Alcohol and Alcoholism, 2021, 56(3) 284–290
doi: 10.1093/alcalc/agaa074
Advance Access Publication Date: 18 August 2020
Article

Article

Evaluation of Risks of Bias in Addiction Medicine Randomized Controlled Trials

Cole Wayant*, Daniel Tritz, Jarryd Horn, Matt Crow, and Matt Vassar

Oklahoma State University Center for Health Sciences, Tulsa, OK 74104, USA

*Corresponding author: Oklahoma State University Center for Health Sciences, 1111 W 17th Street, Tulsa, OK 74104, USA. E-mail: cole.wayant@okstate.edu

Received 20 November 2019; Revised 6 July 2020; Editorial Decision 7 July 2020; Accepted 7 July 2020

Abstract

Aims: Perhaps the most important step when designing and conducting randomized controlled trials (RCTs) in addiction is to put methodological safeguards in place to minimize the likelihood for bias to affect trial outcomes. In this study, we applied the revised Cochrane risk of bias tool (ROB 2) to RCTs of drug, alcohol or tobacco interventions.

Methods: We searched for trials published in 15 addiction medicine journals over a 7-year period. Our primary endpoint is the risk of bias of included studies. We conducted a sensitivity analysis of publicly funded trials.

Results: Overall, included RCTs were most often at high risk of bias per our judgments (244/487, 50.1%). However, significant proportions of included RCTs were at low risk of bias (123/487, 25.3%) or some concerns for bias (120/497, 24.6%). RCTs with behavioral modification interventions (19/44, 43.2%) and alcohol interventions (80/150, 53.3%) had the highest proportion of high-risk judgments. In a sensitivity analysis of publicly funded RCTs), 195/386 (50.5%) were at high risk of bias.

Conclusions: Approximately half of included drug, alcohol or tobacco RCTs in our sample were judged to be at high risk of bias with the most common reason being a lack of proper blinding or proper description of blinding. Key action items to reduce bias in future addiction RCTs include adequate randomization, blinding and inclusion of a trial registry number and protocol.

INTRODUCTION

Randomized controlled trials (RCTs) have led to a number of significant advancements in the treatment of drug, alcohol and tobacco dependence. For example, numerous trials have shown that varenicline is capable of increasing rates of smoking cessation (Koegelenberg, Noor, and Bateman, 2014; Ebbert *et al.*, 2015; Jhanjee *et al.*, 2015). At the request of the National Institute on Drug Abuse (NIDA), the National Academies of Medicine, in a report about the behavioral, ethical, legal and social issues surrounding novel immunotherapeutic medications, stated that 'NIDA may wish to encourage research into these issues in parallel with—if not integrated into—clinical trials' (National Research Council *et al.*, 2004). Establishing RCTs as a funding priority may increase the validity of new research findings, since RCTs are generally believed to be the highest level of primary research evidence (Burns, Rohrich, and Chung, 2011). However, for RCTs to generate evidence-based results, they must first exhibit robust design to prevent bias.

Perhaps the most important step when designing and conducting RCTs in addiction is to put methodological safeguards in place to minimize the likelihood for bias to affect trial outcomes. Bias may arise from low-quality *a priori* design, poor study conduct and poor data analysis. For example, if authors choose *a priori* to conduct an unblinded trial, wherein the investigators, patient or both know what intervention is being given, the subsequent trial results may be biased (Schulz *et al.*, 1995; Moher *et al.*, 1998). Similarly, if during the study efforts to maintain strong intervention compliance are not put forth, unacceptable rates of attrition may ensue, which can result in unbalanced treatment groups and potential bias

(Nunan, Aronson, and Bankhead, 2018). Last, if measures are not taken to adhere to predefined endpoints and analyses, selective outcome reporting and p-hacking may occur, wherein authors choose to highlight statistically significant findings at the expense of non-significant results (Chan *et al*., 2004; Head *et al*., 2015; Wayant *et al*., 2017). Altogether, conducting an RCT that is at low risk of bias is a weighty endeavor; however, the consequences of RCTs at high risk of bias are severe enough to mandate investigations.

While risk of bias has been studied in other fields, including HIV/AIDS (Wayant *et al*., 2019), otolaryngology (Skinner *et al*., 2019), emergency medicine (Brown *et al*., 2019) and antithrombotic therapy (Edwards *et al*., 2018), we know very little about the risk of bias of addiction trials. The literature is surprisingly absent in this regard. Cochrane reviews performed by the Drug and Alcohol group provide some initial indications that systematic review summary effects are often limited by studies judged to be at high or moderate risk of bias due to multiple factors, such as outcome reporting or allocation concealment (Minozzi *et al*., 2010, 2015, Rösner et al., 2010; Gowing *et al*., 2014; Pani et al., 2014). The absence of dedicated risk of bias assessments for a large cohort of RCTs of addiction medicine interventions forms the basis for the present study.

Recently, we have shown that data-sharing practices in addiction RCTs are lacking (Vassar *et al*., 2020), RCT reporting may be improved (Vassar *et al*., 2019a), and bias from selective outcome reporting is prevalent (Vassar *et al*., 2019b). This investigation builds on those previous by examining global risk of bias in a cross section of RCTs published in addiction medicine journals. We applied the revised Cochrane risk of bias tool (ROB 2) to RCTs of drug, alcohol or tobacco interventions. The revised ROB 2 tool is currently the most comprehensive and robust risk of bias tool for clinical trials. ROB 2 was designed, and is recommended, by members of the Cochrane Collaboration to assess five Domains of bias that have been empirically shown to affect RCT results (Higgins *et al*., 2019). The structure of ROB 2 is unique and improves upon previous risk of bias tools. For example, ROB 2 uses decision tree algorithms and signaling questions to objectively guide an investigator to a risk of bias judgment, rather than relying on the investigator's subjective judgment or feeling. In this study, our primary objective is to investigate the overall risk of bias of included trials. Our secondary objective is to explore the Domains of bias that contribute to overall judgments. We further conducted a sensitivity analysis of publicly funded RCTs to determine the risk of bias for RCTs dependent on tax-payer dollars.

METHODS

Our study did not meet the regulatory definition of human subject research and did not require Institutional Review Board oversight. We followed the published guidance for reporting meta-epidemiological (Murad and Wang, 2017) and, where relevant, systematic reviews (Moher *et al*., 2009).

Using PubMed (which includes Medline), we searched for trials published in addiction medicine journals over a 7-year period (2013–2019). We selected the 15 highest-ranking addiction journals, as indexed by Google Scholar, by sequentially searching each journal to determine whether at least 10 clinical trials were published in them. Our database search strategy was chosen to maximize sensitivity across the included journals and date parameter (Hoogendam *et al*., 2009). The exact strategy was as follows: (*Addiction* (Abingdon, England) [Journal] OR *Drug and Alcohol Dependence* [Journal] OR *Nicotine & Tobacco Research*: official journal of the Society for Research on Nicotine and Tobacco [Journal] OR *Addictive Behaviors* [Journal] OR *Alcoholism: Clinical and Experimental Research* [Journal] OR *Psychology of Addictive Behaviors*: *Journal of the Society of Psychologists in Addictive Behaviors* [Journal] OR *Addiction Biology* [Journal] OR *Journal of Studies on Alcohol and Drugs* [Journal] OR *International Journal on Drug Policy* [Journal] OR *Drug and Alcohol Review* [Journal] OR *Alcohol and Alcoholism* (Oxford, Oxfordshire) [Journal] OR *Journal of Substance Abuse Treatment* [Journal] OR *Alcohol* (Fayetteville, NY) [Journal] OR *The American Journal on Addictions* [Journal] OR *Substance Use & Misuse* [Journal] AND Clinical Trial[ptyp] AND '2013/01/01' [PDat]: '2017/12/31' [PDat]). After peer review feedback, we supplemented our search by duplicating the search strategy above with the exception of a date range filter of 01 January 2018 to 31 December 2019.

Included studies were screened independently by two investigators who were unaware of each other's responses, consistent with best practice methods (Checketts *et al*., 2018; Wayant *et al*., 2018; Austin *et al*., 2018), using Rayyan (Ouzzani *et al*., 2016) to determine whether studies met the inclusion criteria. After initial screening, all discrepancies were resolved by group discussion. These criteria were randomized clinical trials that addressed one of the following related to drugs, alcohol or tobacco: (1) addiction prevention, (2) stabilization following excessive use of a substance, (3) relapse prevention and (4) recovery maintenance. We used the National Institutes of Health definition of an RCT (National Institutes of Health, 2019). Thus, we excluded observational studies, letters to the editor, meta-analyses and other studies not using RCT designs.

Data were extracted from all included studies independently by two investigators who were unaware of each other's responses. Extracted data were the journal in which a trial was published, the trial's funding source and items comprising the ROB 2 tool. The ROB 2 tool covers five Domains of bias, each with signaling questions that guide an investigator through a decision tree algorithm to arrive at a Domain risk of bias judgment. The risk of bias Domains are as follows (Box 1 in the supplement): Domain 1 (randomization process), Domain 2 (deviations from the intended interventions), Domain 3 (missing outcome data), Domain 4 (measurement of the outcome) and Domain 5 (selection of the reported results). The signaling questions regarding these bias Domains prompt bias assessors to respond to each question using one of the five anchors: yes, probably yes, no, probably no or no information. Based on a bias assessor's responses to all signaling questions within a Domain, they are directed to a Domain bias judgment via a decision algorithm. To view these decision algorithms and signaling questions, we recommend the readers review the ROB 2 guidance document (Sterne *et al*., 2019).

After each investigator completed risk of bias assessments for all trials, the two investigators resolved discrepancies. We did not resolve discrepancies for each signaling question; rather, discrepancies were resolved only for Domain bias judgments. The rationale for this decision was that the Domain judgment, rather than the signaling question response, contributes to the overall bias judgment. Overall bias judgments were decided based on prespecified criteria. To be judged as 'Low Risk' overall, a trial must be rated as 'low risk' for each bias Domain. To be judged 'High Risk', a trial must be rated 'high risk' for a single bias Domain or be rated as 'some concerns' for a multiple (2+/5) Domains. To be rated as 'Some Concerns', a trial must be rated as 'some concerns' for exactly 1 bias Domain.

We categorized the included RCTs by (1) substance (tobacco, alcohol drug); (2) intervention modality; and (3) treatment setting (e.g. addiction prevention, stabilization after excessive consumption,

Box 1. Signaling questions

Domain 1: Randomization process	Was the allocation sequence random?	Was the allocation sequence properly concealed?	Are there baseline imbalances that suggest a problem with the randomization process?		
Domain 2: Deviations from intended interventions	Were participants aware of their group assignment?	Were those delivering the intervention aware of group assignments?	If yes, were there deviations from the intended intervention because of trial context?	If yes, were these deviations likely to have affected the trial outcome?	
Domain 3: Missing outcome data	Were data available for all or nearly all randomized participants?	If no, is there evidence that the outcome was not biased?	If no, could missingness in the outcome depend on its true value?	If yes, is it likely that missingness in the outcome depended on its true value?	
Domain 4: Outcome measurement	Was the outcome measurement appropriate?	Could measurement have differed between groups?	If no to both previous questions, were outcome assessors aware of the group assignments?	If yes, could outcome assessment be affected by this knowledge?	If yes, is it likely that outcome assessment was influenced by knowledge of intervention received?
Domain 5: Selection of reported result	Were data analyzed according to pre-specified plans?	Is the numerical result likely to have been selected, on the basis of results, from multiple eligible outcomes or multiple analyses of outcomes?			

relapse prevention, recovery maintenance). Summary statistics were calculated using Google Sheets. We conducted a sensitivity analysis of publicly funded RCTs. No further statistical analyses were planned.

RESULTS

Our database search returned 1846 results, of which 487 randomized controlled trials were included (Fig. 1). The median sample size of included RCTs was 130 patients (IQR 75–291.5). These RCTs were most often published in the journals *Drug and Alcohol Dependence* ($n = 99$, 20.3%), *Nicotine & Tobacco Research* ($n = 74$, 15.2%) and *Addiction* ($n = 71$, 14.6%) (Table 1). Included trials were most often funded by public entities (e.g. government; $n = 386$, 79.3%), followed by private sources (e.g. foundation or nonprofit = 29, 6.0%), mixed sources [$n = 19$ (3.9%) that did not include industry; $n = 16$ (3.3%) that included industry] and industry alone ($n = 16$, 3.3%). Interventions related to tobacco were most common ($n = 158$), 32.4%, followed by alcohol ($n = 150$, 30.8%)- and drug-related ($n = 126$, 25.9%) interventions. Interventions were most often designed to reduce use or cravings ($n = 274$, 56.3%).

Overall, included RCTs were most often at high risk of bias (Table 2; Fig. 2), per our judgments (244/487, 50.1%). However, significant proportions of included RCTs were at low risk of bias (123/487, 25.3%) or some concerns for bias (120/497, 24.6%). Two risk of bias Domains were overwhelmingly low risk (Domain 2 = 442/497 (90.8%) (Domain 4 = 456/487 (93.6%)). Domain 3, while still mostly low risk (403/487 (82.8%)), had higher rates of high-risk judgment (44/487 (9.0%)). Finally, Domains 1 and 5 had the lowest proportion of low risk of bias judgments. Domain 1 had the highest rate of high-risk judgments (53/497 (10.9%)), often due to a lack of allocation concealment and baseline imbalances between groups. Domain 5 had the highest rate of some concerns (237/487 (48.7%)). The risk of bias of included intervention modalities is shown in Table 3, with RCTs with behavioral modification interventions most often being at high risk of bias (19/44, 43.2%). Analysis by substance type is shown in Table 4, with alcohol interventions most often being at high risk of bias (80/150, 53.3%).

In a sensitivity analysis of publicly funded RCTs (Table 5), 195/386 (50.5%) were at high risk of bias, nearly identical to the overall sample. High risk of bias was most often given to Domain 1 (randomization and blinding) (46/386, 11.9%). Similar to the overall population of included trials, Domain 5 was at some concerns for bias most often (95/386, 24.6%). When comparing publicly funded RCTs by intervention modality, medication interventions were most common (111/386, 28.8%), and 'Other' interventions, including brain stimulation or situational/simulation exposure interventions, were at the highest risk of bias (17/24, 70.8%).

DISCUSSION

Results of this investigation indicate that drug, alcohol or tobacco addiction RCTs comprising our sample were most often at high risk of bias. Approximately half of included RCTs were judged to be at high risk of bias with the most common reason being a lack of proper blinding or proper description of blinding. The way in which trial authors selected their reported results was often suspected for bias, because such determinations require a trial registration or protocol. Neither of these items were included in a majority of RCTs comprising our sample. Previous studies in which the Cochrane ROB 2 tool has been applied have found that trials across biomedicine commonly have some concerns or are at high risk of bias, with rates of high-risk trials ranging from 7.6% to 48.6% (Bowers *et al.*, 2018; Skinner *et al.*, 2019; Wayant *et al.*, 2019; Meyer *et al.*, 2019, 2020; Goerke *et al.*, 2020). Our study has significantly more included RCTs and differs from these previous studies by evaluating all journals' published RCTs, rather than those cited by clinical practice guidelines. All of these findings together have significant implications for the design and interpretation of drug, alcohol or tobacco addiction RCTs.

The first implication of RCTs that are at high risk for bias is that results from such studies might yield unstable results and potentially spurious conclusions. Bias in the design, analysis or reporting of trial results may lead to conclusions that are the result of the bias itself rather than the intervention effect. For example, a lack of blinding is known to lead to potentially flawed conclusions, since a lack of blinding is known to exaggerate treatment effects (Schulz *et al.*, 1995; Moher *et al.*, 1998). Thus, a lack of blinding in addiction medicine RCTs may lead authors to conclude that strategies to treat addiction are more effective than they actually are. To mitigate the effect of

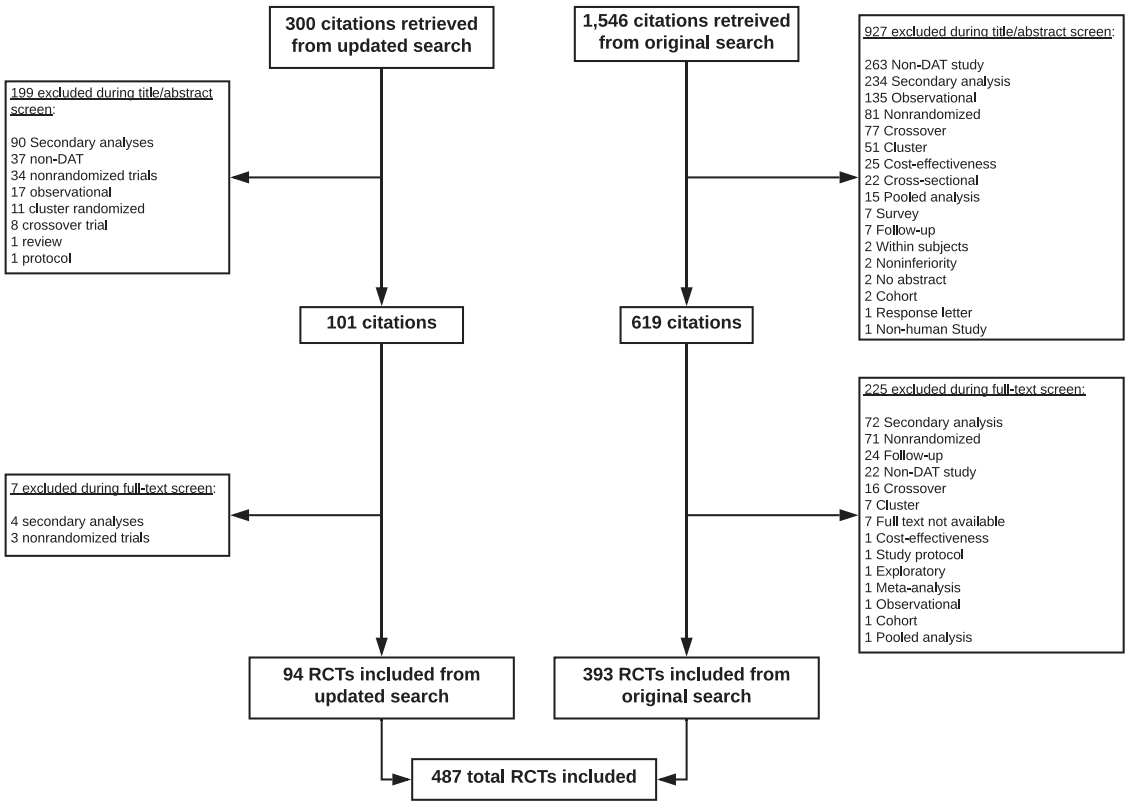

Fig. 1. Flow diagram of included and excluded studies from our original search and the updated search conducted during the peer review process. DAT = drug, alcohol or tobacco; RCT = randomized controlled trial.

unblinded interventions, authors may choose to measure objective endpoints, such as mortality (Evans, 2010). In addiction medicine trials, self-reported measures may be subject to bias if blinding is not maintained. Our finding that publicly funded RCTs were less often at risk of bias in Domain 2—which considers whether consequences from a lack of blinding were present—than the overall population of included RCTs is significant because it indicates that publicly funded trials are more likely to measure endpoints that are not likely to be affected by a lack of blinding. However, the rate of high risk of bias judgments in Domain 1 (randomization and blinding) must be further explored.

In our sample, urn randomization was commonly employed. To illustrate urn randomization, consider an urn with equal numbers of identical red and black balls. To randomize a patient, an investigator would reach into the urn and select one ball. If red, the patient would be randomized to the red group. The red ball would then be placed back into the urn along with an additional black ball. Now, the urn would contain n red balls and $n + 1$ black balls. Thus, group sample size differences should resolve over time if imbalances occur. Urn randomization was common in our sample and may have been a contributor to the high risk of bias judgments. Previous simulation studies have shown that urn randomization exhibits fewer baseline imbalances than other forms of randomization in substance abuse trials (Hedden et al., 2006). However, it has also been shown that urn randomization is susceptible to temporal trends in the recruitment of patients, which may lead to baseline imbalances and bias (Friedman *et al.*, 2015), which were common in our sample. Thus, it is possible that the presence of urn randomization contributed to the high risk of bias judgments in our sample. In future clinical trials of substance abuse interventions, we recommend authors employ safeguards to balance trial groups and protect urn randomization from temporal trends. Similarly, since lack of blinding was also common, we recommend that authors carefully describe how blinding is maintained and that authors choose to measure objective endpoints that are not subject to bias should blinding toward the intervention not be possible.

The second implication, related to Domain 5 (selection of the reported result), is that included drug, alcohol or tobacco addiction RCTs infrequently reported a trial registry number or protocol such that reported results could be compared to analyses planned *a priori*. Similar findings have been shown previously when ROB 2 was applied to RCTs in other areas of medicine (Chase Kruse and Matt Vassar, 2017; Edwards *et al.*, 2018; Bowers *et al.*, 2018; Wayant *et al.*, 2019). The goal of trial registration is to reduce two key forms of bias: publication bias and selective outcome reporting bias. Publication bias occurs when studies with nonsignificant results are published less often than studies with statistically significant findings (DeVito and Goldacre, 2019). The result of publication bias is a predominance of statistically significant findings in the published literature, which has been shown to bias notions of perceived intervention efficacy (Turner *et al.*, 2008). Similarly, selective outcome reporting bias is the addition, omission or modification of study endpoints in published trial reports to highlight statistically significant findings (Hutton and Williamson, 2000). Results of selective outcome reporting bias are published articles that distract from, or omit, nonsignificant results. Trial registration allows authors to state the planned analyses and endpoints prior to trial commencement. Thus, trial registration may decrease the risk of bias for RCTs, and we recommend all journals

Table 1. Demographic characteristics of included RCTs ($n = 487$)

Journal	Drug and Alcohol Dependence	99 (20.3%)
	Nicotine and Tobacco Research	74 (15.2%)
	Addiction	71 (14.6%)
	Journal of Substance Abuse Treatment	62 (12.7%)
	Addictive Behaviors	48 (9.9%)
	Alcoholism: Clinical and Experimental Research	42 (8.6%)
	Psychology of Addictive Behaviors	24 (4.9%)
	Alcohol and Alcoholism	18 (3.7%)
	The American Journal on Addictions	16 (3.3%)
	Journal of Studies on Alcohol and Drugs	12 (2.5%)
	Substance Use & Misuse	12 (2.5%)
	Drug and Alcohol Review	5 (1.0%)
	Addiction Biology	3 (0.6%)
	International Journal on Drug Policy	1 (0.2%)
Funding source	Public	386 (79.3%)
	Private	29 (6.0%)
	Mixed (without industry)	19 (3.9%)
	Industry	16 (3.3%)
	Mixed (with industry)	16 (3.3%)
	None	9 (1.8%)
	Not mentioned	8 (1.6%)
	University	3 (0.6%)
	Self	1 (0.2%)
Intervention modality	Medication	158 (32.4%)
	Intervention delivery system	120 (24.6%)
	Therapy	104 (21.4%)
	Behavioral modification	44 (9.0%)
	Other	31 (6.4%)
	Financial incentive	30 (6.2%)
Substance Domain	Tobacco	158 (32.4%)
	Alcohol	150 (30.8%)
	Drug	126 (25.9%)
	Mixed	53 (10.9%)
Treatment strategy	Reduced use/craving	274 (56.3%)
	Recovery maintenance	151 (31.0%)
	Mixed	22 (4.5%)
	Other (e.g. improved medication adherence)	20 (4.1%)
	Prevention (of initial addiction)	14 (2.9%)
	Stabilization after excess ingestion (overdose)	6 (1.2%)
Sample size	Median (IQR) = 130 (IQR 75–291.5)	

Table 2. Risk of bias for all studies ($n = 487$) for each bias domain

	Total	Domain 1 randomization process	Domain 2 deviations from intended intervention	Domain 3 missing data	Domain 4 outcome measurement	Domain 5 selection of results
Low risk	123 (25.3%)	246 (50.5%)	442 (90.8%)	403 (82.8%)	456 (93.6%)	242 (49.7%)
Some concern	120 (24.6%)	188 (38.6%)	29 (6.0%)	40 (8.2%)	23 (4.7%)	237 (48.7%)
High risk	244 (50.1%)	53 (10.9%)	16 (3.3%)	44 (9.0%)	8 (1.6%)	8 (1.6%)

Note: A trial may be high risk of bias overall if any single Domain is at high risk or if at least two Domains are some concerns.

require authors to publish a trial registration number, protocol and statistical analysis plan. An explanation of deviations from the registration or protocol may further mitigate bias.

This study has several strengths and limitations. First, our study benefited from dual, independent data extraction, which increases the rigor of our findings and decreases the risk of bias. Second, we investigated RCTs in 15 highly read addiction journals over a 7-year timeframe, which increases the relevance to addiction care providers and the generalizability of our study. However, our study is not completely generalizable. We only included RCTs of drug, alcohol and tobacco interventions, which may limit the external validity of our findings to other fields of medicine. Third, by searching only

Table 3. Risk of bias for all studies ($n = 487$), stratified by intervention modality

	High risk	Some concern	Low risk
Medication ($n = 158$)	37 (23.4%)	30 (19.0%)	91 (57.6%)
Therapy* ($n = 104$)	32 (30.8%)	29 (27.9%)	43 (41.3%)
Intervention delivery system ($n = 120$)	30 (25.0%)	39 (32.5%)	51 (42.5%)
Behavioral modification ($n = 44$)	19 (43.2%)	15 (34.1%)	10 (22.7%)
Financial incentive ($n = 30$)	11 (36.7%)	8 (26.7%)	11 (36.7%)
Other** ($n = 31$)	10 (32.3%)	10 (32.3%)	11 (35.5%)

*Included cognitive–behavioral or other interpersonal therapies

**includes situational exposures, diagnostic strategies, brain stimulation and simulation exposures.

Table 4. Risk of bias for all studies ($n = 487$), stratified by substance

	High risk	Some concern	Low risk
Drug ($n = 126$)	57 (45.2%)	32 (25.4%)	37 (29.4%)
Alcohol ($n = 150$)	80 (53.3%)	32 (21.3%)	38 (25.3%)
Tobacco ($n = 158$)	80 (50.6%)	43 (27.2%)	35 (22.2%)
Mixed ($n = 53$)	27 (50.9%)	13 (24.5%)	13 (24.5%)

Table 5. Risk of bias for all publicly funded studies ($n = 386$), stratified by intervention modality

	High risk	Some concern	Low risk
Medication ($n = 111$)	40 (36.0%)	31 (27.9%)	40 (36.0%)
Behavioral modification ($n = 36$)	25 (69.4%)	4 (11.1%)	7 (19.4%)
Therapy ($n = 86$)	44 (51.2%)	23 (26.7%)	19 (22.1%)
Financial incentive ($n = 30$)	19 (63.3%)	8 (26.7%)	3 (10.0%)
Intervention delivery system ($n = 99$)	50 (50.5%)	25 (25.3%)	24 (24.2%)
Other ($n = 24$)	17 (70.8%)	4 (16.7%)	3 (12.5%)

addiction journals for RCTs, we have excluded any RCTs of addiction interventions published in general medical journals. Therefore, our study should not be generalized to studies outside the selected journals.

In conclusion, most included drug, alcohol and tobacco addiction RCTs, either had some concerns or high risk of bias. Key action items to reduce bias in future addiction RCTs include adequate randomization, blinding and inclusion of a trial registry number and protocol.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

Austin J, Smith C, Natarajan K, *et al.* (2018) Evaluation of spin within abstracts in obesity randomized clinical trials: A cross-sectional review. *Clin Obes* **e12292**.

Bowers A, Meyer C, Tritz D, *et al.* (2018) Assessing quality of randomized trials supporting guidelines for laparoscopic and endoscopic surgery. *J Surg Res* **224**:233–9.

Brown J, Lane A, Cooper C, *et al.* (2019) The results of randomized controlled trials in emergency medicine are frequently fragile. *Ann Emerg Med* **73**:565–76.

Burns PB, Rohrich RJ, Chung KC. (2011) The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg* **128**:305–10.

Chan A-W, Hróbjartsson A, Haahr MT, *et al.* (2004) Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *JAMA* **291**:2457–65.

Chase Kruse B, Matt Vassar B. (2017) Unbreakable? An analysis of the fragility of randomized trials that support diabetes treatment guidelines. *Diabetes Res Clin Pract* **134**:91–105.

Checketts JX, Scott JT, Meyer C, *et al.* (2018) The robustness of trials that guide evidence-based orthopaedic surgery. *J Bone Joint Surg Am* **100**:e85.

DeVito NJ, Goldacre B. (2019) Catalogue of bias: Publication bias. *BMJ Evidence-Based Medicine*. **24**:53–54. Doi: 10.1136/bmjebm-2018-111107.

Ebbert JO, Hughes JR, West RJ, *et al.* (2015) Effect of varenicline on smoking cessation through smoking reduction: A randomized clinical trial. *JAMA* **313**:687–94.

Edwards E, Wayant C, Besas J, *et al.* (2018) How fragile are clinical trial outcomes that support the CHEST clinical practice guidelines for VTE? *Chest*.

Evans SR. (2010) Fundamentals of clinical trial design. *Journal of Experimental Stroke & Translational Medicine*. **3**:19–27. Doi: 10.4172/1939-067x.1000116.

Friedman LM, Furberg CD, DL DM, *et al.* (2015) *Fundamentals of Clinical Trials*. Cham: Springer.

Goerke K, Parke M, Horn J, *et al.* (2020) Are results from randomized trials in anesthesiology robust or fragile? An analysis using the fragility index. *Int J Evid Based Healthc* **18**:116–24.

Gowing L, Farrell MF, Ali R, White JM. (2014) Alpha2-adrenergic agonists for the management of opioid withdrawal. *Cochrane Database of Systematic Reviews*. **31**:CD002024. Doi: 10.1002/14651858.cd002024.pub4.

Head ML, Holman L, Lanfear R, *et al.* (2015) The extent and consequences of p-hacking in science. *PLoS Biol* **13**:e1002106.

Hedden SL, Woolson RF, Malcolm RJ. (2006) Randomization in substance abuse clinical trials. *Subst Abuse Treat Prev Policy* **1**:6.

Higgins JPT, Thomas J, Chandler J, *et al.* (2019) *Cochrane Handbook for Systematic Reviews of Interventions*, 2nd edn. Chichester UK: John Wiley & Sons.

Hoogendam A, de Vries Robbé PF, Stalenhoef AFH, *et al.* (2009) Evaluation of PubMed filters used for evidence-based searching: Validation using relative recall. *J Med Libr Assoc* **97**:186–93.

Hutton JL, Williamson PR. (2000) Bias in meta-analysis due to outcome variable selection within studies. *J R Stat Soc Ser C Appl Stat* **49**:359–70.

Jhanjee S, Jain R, Jain V, *et al.* (2015) Evaluating the effects of Varenicline on craving, withdrawal, and affect in a randomized, double-blind, placebo-controlled clinical trial of Varenicline for smokeless tobacco dependence in India. *J Psychoactive Drugs* **47**:325–30.

Koegelenberg CF, Noor F, Bateman ED. (2014) Efficacy of Varenicline combined with nicotine replacement therapy vs Varenicline alone for smoking cessation: A randomized clinical trial. *Journal of Vascular Surgery*. Doi: 10.1016/j.jvs.2014.08.086.

Meyer C, Bowers A, Jaiswal D, *et al.* (2019) An analysis of randomized controlled trials underpinning ST-elevation myocardial infarction management guidelines. *Am J Emerg Med* **37**:2229–38.

Meyer C, Bowers A, Tritz D, *et al.* (2020) The fragility of randomized trial outcomes underlying management of dyspepsia and helicobacter pylori infections. *Int J Evid Based Healthc* **18**:125–37.

Minozzi S, Amato L, Vecchi S, *et al.* (2010) Anticonvulsants for alcohol withdrawal. *Cochrane Database Syst Rev* CD005064.

Minozzi S, Cinquini M, Amato L, *et al.* (2015) Anticonvulsants for cocaine dependence. *Cochrane Database Syst Rev* CD006754.

Moher D, Liberati A, Tetzlaff J, *et al.* (2009) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann Intern Med* **151**:264W64–9 W64.

Moher D, Pham B, Jones A, *et al.* (1998) Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* **352**:609–13.

Murad MH, Wang Z. (2017) Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med* **22**:139–42.

National Institutes of Health. (2019) NIH Definition of Clinical Trial Case Studies. Https://grants.nih.gov/policy/clinical-trials/CT-Definition-Case-Studies_1.7.19.pdf (10 April 2019, date accessed).

National Research Council, Institute of Medicine, Board on Neuroscience and Behavioral Health *et al.* (2004) *New Treatments for Addiction: Behavioral, Ethical, Legal, and Social Questions*. Washington, DC: National Academies Press.

Nunan D, Aronson J, Bankhead C. (2018) Catalogue of bias: Attrition bias. *BMJ Evid Based Med* **23**:21–2.

Ouzzani M, Hammady H, Fedorowicz Z, *et al.* (2016) Rayyan—A web and mobile app for systematic reviews. *BioMed Central* **5**:210.

Pani PP, Trogu E, Pacini M, *et al.* (2014) Anticonvulsants for alcohol dependence. *Cochrane Database Syst Rev* CD006754.

Rösner S, Hackl-Herrwerth A, Leucht S, *et al.* (2010) Acamprosate for alcohol dependence. *Cochrane Database Syst Rev* CD004332.

Schulz KF, Chalmers I, Hayes RJ, *et al.* (1995) Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* **273**:408–12.

Skinner M, Tritz D, Farahani C, *et al.* (2019) The fragility of statistically significant results in otolaryngology randomized trials. *Am J Otolaryngol* **40**:61–6.

Sterne JAC, Savović J, Page MJ, *et al.* (2019) RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **28**;366:l4898.

Turner EH, Matthews AM, Linardatos E, *et al.* (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* **358**:252–60.

Vassar M, Jellison S, Wendelbo H, *et al.* (2020) Data sharing practices in randomized trials of addiction interventions. *Addict Behav* **102**:106193.

Vassar M, Jellison S, Wendelbo H, *et al.* (2019a) Using the CONSORT statement to evaluate the completeness of reporting of addiction randomised trials: A cross-sectional review. *BMJ Open* **9**:e032024.

Vassar M, Roberts W, Cooper CM, *et al.* (2019b) Evaluation of selective outcome reporting and trial registration practices among addiction clinical trials. *Addiction* **115**:1172–79.

Wayant C, Meyer C, Gupton R, *et al.* (2019) The fragility index in a cohort of HIV/AIDS randomized controlled trials. *J Gen Intern Med* **34**:1236–43.

Wayant C, Scheckel C, Hicks C, *et al.* (2017) Evidence of selective reporting bias in hematology journals: A systematic review. *PLoS One* **12**: e0178379.

Wayant C, Scott J, Vassar M. (2018) Evaluation of lowering the P value threshold for statistical significance from .05 to .005 in previously published randomized clinical trials in major medical journals. *JAMA* **320**:1813–5.