes in substance use. Compared to under- and over-matched patients, matched patients reported significantly lower stimulant use at follow-up. See Table 3 for their findings across the groups in ASI subscales.

Merkx et al. (2011, 2013) evaluated the validity of patient-treatment matching guidelines (using an algorithm to derive a recommended LOC based on a patient's treatment history, addiction severity, psychiatric impairment, and social stability) in predicting outpatient treatment outcomes for people with an AUD. In their 2011 study, Merkx et al. matched participants to one of two LOCs: LOC1 consisted of a brief outpatient treatment comprised of 4–6 individual or group MI-based sessions; LOC2 comprised standard outpatient treatment consisting of 10–12 group or individual sessions based in MET, behavioral techniques and coping skills. Their 2013 study matched participants to one of three LOCs: LOC1 and LOC2 from their 2011 study, and LOC3 which offered intensive inpatient and outpatient options, including detoxification and 12–18 group sessions.

The 2011 study did not find any significant differences between matched, over- and under-matched participants, but in their 2013 study, Merkx et al. found that a significantly greater proportion of over-matched participants (55.3 %) were no longer drinking at harmful levels at follow-up, compared to matched (43.9 %) and under-matched (38.3 %) participants. Over-matched patients also reported a significantly larger reduction in their heavy drinking days (−12.5 days) than matched (−10.5) and under-matched (−8.7) patients, and a significantly greater increase in their abstinent days (+10.5) than matched (+7.6) and under-matched (+5.5).

3.2.3. Summary

Very few studies have examined within-group change over time; however, those that have report improvements among those allocated to higher levels of care (Laporte et al., 2018) or those who are matched or overmatched (Stallvik et al., 2015). No change over time seemed to occur for patients allocated to lower levels of care or those who were under-matched. Examinations of between-group differences demonstrate mixed findings with some reporting no between-group differences (Merkx et al., 2011; Prendergast et al., 2017; Woodruff et al., 2014), others reporting positive findings in relation to being over-matched relative to being matched/under-matched (Merkx et al., 2013) or matched relative to over-matched (Stallvik et al., 2015).

3.2.4. Combined adaptive-/matched-care models

One study of industrial workers in Spain evaluated a combined adaptive/matched care model in a single-group pre/post study. Gómez-Recasens et al. (2018) implemented a health promotion and substance use monitoring program at multiple worksites at the same company. Employees who were identified as using alcohol at risky levels, or using other drugs, through monitoring were stepped up into a matched-care model comprised of three LOCs. LOC1 consisted of company-wide education and drug testing; LOC2, a 10–15 min BI, personalised advice and referral to specialist services; and LOC3 referred participants to specialist substance use treatment. Gómez-Recasens et al. (2018) found that the proportion of employees reporting risky alcohol use significantly reduced from baseline (14.6 %) to 1-year follow-up (10.6 %), and from 1-year to 2-year follow-up (9.3 %). This reduction was maintained at 3-year follow-up (10.7 %). They did not find a significant change in other drug use from baseline to 1-year follow-up, and reported a modest but significant increase from 1-year to 2-year follow-up.

3.3. Cost-effectiveness of stepped-care models

One study examined the cost-effectiveness of an adaptive-care approach. In their examination of an adaptive-care model among adults aged >55 years, Watson et al. (2013) found that adaptive-care was not significantly more expensive or more effective than a minimal control intervention at 6-month follow-up. At 12-month follow-up, adaptive-care participants had incurred fewer treatment costs and gained marginally more QALYs than control participants, although these differences were not significant. The probability the adaptive-care would fall below the National Institute for Health and Care Excellence's willingness to pay threshold (£20,000-20,000/QALY) was 81.3–86.4 % at 6-months and 93.5–93.8 % by 12-months. No cost-effectiveness studies have been conducted in relation to matched-care.

3.4. Characteristics of stepped-care models associated with improvements in outcomes

3.4.1. Intervention modality

Motivational interviewing/enhancement therapy (MI/MET) was by far the most common intervention modality ($n = 15$), followed by brief interventions (BIs; $n = 10$), pharmacotherapy ($n = 7$), CBT ($n = 5$), referral to specialist substance use services ($n = 5$), and other counselling ($n = 3$). Note that these intervention modalities were not exclusive, and some studies offered them in combination (e.g. an MI-based BI). Due to methodological limitations within studies, and differences between studies, direct comparisons between interventions were not possible.

MI/MET tended to be delivered as a BI over 1–3 sessions of 15–45 min in duration in the first LOC of the model (Borsari et al., 2012; Borsari et al., 2014; Prendergast et al., 2017; Satre et al., 2019; Watson et al., 2013; Woodruff et al., 2014) but also featured in later steps as a component of more intensive counselling (3+ sessions) as either a stand-alone intervention (Edelman et al., 2019a, 2019b; Edelman et al., 2020; Watson et al., 2013) or within the context of CBT or other counselling (Kay-Lambkin et al., 2010; McKay et al., 2010; McKay et al., 2015; Merkx et al., 2011; Merkx et al., 2013; Woodruff et al., 2014). Of the 15 studies that included an MI/MET component, eight found an association between allocation to receive the stepped-care model and improvements in at least one substance use outcome (including overall use, drinks per drinking day, heavy drinking days, or alcohol-related problems) relative to baseline (Gómez-Recasens et al., 2018; Laporte et al., 2018; Satre et al., 2019) or a control condition (Borsari et al., 2012, 2014; McKay et al., 2010, 2015; Edelman et al., 2019a).

The 10 studies that used a BI in their model found little evidence to suggest that BIs are effective as part of a stepped-care model for substance use. Five studies (Borsari et al., 2012, 2014; Buchholz et al., 2020; Gómez-Recasens et al., 2018; Woodruff et al., 2014) analyzed within-group changes from baseline to follow-up, one of which found that stepped-care participants significantly reduced their substance use, but did not compare them to a control group. Borsari et al. (2012, 2014) found that most (78 %) of their participants failed to respond to a BI and continued to report risky drinking 6-weeks post-BI. The remaining five studies did not analyze within-group changes from baseline to follow-up, and did not find or report between-group differences (Kay-Lambkin et al., 2010; Watson et al., 2013; Prendergast et al., 2017; Edelman et al., 2019b, 2020).

Three of the five studies that examined a stepped-care approach that included CBT used it in combination with MI (Kay-Lambkin et al., 2010; McKay et al., 2010; Woodruff et al., 2014), making it difficult to disentangle the effects of CBT from that of MI. Of the two that did not include an MI component, the findings of one suggest that a greater number of CBT sessions was associated with an increased likelihood of achieving abstinence (Merkx et al., 2013). All three studies of individual and/or group counselling that was not based in MI or CBT reported similar findings (Laporte et al., 2018; Marlowe et al., 2012; Stallvik et al., 2015).

Referral to specialist substance use services was only used as the most intensive LOC in the stepped-care models reviewed (Edelman et al., 2019a; Gómez-Recasens et al., 2018; Prendergast et al., 2017; Watson et al., 2013; Woodruff et al., 2014). As such, the number of participants who were offered, and/or took up, a referral was typically small, making it impossible to examine the effect of this aspect of the stepped-care

model relative to other components.

Seven studies offered pharmacotherapies for alcohol use, but few reported on uptake, and those that did reported mixed findings with respect to the efficacy of the stepped-care intervention relative to TAU on uptake and alcohol-related outcomes (Edelman et al., 2019a).

3.4.2. Intervention intensity

No clear relationship exists between differences in LOC intensity and effectiveness. For example, Watson et al. (2013) and Edelman et al. (2019b, 2020) had some of the largest differences in intensity in their models: all three used a BI at LOC1, 3–4 sessions of MET at LOC2, and specialist substance use treatment at LOC3. However, none of these studies found significant differences in substance use between stepped-care and control at follow-up. Three studies explicitly investigated the effect of treatment engagement (i.e., actual receipt of more treatment) on outcomes, two of which found no association between amount of treatment attended and substance use at follow-up (Edelman et al., 2020; Merkx et al., 2011). One study (Edelman et al., 2019a) found in post-hoc analyses that stepped-care participants who attended at least 30 % of treatment sessions consumed significantly fewer drinks per week than those randomised to TAU at 52-week follow-up.

3.4.3. Decision rules

All adaptive-care models in this review had decision rules for stepping participants up to a more intensive LOC; however, only one model (Kay-Lambkin et al., 2010) had a rule for stepping participants down if they responded to treatment. Most adaptive models (n = 8) used a measure of problematic substance use to classify participants as non-responders and step them up to a more intensive LOC. All matched care models used a measure of substance use severity to match participants to the appropriate LOC. Two models (Marlowe et al., 2012; McKay et al., 2015) also used treatment engagement as part of their decision rules, and four included participant preference in their decision rules (Buchholz et al., 2020; Kay-Lambkin et al., 2010; Laporte et al., 2018; McKay et al., 2015). Five studies considered psychiatric comorbidity and social functioning in their decision rules, four of which were matched-care studies (Buchholz et al., 2020; Kay-Lambkin et al., 2010; Laporte et al., 2018; Merkx et al., 2011; Merkx et al., 2013). No studies systematically assessed the efficacy or effectiveness of different decision rules. No evidence was found for a relationship between decision rules and model effectiveness.

3.4.4. Summary

We have insufficient evidence to determine whether intervention modality, intervention intensity or decision rules used to determine movement between LOCs impact upon outcome.

3.5. Study quality

See Table 3 for all study quality ratings. Most studies (13/19) were rated “good” quality according to Korakakis et al.'s (2018) modified version of the Downs and Black (1998), three were “fair” quality, and three were “poor” quality. The most common methodological limitations in studies rated poor quality were failure to blind participants to their treatment condition, variable follow-up periods between and within study groups, and either no power calculation was reported or the study was underpowered. The economic evaluation by Watson et al. (2013) was rated good quality, meeting all 10 criteria of the Drummond Checklist (Drummond et al., 2015).

4. Discussion

4.1. Efficacy, effectiveness and cost-effectiveness of stepped-care models for the prevention and treatment of problematic substance use

Building on the findings of Jaehne et al. (2012), this review provides a synthesis of the international evidence regarding the efficacy, effectiveness and cost-effectiveness of stepped-care interventions for the prevention and treatment of problematic substance use. Over the last decade, stepped approaches to substance use treatment have received increased attention across populations and intervention settings, from opportunistic identification of substance use in primary care (Watson et al., 2013) to participants already engaged in intensive outpatient treatment for substance use (McKay et al., 2010; McKay et al., 2015).

Notably, most studies focused on participant samples identified as already having some level of problematic alcohol or other drug use at baseline. Only three studies implemented stepped-care universally, among prison inmates (Prendergast et al., 2017), patients diagnosed with borderline personality disorder (Laporte et al., 2018), and employees (Gómez-Recasens et al., 2018), and could be considered to have any element of prevention. Only one study conducted an economic evaluation (Watson et al., 2013), a limitation found in reviews of stepped-care for other mental health conditions such as depression (van Straten et al., 2015).

Although most studies were good quality, the degree to which the data could be synthesized to draw robust conclusions was limited by the diversity of stepped-care models that the studies examined (e.g., whether they were adaptive or matched; the number of steps/LOCs and the types and intensities of interventions in the model; decision rules regarding allocation to and/or movement between LOCs) and the methodologies used to examine them (e.g., study design; measures used; analysis of within- and/or between-groups differences). For example, most adaptive-care evaluations used RCT designs (9/10) but less than half of matched-care studies (3/7); and only four studies (Edelman et al., 2019a; McKay et al., 2010; Merkx et al., 2013; Watson et al., 2013) reported a power calculation *and* recruited a sufficient sample. Reviews of stepped-care approaches to mental health conditions including anxiety and depression have met similar challenges (Firth et al., 2015) and found mixed results (Ho et al., 2016).

Considering these limitations, and consistent with the findings of Jaehne et al.'s (2012) systematic review, we have insufficient evidence to determine the overall efficacy, effectiveness, or cost-effectiveness of stepped-care for the prevention and treatment of problematic substance use. However, some notable findings indicate where stepped-care approaches show promise and highlight areas that warrant further investigation.

4.1.1. Adaptive- vs matched-care

The most promising model to date appears to be adaptive-care for substance use delivered as part of a broader health care program. Four of six studies found that adaptive-care interventions delivered in this context produced greater benefit than TAU or other control conditions in relation to at least one alcohol-related outcome (Edelman et al., 2019a; McKay et al., 2010; McKay et al., 2015; Satre et al., 2019). The convenience of accessing stepped-care through existing health care may have increased participant adherence (Brunner et al., 2013; Drainoni et al., 2014), or the effect of the stepped-care intervention may have been enhanced by targeting multiple health behaviors in the same treatment program (Prochaska et al., 2008).

Limited evidence exists to suggest that either adaptive-care in the absence of other systematic care or matched-care outperform TAU or other control conditions. The findings of the only study to directly compare adaptive- and matched-care models also favored the adaptive-care approach with respect to abstinence, although both groups demonstrated reductions in substance use (Marlowe et al., 2012). Some evidence also exists to suggest that matched-care in combination with adaptive-care may be effective in reducing risky alcohol use (but not other drug use), but to date, this approach has been examined in only one uncontrolled trial (Gómez-Recasens et al., 2018).

4.1.2. Cost-effectiveness

Consistent with the findings of Jaehne et al.'s (2012) systematic

review, Watson et al.'s (2013) economic evaluation found adaptive-care to be associated with lower treatment costs than a minimal intervention control over 12 months, but this difference did not reach statistical significance. The study did find, however, a high probability (>93 %) that adaptive-care would fall below the National Institute for Health and Care Excellence's willingness to pay threshold. However, no studies examine the cost-effectiveness of matched-care interventions.

4.2. Characteristics of stepped-care models associated with improvements in outcomes

4.2.1. Intervention type

Due to the nature of their design, it is difficult to ascertain the relative effectiveness of different intervention modalities and intensities within stepped-care models. Further, although most studies reported the proportion of participants who accessed each LOC in the model, none examined treatment outcomes based on which LOC/s participants were offered during the trial. Attribution of effects to any one treatment modality or intensity is also complicated by the fact that participants with more severe substance use are more likely to be matched, or progressively stepped-up to, the more intensive LOCs offered in a model than those with less severe symptoms.

Despite these limitations, some evidence exists to support including MI-based interventions and psychosocial counselling (both group and individual) in stepped-care models. MI/MET was by far the most common intervention type included in the models ($n = 15$) and was used across the intensity spectrum from BIs to ongoing counselling. Four of the five studies that evaluated stepped-care models offering CBT or other psychosocial counselling found evidence of an association between a greater number of sessions and lower substance use at follow-up (Laporte et al., 2018; Marlowe et al., 2012; Merkx et al., 2013; Stallvik et al., 2015). However, none of these studies directly examined treatment attendance as an outcome mediator, and these studies did not describe in sufficient detail the therapeutic approach they took to determine why it was effective. A systematic review and meta-regression of the effect of psychosocial intervention duration on outcomes found no relationship between number of sessions and alcohol use (Schmidt et al., 2018), possibly because patients with more severe symptoms are allocated to receive more sessions. Considered in this context, the matched-care studies show that stepped-care models should avoid under-treating patients. We have insufficient evidence in this review to draw conclusions on the effects of over-treatment for substance use. This issue warrants further investigation as over-treatment can incur unnecessary costs to the healthcare system (Hammersley, 2011) and evidence from the wider mental health literature that overtreatment can negatively affect patients (Doblytė, 2020; Jerant et al., 2014).

This review found little evidence to support use of BIs, pharmacotherapy and referral to specialist treatment in stepped-care models. As BIs are often the first LOC in adaptive-care models, most if not all participants offered a BI in these studies and treatment effects can at least be attributed to the BI. It is therefore telling that only one of the 10 studies that included a BI found some evidence of model effectiveness, which did not include a control group or offer the BI to all participants (Gómez-Recasens et al., 2018). In fact, Borsari et al. (2012) found that most participants (78 %) did not respond to a BI and required additional care to address risky drinking. The lack of evidence of effectiveness of BIs may be due to most studies (15/19) recruiting participants identified as engaging in problematic substance use (i.e., risky/hazardous/harmful use, or substance use disorder), suggesting that BIs are not sufficient to address substance use that has already become problematic. However, evidence does exist from the broader substance use literature that BIs can reduce even excessive alcohol use (Tansil et al., 2016), any may be more effective when MI-based (DiClemente et al., 2017; Lundahl et al., 2013; VanBuskirk & Wetherell, 2014). McKellar et al. (2012) argue that low-intensity interventions are a key first step in adaptive models as they can lead to further help-seeking. Indeed, a high proportion of Borsari et al.'s (2012) participants (93 %) who did not respond to the BI engaged in the step-up intervention. Patient engagement is a particular challenge in stepped-care models for substance use (Brooner, 2008), and warrants further investigation.

Although several studies offered pharmacotherapies and referral to specialist substance use treatment, the number of participants who engaged in these steps was typically small or unknown. Further, few studies reported on the types of medications offered, and none provided information regarding dosing. Collectively, these constraints limit the degree to which conclusions can be drawn regarding their effectiveness within the context of stepped-care. Further research examining the use of pharmacotherapies and specialist referrals within stepped-care models is needed, including investigation of whether stepped-care models may provide a supportive framework for medication delivery.

Although not the focus of this review, it is worth noting that very few studies incorporated any technology-based interventions, and for those that did, the technology-based aspects were minimal (i.e., provision of feedback via email or web-based psychoeducation) (Edelman et al., 2019a, b; Edelman et al., 2020; Satre et al., 2019). The importance of having readily available evidence-based e-health interventions has come to the fore during the COVID-19 pandemic and should be a priority for inclusion in future research examining stepped-care models.

4.2.2. Decision rules

A notable finding to emerge in studies of matched-care models is the importance of using decision rules that avoid under-treating patients, and match them to either the appropriate LOC or a more intensive LOC (Buchholz et al., 2020; Merkx et al., 2013; Stallvik et al., 2015). Most studies used multifaceted decision rules to determine whether a participant would be stepped-up, or matched to a higher LOC. Notably, none of the models examined incorporated options to step-down, despite evidence for the need for treatment tapering from other mental health domains (Grenyer, 2014; Ngo et al., 2020; Roos et al., 2018). Only five studies considered psychiatric comorbidity in decision rules, a notable omission given the high comorbidity between substance use and other mental health disorders (Kingston et al., 2017). Only four studies included patient preferences in their decision rules (Buchholz et al., 2020; Kay-Lambkin et al., 2010; Laporte et al., 2018; McKay et al., 2015). Although the evidence regarding the impact of patient preference and shared decision-making on clinical outcomes is limited, they are considered essential to the provision of good clinical care in the context of substance use and co-occurring mental health conditions (Fisher et al., 2021).

5. Conclusions and future directions

Despite the ubiquity with which stepped-care approaches appear in clinical guidelines and policy documents, implementation into practice is limited and further research is needed to guide clinical guidelines and policy directives. Research to support the efficacy, effectiveness, and cost-effectiveness of these approaches is limited and questions remain regarding the optimal model framework; levels and components of care to be provided; modes of delivery; and decision rules regarding movement between levels of care. However, the findings of this review indicate that promising evidence does exist with regard to adaptive-care models delivered in the context of other health care programs; motivation-focused interventions and psychosocial counselling as part of stepped models; and designing decision rules to avoid under-treating patients. Sample recruitment and statistical power is a significant challenge for stepped-care models. In addition, a need exists for innovative trial designs like sequential randomization trials (e.g., Borsari et al., 2012; McKay et al., 2015) to achieve robust evidence with regard to their efficacy, effectiveness and cost-effectiveness. For example, models could take a prevention focus and recruit a larger sample with lower levels of substance use, and/or combine adaptive and matched approaches; however, these studies require significant resources to

conduct. Future studies should also examine a range of other drug use and related outcomes (e.g., mental health, wellbeing), and models that include step-down decision rules, to guide the development and implementation of evidence-based approaches to stepped-care for substance use. Greater consistency in study design and methodology would also facilitate comparisons and allow for meta-analyses.

CRediT authorship contribution statement

Ashleigh K Morse: Conceptualization; Methodology; Analysis; Investigation; Writing - Original Draft; Writing - Review & Editing; Supervision; Project administration.

Jayden Sercombe: Methodology; Analysis; Investigation; Writing - Original Draft; Writing - Review & Editing.

Mina Askovic: Methodology; Analysis; Investigation; Writing - Original Draft; Writing - Review & Editing.

Alana Fisher: Conceptualization; Methodology; Writing - Review & Editing; Supervision; Funding acquisition.

Christina Marel: Conceptualization; Methodology; Writing - Review & Editing; Supervision; Funding acquisition.

Mary-Lou Chatterton: Methodology; Analysis; Investigation; Writing - Original Draft; Writing - Review & Editing; Funding acquisition.

Frances Kay-Lambkin: Conceptualization; Methodology; Writing - Review & Editing; Funding acquisition.

Emma Barrett: Conceptualization; Methodology; Writing - Review & Editing; Funding acquisition.

Matthew Sunderland: Conceptualization; Methodology; Writing - Review & Editing; Funding acquisition.

Logan Harvey: Conceptualization; Methodology; Writing - Review & Editing; Funding acquisition.

Natalie Peach: Conceptualization; Methodology; Writing - Review & Editing; Funding acquisition.

Maree Teesson: Conceptualization; Methodology; Writing - Review & Editing; Funding acquisition.

Katherine L Mills: Conceptualization: Methodology; Analysis; Writing - Original Draft; Writing - Review & Editing; Supervision; Funding acquisition.

Declaration of competing interest

The authors declare no financial or non-financial competing interests.

Acknowledgements

Funding for this review was provided by icare NSW. Katherine Mills, Christina Marel, Frances Kay-Lambkin, Emma Barrett, Maree Teesson receive fellowship funding from the Australian National Health and Medical Research Council.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsat.2022.108928.

References

Australian Government Department of Health. (2016). *PHN Primary Mental Health Care Flexible Funding Pool Implementation Guidance: Stepped Care*.

Bair, M. J., Ang, D., Wu, J., Outcalt, S. D., Sargent, C., Kempf, C., Froman, A., Schmid, A. A., Damush, T. M., & Yu, Z. (2015). Evaluation of stepped care for chronic pain (ESCAPE) in veterans of the Iraq and Afghanistan conflicts: A randomized clinical trial. *JAMA Internal Medicine, 175*(5), 682–689.

Borsari, B., Hustad, J. T., Mastroleo, N. R., Tevyaw, T. O. L., Barnett, N. P., Kahler, C. W., Short, E. E., & Monti, P. M. (2012). Addressing alcohol use and problems in mandated college students: A randomized clinical trial using stepped care. *Journal of Consulting and Clinical Psychology, 80*(6), 1062.

Borsari, B., Magill, M., Mastroleo, N. R., Hustad, J. T., Tevyaw, T. O. L., Barnett, N. P., Kahler, C. W., Eaton, E., & Monti, P. M. (2016). Mandated college students' response to sequentially administered alcohol interventions in a randomized clinical trial using stepped care. *Journal of Consulting and Clinical Psychology, 84*(2), 103.

Borsari, B., Short, E. E., Mastroleo, N. R., Hustad, J. T., Tevyaw, T. O. L., Barnett, N. P., Kahler, C. W., & Monti, P. M. (2014). Phone-delivered brief motivational interventions for mandated college students delivered during the summer months. *Journal of Substance Abuse Treatment, 46*(5), 592–596.

Bower, P., & Gilbody, S. (2005). Stepped care in psychological therapies: Access, effectiveness and efficiency: Narrative literature review. *The British Journal of Psychiatry, 186*(1), 11–17.

Brooner, R. K. (2008). Improving treatment engagement in opioid-dependent outpatients with a motivated stepped-care adaptive treatment model. *The Joint Commission Journal on Quality and Patient Safety, 34*(4), 209–213.

Brunner, N., Senn, O., Rosemann, T., Falcato, L., & Bruggmann, P. (2013). Hepatitis C treatment for multimorbid patients with substance use disorder in a primary care-based integrated treatment centre: A retrospective analysis. *European Journal of Gastroenterology & Hepatology, 25*(11), 1300–1307.

Buchholz, A., Dams, J., Rosahl, A., Hempleman, J., König, H.-H., Konnopka, A., Kraus, L., Kriston, L., Piontek, D., & Reimer, J. (2020). Patient-centered placement matching of alcohol-dependent patients based on a standardized intake assessment: Primary outcomes of an exploratory randomized controlled trial. *European Addiction Research, 26*(3), 109–121.

Cross, S. P., & Hickie, I. (2017). Transdiagnostic stepped care in mental health. *Public Health Research & Practice, 27*(2).

DiClemente, C. C., Corno, C. M., Graydon, M. M., Wiprovnick, A. E., & Knoblach, D. J. (2017). Motivational interviewing, enhancement, and brief interventions over the last decade: A review of reviews of efficacy and effectiveness. *Psychology of Addictive Behaviors, 31*(8), 862.

Doblytė, S. (2020). Under-or overtreatment of mental distress? Practices, consequences, and resistance in the field of mental health care. *Qualitative Health Research, 30*(10), 1503–1516.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health, 52*(6), 377–384.

Drainoni, M.-L., Farrell, C., Sorensen-Alawad, A., Palmisano, J. N., Chaisson, C., & Walley, A. Y. (2014). Patient perspectives of an integrated program of medical care and substance use treatment. *AIDS Patient Care and STDs, 28*(2), 71–81.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes*. Oxford University Press.

Edelman, E. J., Maisto, S. A., Hansen, N. B., Cutter, C. J., Dziura, J., Deng, Y., Fiellin, L. E., O'Connor, P. G., Bedimo, R., & Gibert, C. L. (2019a). Integrated stepped alcohol treatment for patients with HIV and alcohol use disorder: A randomised controlled trial. *The Lancet HIV, 6*(8), e509–e517.

Edelman, E. J., Maisto, S. A., Hansen, N. B., Cutter, C. J., Dziura, J., Deng, Y., Fiellin, L. E., O'Connor, P. G., Bedimo, R., & Gibert, C. L. (2019b). Integrated stepped alcohol treatment for patients with HIV and liver disease: A randomized trial. *Journal of Substance Abuse Treatment, 106*, 97–106.

Edelman, E. J., Maisto, S. A., Hansen, N. B., Cutter, C. J., Dziura, J., Deng, Y., Fiellin, L. E., O'Connor, P. G., Bedimo, R., & Gibert, C. L. (2020). Integrated stepped alcohol treatment for patients with HIV and at-risk alcohol use: A randomized trial. *Addiction Science & Clinical Practice, 15*(1), 1–12.

Firth, N., Barkham, M., & Kellett, S. (2015). The clinical effectiveness of stepped care systems for depression in working age adults: A systematic review. *Journal of Affective Disorders, 170*, 119–130. https://doi.org/10.1016/j.jad.2014.08.030

Fisher, A., Mills, K., Teesson, M., & Marel, C. (2021). Shared decision-making among people with problematic alcohol/other drug use and co-occurring mental health conditions: A systematic review. *Drug and Alcohol Review, 40*(2), 307–324.

Gómez-Recasens, M., Alfaro-Barrio, S., Tarro, L., Llauradó, E., & Solà, R. (2018). A workplace intervention to reduce alcohol and drug consumption: A nonrandomized single-group study. *BMC Public Health, 18*(1), 1–10.

Gonzalez-Perez, J. (2002). Developing a scoring system to quality assess economic evaluations. *The European Journal of Health Economics, 3*(2), 131–136.

Graham, K., Cheng, J., Bernards, S., Wells, S., Rehm, J., & Kurdyak, P. (2017). How much do mental health and substance use/addiction affect use of general medical services? Extent of use, reason for use, and associated costs. *The Canadian Journal of Psychiatry, 62*(1), 48–56.

Grenyer, B. F. (2014). An integrative relational step-down model of care: The project air strategy for personality disorders. *The ACPARIAN, 9*(8–13).

Hammersley, R. (2011). Pathways through drugs and crime: Desistance, trauma and resilience. *Journal of Criminal Justice, 39*(3), 268–272.

Ho, F. Y.-Y., Yeung, W.-F., Ng, T. H.-Y., & Chan, C. S. (2016). The efficacy and cost-effectiveness of stepped care prevention and treatment for depressive and/or anxiety disorders: A systematic review and meta-analysis. *Scientific Reports, 6*(1), 29281. https://doi.org/10.1038/srep29281

Jaehne, A., Loessl, B., Frick, K., Berner, M., Hulse, G., & Balmford, J. (2012). The efficacy of stepped care models involving psychosocial treatment of alcohol use disorders and nicotine dependence: A systematic review of the literature. *Current Drug Abuse Reviews, 5*(1), 41–51.

Jerant, A., Kravitz, R. L., Garcia, E. F., Feldman, M. D., Cipri, C., Nishio, D., Knoepfler, A., Wooddell, M. K., Baquero, V., & Franks, P. (2014). Potential antidepressant overtreatment associated with office use of brief depression symptom measures. *The Journal of the American Board of Family Medicine, 27*(5), 611–620.

Kalarchian, M. A., & Marcus, M. D. (2018). The case for stepped care for weight management after bariatric surgery. *Surgery for Obesity and Related Diseases, 14*(1), 112–116. https://doi.org/10.1016/j.soard.2017.07.023

Kay-Lambkin, F. J., Baker, A. L., McKetin, R., & Lee, N. (2010). Stepping through treatment: Reflections on an adaptive treatment strategy among methamphetamine users with depression. *Drug and Alcohol Review, 29*(5), 475–482.

Kingston, R. E., Marel, C., & Mills, K. L. (2017). A systematic review of the prevalence of comorbid mental health disorders in people presenting for substance use treatment in Australia. *Drug and Alcohol Review, 36*(4), 527–539.

Korakakis, V., Whiteley, R., Tzavara, A., & Malliaropoulos, N. (2018). The effectiveness of extracorporeal shockwave therapy in common lower limb conditions: A systematic review including quantification of patient-rated pain reduction. *British Journal of Sports Medicine, 52*(6), 387–407.

Laporte, L., Paris, J., Bergevin, T., Fraser, R., & Cardin, J. F. (2018). Clinical outcomes of a stepped care program for borderline personality disorder. *Personality and Mental Health, 12*(3), 252–264.

Lundahl, B., Moleni, T., Burke, B. L., Butters, R., Tollefson, D., Butler, C., & Rollnick, S. (2013). Motivational interviewing in medical care settings: A systematic review and meta-analysis of randomized controlled trials. *Patient Education and Counseling, 93*(2), 157–168. https://doi.org/10.1016/j.pec.2013.07.012

Maehder, K., Löwe, B., Härter, M., Heddaeus, D., Scherer, M., & Weigel, A. (2018). Management of comorbid mental and somatic disorders in stepped care approaches in primary care: A systematic review. *Family Practice, 36*(1), 38–52. https://doi.org/10.1093/fampra/cmy122

Marlowe, D. B., Festinger, D. S., Dugosh, K. L., Benasutti, K. M., Fox, G., & Croft, J. R. (2012). Adaptive programming improves outcomes in drug court: An experimental trial. *Criminal Justice and Behavior, 39*(4), 514–532.

McKay, J. R., Drapkin, M. L., Van Horn, D. H., Lynch, K. G., Oslin, D. W., DePhilippis, D., Ivey, M., & Cacciola, J. S. (2015). Effect of patient choice in an adaptive sequential randomization trial of treatment for alcohol and cocaine dependence. *Journal of Consulting and Clinical Psychology, 83*(6), 1021.

McKay, J. R., Van Horn, D. H., Oslin, D. W., Lynch, K. G., Ivey, M., Ward, K., Drapkin, M. L., Becher, J. R., & Coviello, D. M. (2010). A randomized trial of extended telephone-based continuing care for alcohol dependence: Within-treatment substance use outcomes. *Journal of Consulting and Clinical Psychology, 78*(6), 912.

McKellar, J., Austin, J., & Moos, R. (2012). Building the first step: A review of low-intensity interventions for stepped care. *Addiction Science & Clinical Practice, 7*(1), 1–14.

Merkx, M. J., Schippers, G. M., Koeter, M. W., Vuijk, P. J., Oudejans, S. C., Stam, R. K., & van den Brink, W. (2011). Guidelines for allocating outpatient alcohol abusers to levels of care: Predictive validity. *Addictive Behaviors, 36*(6), 570–575.

Merkx, M. J., Schippers, G. M., Koeter, M. W., Vuijk, P. J., Poch, M., Kronemeijer, H., & van den Brink, W. (2013). Predictive validity of treatment allocation guidelines on drinking outcome in alcohol-dependent patients. *Addictive Behaviors, 38*(3), 1691–1698.

Mewton, L., Visontay, R., Chapman, C., Newton, N., Slade, T., Kay-Lambkin, F., & Teesson, M. (2018). Universal prevention of alcohol and drug use: An overview of reviews in an Australian context. *Drug and Alcohol Review, 37*, S435–S469.

Ngo, H., Ennals, P., Turut, S., Geelhoed, E., Celenza, A., & Wolstencroft, K. (2020). Step-up, step-down mental health care service: Evidence from western Australia's first–A mixed-method cohort study. *BMC Psychiatry, 20*, 1–13.

Prendergast, M. L., McCollister, K., & Warda, U. (2017). A randomized study of the use of screening, brief intervention, and referral to treatment (SBIRT) for drug and alcohol use with jail inmates. *Journal of Substance Abuse Treatment, 74*, 54–64.

Peterson, K., Anderson, J., Bourne, D., Mackey, K., & Helfand, M. (2018). Effectiveness of models used to deliver multimodal care for chronic musculoskeletal pain: a Rapid evidence review. *Journal of General Internal Medicine, 33*(1), 71–81. https://doi.org/10.1007/s11606-018-4328-7

Prochaska, J. J., Spring, B., & Nigg, C. R. (2008). Multiple health behavior change research: An introduction and overview. *Preventive Medicine, 46*(3), 181–188.

Reeves, P., Szewczyk, Z., Proudfoot, J., Gale, N., Nicholas, J., & Anderson, J. (2019). Economic evaluations of stepped models of care for depression and anxiety and associated implementation strategies: A review of empiric studies. *International Journal of Integrated Care, 19*(2).

Roos, E., Bjerkeset, O., & Steinsbekk, A. (2018). Health care utilization and cost after discharge from a mental health hospital; an RCT comparing community residential aftercare and treatment as usual. *BMC Psychiatry, 18*(1), 1–13.

Sacks, J. J., Gonzales, K. R., Bouchery, E. E., Tomedi, L. E., & Brewer, R. D. (2015). 2010 national and state costs of excessive alcohol consumption. *American Journal of Preventive Medicine, 49*(5), e73–e79. https://doi.org/10.1016/j.amepre.2015.05.031

Satre, D. D., Leibowitz, A. S., Leyden, W., Catz, S. L., Hare, C. B., Jang, H., Lam, J. O., Bryant, K. J., Weisner, C. M., & Sterling, S. A. (2019). Interventions to reduce unhealthy alcohol use among primary care patients with HIV: The health and motivation randomized clinical trial. *Journal of General Internal Medicine, 34*(10), 2054–2061.

Schmidt, L. K., Bojesen, A. B., Nielsen, A. S., & Andersen, K. (2018). Duration of therapy – Does it matter?: A systematic review and meta-regression of the duration of psychosocial treatments for alcohol use disorder. *Journal of Substance Abuse Treatment, 84*, 57–67. https://doi.org/10.1016/j.jsat.2017.11.002

Stallvik, M., Gastfriend, D. R., & Nordahl, H. M. (2015). Matching patients with substance use disorder to optimal level of care with the ASAM criteria software. *Journal of Substance Use, 20*(6), 389–398.

Tansil, K. A., Esser, M. B., Sandhu, P., Reynolds, J. A., Elder, R. W., Williamson, R. S., Chattopadhyay, S. K., Bohm, M. K., Brewer, R. D., McKnight-Eily, L. R., Hungerford, D. W., Toomey, T. L., Hingson, R. W., & Fielding, J. E. (2016). Alcohol electronic screening and brief intervention: A community guide systematic review. *American Journal of Preventive Medicine, 51*(5), 801–811. https://doi.org/10.1016/j.amepre.2016.04.013

van Straten, A., Hill, J., Richards, D. A., & Cuijpers, P. (2015). Stepped care treatment delivery for depression: A systematic review and meta-analysis. *Psychological Medicine, 45*(2), 231–246.

Van Straten, A., Tiemens, B., Hakkaart, L., Nolen, W., & Donker, M. (2006). Stepped care vs. Matched care for mood and anxiety disorders: A randomized trial in routine practice. *Acta Psychiatrica Scandinavica, 113*(6), 468–476.

VanBuskirk, K. A., & Wetherell, J. L. (2014). Motivational interviewing with primary care populations: A systematic review and meta-analysis. *Journal of Behavioral Medicine, 37*(4), 768–780.

Watson, J., Crosby, H., Dale, V., Tober, G., Wu, Q., Lang, J., McGovern, R., Newbury-Birch, D., Parrott, S., & Bland, J. (2013). AESOPS: A randomised controlled trial of the clinical effectiveness and cost-effectiveness of opportunistic screening and stepped care interventions for older hazardous alcohol users in primary care. *Health Technology Assessment, 17*(25).

Woodruff, S. I., Clapp, J. D., Eisenberg, K., McCabe, C., Hohman, M., Shillington, A. M., Sise, C. B., Castillo, E. M., Chan, T. C., & Sise, M. J. (2014). Randomized clinical trial of the effects of screening and brief intervention for illicit drug use: The life shift/shift gears study. *Addiction Science & Clinical Practice, 9*(1), 1–11.

Yurasek, A. M., Borsari, B., Magill, M., Mastroleo, N. R., Hustad, J. T., Tevyaw, T. O. L., Barnett, N. P., Kahler, C. W., & Monti, P. M. (2015). Descriptive norms and expectancies as mediators of a brief motivational intervention for mandated college students receiving stepped care for alcohol use. *Psychology of Addictive Behaviors, 29*(4), 1003.

Yurasek, A. M., Merrill, J. E., Metrik, J., Miller, M. B., Fernandez, A. C., & Borsari, B. (2017). Marijuana use in the context of alcohol interventions for mandated college students. *Journal of Substance Abuse Treatment, 79*, 53–60.