Discrepancies from registered protocols and spin occurred frequently in randomized psychotherapy trials

Stoll, Marlene; Mancini, Alexander; Hubenschmid, Lara; Dreimüller, Nadine; König, Jochem; Cuijpers, Pim; Barth, Jürgen; Lieb, Klaus

published in
Journal of clinical epidemiology
2020

DOI (link to publisher)
10.1016/j.jclinepi.2020.08.013

document version
Publisher's PDF, also known as Version of record

Journal of Clinical Epidemiology 128 (2020) 49−56

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Discrepancies from registered protocols and spin occurred frequently in randomized psychotherapy trials—A meta-epidemiologic study

Marlene Stoll[a,b], Alexander Mancini[a], Lara Hubenschmid[b], Nadine Dreimüller[a], Jochem König[c], Pim Cuijpers[d], Jürgen Barth[e,1,*], Klaus Lieb[a,b,1]

[a]*Department of Psychiatry and Psychotherapy, University Medical Center Mainz, Untere Zahlbacher Straße 8, 55131 Mainz, Germany*
[b]*Leibniz Institute for Resilience Research, Wallstraße 7, 55122 Mainz, Germany*
[c]*Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Mainz, Obere Zahlbacher Straße 69, 55131 Mainz, Germany*
[d]*Department of Clinical, Neuro, and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorstraat 7, 1081 BT Amsterdam, the Netherlands*
[e]*Institute for Complementary and Integrative Medicine, University Hospital Zurich and University Zurich, Sonneggstraße 6, 8091 Zurich, Switzerland*

Accepted 18 August 2020; Published online 20 August 2020

Abstract

Objectives: This study aimed to investigate the relationship between trial registration, trial discrepancy from registered protocol, and spin in nonpharmacological trials.

Study Design and Setting: Recent psychotherapy trials on depression (2015−2018) were analyzed regarding their registration status and its relationship to discrepancies between registered and published primary outcomes and to spin (discrepancy between the nonsignificant finding in a study and an overly beneficial interpretation of the effect of the treatment).

Results: A total of 196 trials were identified, of which 78 (40%) had been registered prospectively and 56 (29%) had been registered retrospectively. In 102 (76%) of 134 registered trials, discrepancies between trial and protocol were present. Of 72 trials with a nonsignificant difference between treatments for the primary outcome, 68 trials (94%) showed spin. Discrepancies from protocol were less frequent in prospectively than in retrospectively registered trials (odds ratio= 0.19; 95% confidence interval [CI]: 0.07−0.52), but regarding the amount of spin, there was no difference between prospectively and retrospectively registered trials (r_b = −0.12; 95% CI: −0.41 to 0.19) or between registered and unregistered trials (r_b = −0.22, 95% CI −0.49 to 0.08).

Conclusion: Protocol discrepancies and spin have a high prevalence in psychotherapy outcome research. The results show no relation between registration and spin, but prospective registration may prevent discrepancies from protocol.

Keywords: Psychotherapy; Depression; Reporting bias; Spin in research; Conflict of interest; Review

Funding: This work was funded by the Volkswagen Foundation, Germany (grant no. A118085, ref. 91498 to K.L.). The Foundation did not take role in the study design, data collection, analysis or interpretation, or in writing of the report of the submitted article. All researchers are independent from funders and all authors had full access to the data.

Declarations of interest: All authors have completed the International Committee of Medical Journal Editors uniform disclosure form at www.icmje.org/coi_disclosure.pdf. M.S., A.M., and L.H. declared that they had received salary from the Volkswagen Foundation. K.L. declared that he received a research grant by Volkswagen Foundation. However, the funder had no influence on the study findings and interpretation. J.B., P.C., N.D., and J.K. have nothing to disclose. All authors declared no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced submitted work.

Data statement: All data are available in the online appendix (eTables 1−3).

Authors' contributions: M.S. contributed to conceptualization, methodology, investigation, writing, reviewing, and editing the article, visualization, and project administration. A.M. contributed to methodology, investigation, review and editing the article, and formal analysis. L.H. contributed to investigation, reviewing and editing the article, and project administration. N.D. contributed to reviewing and editing the article. J.K. contributed to formal analysis and reviewing and editing the article. P.C. contributed to conceptualization, resources, data curation, and reviewing and editing the article. J.B. contributed to conceptualization, methodology, validation, reviewing and editing the article, and supervision. K.L. contributed to conceptualization, methodology, writing, reviewing and editing the article, supervision, and funding acquisition.

[1] These authors contributed equally.

* Corresponding author. Tel.: +41763244698; fax: +41442554394.
E-mail address: mail@juergen-barth.de (J. Barth).

https://doi.org/10.1016/j.jclinepi.2020.08.013
0895-4356/

What is new?

Key findings

- Discrepancies from the registered protocol manifest in 76% of psychotherapy trials, and nearly all trials with nonsignificant effects between treatments show some form of spin.

What does this add to what was known?

- Protocol discrepancies are less frequent in psychotherapy trials which are registered prospectively compared with retrospectively registered trials.

What is the implication and what should change now?

- Spin is not prevented by registration of trials.
- Policies such as reporting guidelines should be promoted by relevant stakeholders.

1. Introduction

Since 2013, the World Medical Association declaration of Helsinki states that "every research study involving human subjects must be registered in a publicly accessible database before recruitment of the first subject" [1]. In medicine as well as in other disciplines such as psychology, the (pre-) registration of clinical trials or studies in public registries is seen as an effective tool to improve conduct, analysis, and interpretation of studies and to increase the reliability and consequently credibility of research [2−5]. To publish a study protocol in public registries includes the registration of study characteristics such as study design, main outcomes, and main analyses [6]. Registration of studies thereby is intended to prevent selective publication and selective reporting of outcomes [3,7], to prevent unnecessary double research effort, and to give patients, the public, and other stakeholders such as ethical review boards an overview of ongoing or planned trials [7]. In the medical sciences, up to 90% of clinical trials are registered, but only 60−77% are registered prospectively, that is, before enrolling the first participant [8,9]. These numbers are based mainly on pharmacological trials published in journals that endorse the International Committee of Medical Journal Editors' guidelines [7]. Registration of nonpharmacological trials has not been well investigated yet. One study of psychotherapy trials found that 60% of the investigated trials had been registered, 24% of them prospectively [10]. More evidence is available for health intervention trials, which include pharmacological as well as nonpharmacological trials with about half of them being registered [11−15].

If there are systematic discrepancies between the information given in published trials and their respective registered protocols, this is called outcome reporting bias. Outcome reporting bias is frequently found (28−62%) in pharmacological as well as nonpharmacological trials [14,16−19]. Another form of bias in the reporting of results can occur if the publications themselves show an overly beneficial interpretation of the reported effect of the treatment [20]. This interpretation bias is called "spin" and is also highly prevalent in the biomedical literature [20,21]. It has been estimated that 56% of the abstracts published in psychiatry and psychology journals contain spin [22]. Different biases can accumulate and interact: Trials with positive outcomes are more likely to be published (publication bias) and significant outcomes are more likely to be included in a published trial, whereas negative outcomes are changed or omitted (another form of outcome reporting bias). In case negative outcomes are reported, they are sometimes reported in an overly beneficial way (spin). These biases may arise in conjunction and thereby leave nonsignificant results out of eyeshot, which creates a risk of a distorted image of the actual evidence in the published literature [23].

The field of psychotherapy research is of particular interest regarding reporting biases: Contrary to most medical trials, a typical psychotherapy trial is conducted by a researcher with clinical expertise who works as a therapist and whose school of thought is exceptionally shaped by a long education in this therapy [24]. Although in medical research the pharmaceutical industry as an external factor may play a relevant role in the conduction of trials, industry is less involved in the conduction of psychotherapy trials. Therefore, psychotherapy trials are more dependent on the individual researcher, and the researcher's personal interests in the outcome of the trial might play a more important role. These interests are discussed in terms of researcher allegiance. Evidence shows that researchers with higher researcher allegiance often published studies with larger effects [25].

To the best of our knowledge, no studies have investigated the registration of psychotherapy randomized trials and its relationship to both protocol discrepancies and spin. We, therefore, investigated in the present study registration status and registration time point, discrepancies between trials and protocols, and spin in a larger number of psychotherapy trials. The objectives of the present study were (1) to investigate the extent to which recent psychotherapy trials on depression are prospectively or retrospectively registered; (2) to investigate the respective prevalence of protocol discrepancies and spin; and (3) to examine the relationship of registration status and registration time point to protocol discrepancies and spin. Furthermore, the relationship between protocol discrepancies and trial effectiveness was explored to provide preliminary evidence for the prevalence of outcome reporting bias in trials with nonpharmacological interventions.

2. Methods

2.1. Selection of trials

Trials were retrieved from a collection of psychotherapy trials on depression provided by Cuijpers et al. [26]. We

focused on one particular disorder to minimize variation and decided upon depression because it is a highly prevalent disorder [27] with a high level of disease burden [28], leading to an ongoing development of treatments, and therefore a large evidence base for research on treatments is available [29]. In the trial collection by Cuijpers et al. [26], studies were eligible to be included if they investigated the treatment of a depressive disorder or an elevated level of depressive symptomatology and at least one treatment arm is psychological and for adults. The systematic literature search is updated every year and was conducted in the databases PubMed, Embase, PsycINFO, and Cochrane Register of Controlled Trials up to January 1, 2019, with no language restrictions. For more details regarding the database, refer to studies by Cuijpers [29], Cuijpers et al. [30], and Cuijpers [31]. Specific inclusion criterion for the present study was that randomized trials of the described collection were published between January 1, 2015, and December 31, 2018. Studies published before 2015 were excluded to get a current picture that is not interfered by earlier standards.

2.2. Data extraction

For each trial, two independent reviewers (A.M. and M.S.) extracted information and conducted assessments for protocol discrepancies and spin. Disagreements were resolved by discussion and consultation with a third investigator (J.B.), if necessary. Reviewers extracted the following items from the trial protocol and/or the published article: registration number and time point of registration, primary outcomes, and statistical significance on the published primary outcome. If available, the definition of the primary outcome was extracted, including measurement scale, time point, and time frame (i.e., involving ''baseline'' time point). To find the respective information, texts were searched manually.

2.3. Assessing trial registration status

To assess trial registration status, the trial was screened to identify a registration number. If none was found, we searched online registries in the following order by using the surname of the first author, the name of the treatment, and mental health condition as search parameters: ClinicalTrials.gov, www.who.int/trialsearch, www.isrctn.com, and author's local registry. If a registration number and the respective registration were found, the trial was rated as *registered*. Trial registration search was conducted in October 2019.

2.4. Assessing data in the registrations

Trials were considered *prospectively registered* if the registration date preceded participant enrollment date or if the trial was registered within 1 month of participant enrollment (e.g., participant enrollment date: May 1; registration date: May 15). Trials were considered *retrospectively registered* if the trial was registered more than 1 month after participant enrollment had begun or if registration within 1 month was unclear (e.g., participant enrollment date: May; registration date: June) or if participant enrollment began in the same year the trial was registered and no specific month was specified.

Registrations were screened for the definition (method of measurement, time point, or time frame) of the primary outcome. If the primary outcome definition had been changed, we extracted information of the ''original'' rather than a changed version of the primary outcome definition. The registered primary outcome was deemed to be defined *exactly* if one scale and one time point were defined and *inexactly* if none or more than one scale or none or more than one time point or a time frame was defined.

2.5. Assessing data in the publications

We screened publications for the terms ''primary outcome(s),'' ''primary endpoint(s),'' ''main outcome(s),'' or ''main endpoint(s).'' If mentioned, the respective definition (i.e., method, time point, or time frame of measurement) was extracted. For this purpose, we screened the paragraph where the term was mentioned and extracted the information regarding measurement and time point that was semantically the closest. If more than one primary outcome was defined, we extracted all mentioned primary outcomes.

Next, we assessed the exactness of the definition for every mentioned primary outcome (as previously mentioned) and its statistical significance. A primary outcome was regarded as *significant* if it was exactly defined and the statistical test reached $P < 0.05$ (if not otherwise specified) and as *nonsignificant* if this test did not reach statistical significance. In case of an inexactly defined primary outcome, we extracted the statistical significance of this primary outcome at posttest in favor of the intervention group.

Afterward, we assessed the effectiveness of trials. A trial was classified as *effective*, if the reported primary outcome was statistically significant or, in case there was more than one primary outcomes per trial, all of them were statistically significant. It was classified as *not effective*, if at least one primary outcome was statistically nonsignificant.

2.6. Analysis of protocol discrepancies

We analyzed protocol discrepancies in all trials for which a registration could be identified by, first, comparing the respective registered and published primary outcomes. They were classified as *discrepant* if their definitions differed (e.g., different methods of measurement) or if the amount of information differed (e.g., a time point was registered but not reported). They were classified as *concordant* if the registered and reported primary outcomes matched exactly. This was also assumed if the time point of

the reported primary outcome was within the time frame of the registered primary outcome.

In a second step, we assessed protocol discrepancy per trial, which was considered present if there was a discrepancy between the registered and the reported primary outcome, and if there was more than one reported primary outcome, protocol discrepancy per trial was considered present if there was at least one discrepancy between registered and reported primary outcomes in the trial. Protocol discrepancy per trial was considered nonpresent if the primary outcome or, in case of more than one primary outcome, all of them were reported exactly as they had been registered. Notably, this approach to assess protocol discrepancies differs from assessments of outcome reporting bias [17].

2.7. Assessing spin

To assess spin, we examined all trials with at least one nonsignificant primary outcome. We adapted the coding manual by Gewandter et al. [32], which they had developed based on the study by Boutron et al. [20]. Seven forms of spin were investigated: (1) selective reporting (the nonsignificant primary outcome is not mentioned in the screened section), (2) distracting with secondary analyses (the primary outcome is not mentioned but significant secondary analyses are), (3) distracting with within-group differences (the primary outcome is not mentioned but significant within-group differences are), (4) focus on significant secondary analyses ([a] secondary analyses are mentioned before the primary outcome; [b] effect sizes are mentioned instead of primary outcome effect sizes; [c] effect is depicted in figures but primary outcome is not), (5) focus on significant within-group differences over time (a to c, as mentioned earlier), (6) interpreting nonsignificant primary results as showing treatment equivalence in a superiority trial, and (7) claiming or emphasizing the beneficial effect of the treatment despite a nonsignificant outcome. Spin forms were investigated in five sections of the publication: abstract results and conclusions, main text results, discussions, and conclusions (Table 2). In addition to the assessment of the single spin forms in the different publication sections, we calculated the amount of spin per trial by summing up the occurrence of spin forms in that trial. To estimate interrater reliability of the spin scores per trial, we calculated the intraclass correlation coefficient, which was 0.81, based on a mean-rating ($k = 2$), absolute-agreement, two-way random effects model. If two or more primary outcomes per trial were nonsignificant, we assessed spin separately per primary outcome, and the amount of spin per trial was generated by using the average score.

2.8. Statistical analyses

In addition to the descriptive analysis of the sample characteristics, we aimed to investigate the relationship between registration time point and protocol discrepancy, registration status and amount of spin, and registration time point and amount of spin. Furthermore, we investigated the relationship between registration status and effectiveness, registration time point and effectiveness, and protocol discrepancy and effectiveness. For quantification of the relationship between two binary characteristics, we calculated odds ratios (ORs) and corresponding 95% confidence intervals (CIs). For quantification of the relationship between a binary characteristic and the amount of spin, we calculated Mann–Whitney U tests and provide rank biserial correlations (r_b) with 95% CI. Statistics of the distribution of the amount of spin are reported as median and quartiles (interquartile range [IQR]).

3. Results

3.1. Selection of eligible RCTs

A total of 204 trials of the database matched the inclusion criteria. Eight trials had to be excluded because they were duplicates in the database. Finally, we extracted and analyzed 196 trials (characteristics of each trial are shown in eTable1).

3.2. Registration status and registered and reported primary outcomes

Of 196 trials, 134 (68%) had been registered in a clinical trial registry. Of those, 78 (58%) had been registered prospectively and 56 (42%) had been registered retrospectively (Table 1A). In the 134 protocols, 197 primary outcomes were registered, which are, on average, 1.47 registered primary outcomes per protocol. Only 26% of them were exactly defined.

In the 196 published trials, 194 primary outcomes were reported, of which 89 (46%) were statistically significant (Table 1B). In 46 (23%) of 196 published trials, no primary outcome was defined. Overall, we could assess effectiveness in 144 of the 196 trials, and classified 67 (47%) as effective and 77 (53%) as not effective.

3.3. Prevalence of protocol discrepancies

At least one discrepancy between protocol and trial was present in 102 (76%) of 134 registered trials. Prevalence of protocol discrepancies were lower in prospectively (51/78, 65%) than in retrospectively (51/56, 91%) registered trials. Odds of protocol discrepancies were significantly reduced with prospective registration ($OR = 0.19$; 95% CI: 0.07–0.52).

Registered and reported primary outcomes and rating of protocol discrepancies in all 196 included trials are shown in eTable 2.

3.4. Prevalence of spin

Seventy-two trials had at least one nonsignificant primary outcome and were assessed for spin. Spin forms

Table 1. Characteristics of registrations and published trials: registration time point and characteristics of registered primary outcomes, trial registration status, and characteristics of trial and reported primary outcomes

Characteristics	*n/N* (%)
(A) Registration report of 134 registered trials	
Time point of registration	
Prospective	78/134 (58)
Retrospective	56/134 (42)
Registered POs	
At least one PO mentioned	134/134 (100)
No PO mentioned	0/134 (0)
Number of registered PO	197
Quality of registered PO definition	
Exactly defined	51/197 (26)
Inexactly defined	146/197 (74)
(B) Trial report of 196 published randomized trials	
Registration status	
Registered	134/196 (68)
Not registered	62/196 (32)
Trial effectiveness	
Effective	67/196 (34)
Not effective	77/196 (39)
Not rateable[a]	52/196 (27)
Reported POs	
At least one PO mentioned	150/196 (77)
No PO mentioned	46/196 (23)
Number of reported POs	194
Quality of reported PO definition	
Exactly defined	59/194 (30)
Inexactly defined	135/194 (70)
Significance of PO	
Significant	89/194 (46)
Nonsignificant	94/194 (48)
Not rateable[a]	11/194 (6)

Abbreviation: PO, primary outcome.
[a] In some cases, we were not able to identify a primary outcome or a similar equivalent to assess statistical significance.

and amount of spin for each of those 72 trials are listed in eTables 3a and 3b. Sixty-eight trials (94%) showed at least one form of spin (median amount of spin per trial was 5.75, *IQR:* 3–8). As shown in Table 2, the most frequently used forms of spin were that the nonsignificant PO was not mentioned in the abstract conclusion section (selective reporting in 37/69 trials, 54%) and that the beneficial effect of the treatment was claimed in the abstract (30/69 trials, 43%) or the main text (22/48 trials, 46%) conclusion section. The text section with the highest prevalence of spin ratings was the abstract conclusion section, in which 56 of 69 (81%) investigated trials showed some form of spin, and the main text discussion section, in which 58 of 72 (81%) investigated trials showed some form of spin.

We could assess spin in 28 prospectively registered trials, 25 retrospectively registered trials, and 19 unregistered trials. The percentages of trials that showed spin were 89%, 96%, and 100%, respectively. Amount of spin per trial did not differ significantly between unregistered (*median* = 7, *IQR:* 4–8) and registered trials (*median* = 5, *IQR:* 3–8; $r_b = -0.22$; 95% CI: −0.49 to 0.08) nor between retrospectively (*median* = 6, *IQR* 3–8) and prospectively registered trials (*median* = 5; *IQR:* 2.63–8; $r_b = -0.12$; 95% CI: −0.41 to 0.19).

In 4% of investigated trials (3/72), spin was found in only one section of the trial publication, in 10% (7/72) in two, in 28% (20/72) in three, in 31% (22/72) in four, and in 22% (16/72) in five sections of the trial publication.

3.5. Effectiveness of trials

Trials reported effective interventions in 67 (47%) of 144 trials that unambiguously reported about effectiveness. The prevalence rates of effective trials are registered vs. unregistered trials (53/111, 48%, vs. 14/33, 42%), prospectively vs. retrospectively registered trials (36/66, 55%, vs. 17/45, 38%), and trials with vs. trials without protocol discrepancies (36/80, 45%, vs. 17/31, 55%).

The odds to report about an effective treatment did not significantly increase with registered compared with unregistered trials (*OR* = 1.24, 95% CI: 0.57–2.72) or with prospective compared with retrospective registration (*OR* = 1.98, 95% CI: 0.91–4.28). The odds to report about an effective treatment did not significantly decrease with trials that showed protocol discrepancies compared with trials without protocol discrepancies (*OR* = 0.67, 95% CI: 0.29–1.55).

4. Discussion

4.1. Principal findings

Of 196 trials, 40% had been registered prospectively and 29% had been registered retrospectively. Protocol discrepancies were present in 76% of registered trials. Nonsignificant primary outcomes were interpreted with some form of spin in 94% of all trials. We found no differences in protocol discrepancies and spin between registered and unregistered trials, but protocol discrepancies were less likely if trials were registered prospectively.

4.2. Findings in context

Our finding of 68% registered trials is in the range of previous research on clinical trial registration of health interventions [10,14,15], but prevalence of registration appears lower than in comparable studies on pharmacological trials where clinical trial registration is an established standard [8,9]. To compare our findings regarding outcome reporting bias with other studies is

Table 2. Forms of spin in 72 trials showing at least one nonsignificant primary outcome

Spin	***N* (%)**	
Spin in results section	**Abstract (n = 70)**	**Main text (n = 72)**
Some type of spin	51 (73)	44 (61)
Selective reporting	4 (6)	0 (0)
Distracting with secondary analyses	15 (21)	4 (6)
Distracting with within-group differences	11 (16)	4 (6)
Focus on secondary analyses		
Secondary analyses mentioned first	15 (21)	19 (26)
Effect estimates mentioned for secondary analyses only	12 (17)	NA
Only secondary analyses are presented in figures	NA	7 (10)
Focus on within-group differences		
Within-group differences mentioned first	15 (21)	29 (40)
Effect size mentioned for within-group differences only	9 (13)	NA
Only within-group differences are presented in figures	NA	1 (1)
Spin in discussion section		Main text (n = 72)
Some type of spin	NA	58 (81)
Selective reporting	NA	4 (6)
Distracting with secondary analyses	NA	15 (21)
Distracting with within-group differences	NA	14 (19)
Focus on secondary analyses		
Secondary analyses mentioned first	NA	22 (31)
Focus on within-group differences		
Within-group differences mentioned first	NA	24 (33)
Interpreting nonsignificant primary results in a superiority trial as showing treatment equivalence	NA	10 (14)
Claiming or emphasizing beneficial effect of treatment	NA	28 (39)
Spin in conclusion section	Abstract (n = 69)	Main text (n = 48)
Some type of spin	56 (81)	36 (75)
Selective reporting	37 (54)	15 (31)
Distracting with secondary outcomes	8 (12)	11 (23)
Distracting with within-group differences	7 (10)	7 (15)
Focus on secondary analyses		
Secondary analyses mentioned before	4 (6)	5 (10)
Focus on within-group differences		
Within-group differences mentioned before	3 (4)	1 (2)
Interpreting nonsignificant primary results in a superiority trial as showing treatment equivalence	4 (6)	3 (6)
Claiming or emphasizing beneficial effect of treatment	30 (43)	22 (46)

Abbreviation: NA, not applicable.

difficult because the definition of outcome reporting bias often differs between studies. Our approach was different since we measured discrepancies between registered and published primary outcomes in a first step and then related protocol discrepancy to trial effectiveness in a second step, whereas other studies assessed outcome reporting bias as discrepancy that favors the published primary outcomes [14,16,17]. Our findings support the conclusion of other researchers that registrations of pharmacological and nonpharmacological trials are of low quality (e.g., high prevalence of inexactly defined primary outcomes) and that registered and published information often is discordant [10,11,18,19]. Regarding spin, our results are in line with previous research in that most spin was found in the conclusion sections of the investigated trials [20,22,33]. More detailed comparisons are difficult because of different definitions of spin. Further studies should find a standardized spin measurement method and make direct comparisons, for example, between pharmacological and nonpharmacological trials. The reasons for the high prevalence of reporting biases such as spin are still unclear. Chiu et al. [21] showed that funding source is one of the most frequently investigated factors associated with spin, but they did not find a significant association between industry sponsorship

and spin. It might be speculated, especially for psychotherapy trials, that other factors such as researcher allegiance [24,25,34,35] or inappropriate incentives fueled through the academia reward system may contribute to the high prevalence of bias [36,37].

4.3. Strengths and limitations

A strength of the present study is that registration status, protocol discrepancies, spin, and their interactions were investigated for the first time in a large sample of nonpharmacological outcome studies. A further strength was that we investigated spin in a very detailed way, with an extension of the scale developed by Gewandter et al. [32]. Limitations are, first, protocol discrepancies obviously could only be investigated in registered trials. We, therefore, do not know whether registration per se is effective to reduce the risk of outcome reporting bias. Second, our analysis of protocol discrepancies and spin was in general only possible if outcomes were (adequately) reported, which often was not the case.

5. Conclusions and policy implications

The high prevalence of protocol discrepancies and spin and the fact that only the risk of protocol discrepancies, but not of spin, was reduced by prospective registration, suggest that more effective ways than mere registration of studies are needed to increase the trustworthiness of psychotherapy outcome research.

First, better adherence to standards is needed. Many journals that publish nonpharmacological trials do not have policies that require prospective registration. The rate of prospectively registered trials published in these journals is much lower than in journals that endorse such guidelines [12,38]. The most well-known reporting guideline is the Consolidated Standards of Reporting Trials 2010 Statement that requires that primary and secondary outcome measures are completely defined, including "how and when they were assessed," that trial's registration number and name of trial registry are reported and that any changes of outcomes after the trial commenced are mentioned, "with reasons" [39]. We especially encourage journals publishing nonpharmacological trials to implement these guidelines. Reporting guidelines should also be promoted by graduate schools and writing courses, and they should receive appropriate attention by scientific boards of academic associations. Second, other potential sources of bias despite funding, for example, researcher allegiance [34,36], have to be better identified and need transparency. Third, the publication format of registered reports, which is adopted by an increasing number of journals [37,40], should be promoted. In such registered reports, study protocols are submitted to a journal before any data are gathered or analyzed. Study protocols then undergo a peer-review process and, with acceptance, the publication of the trials' results after data collection, analysis, and interpretation is guaranteed independent from the finding [41,42]. To enhance acceptance among researchers, however, a fast reviewing process will be of high importance.

In conclusion, this study shows a high prevalence of protocol discrepancies and spin and low rates of registered trials in psychotherapy research. Prospective registration in this sample was associated with less protocol discrepancies than retrospective registration.

Acknowledgments

The authors would like to thank Marc Himmelmann, Janosch Weisskircher, and Cora Koch for their help and input in the early stages of the project. The authors also thank Ioana Cristea and Whitney Yoder for their feedback on our research project and support regarding the database.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.08.013.

References

[1] World Medical Association. WMA declaration of Helsinki – ethical principles for medical research involving human subjects. Available at https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/. Accessed August 10, 2020.
[2] Allen C, Mehler DMA. Open science challenges, benefits and tips in early career and beyond. PLoS Biol 2019;17(5):e3000246.
[3] Dal-Ré R, Bobes J, Cuijpers P. Why prudence is needed when interpreting articles reporting clinical trial results in mental health. Trials 2017;18(1):1–4.
[4] DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International committee of Medical Journal Editors. Arch Dermatol 2005;141:76–7.
[5] Ioannidis JPA. Why most published research findings are false. PLoS Med 2005;2(8):e124.
[6] World Health O. International standards for clinical trial registries: the registration of all interventional trials is a scientific, ethical and moral responsibility. version 3.0 ed. Geneva: World Health Organization; 2018:48.
[7] ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. Available at http://www.icmje.org/icmje-recommendations.pdf. Accessed August 10, 2020.
[8] Gopal AD, Wallach JD, Aminawung JA, Gonsalves G, Dal-Ré R, Miller JE, et al. Adherence to the International Committee of Medical Journal Editors' (ICMJE) prospective registration policy and implications for outcome integrity: a cross-sectional analysis of trials published in high-impact specialty society journals. Trials 2018;19(1):448.
[9] Huser V, Cimino JJ. Evaluating adherence to the International Committee of Medical Journal Editors' policy of mandatory, timely clinical trial registration. J Am Med Inform Assoc 2013;20:e169–74. https://doi.org/10.1136/amiajnl-2012-001501.

[10] Bradley HA, Rucklidge JJ, Mulder RT. A systematic review of trial registration and selective outcome reporting in psychotherapy randomized controlled trials. Acta Psychiatr Scand 2017;135(1):65−77.
[11] Azar M, Riehm KE, McKay D, Thombs BD. Transparency of outcome reporting and trial registration of randomized controlled trials published in the journal of consulting and clinical psychology. PLoS One 2015;10:e0142894.
[12] Azar M, Riehm KE, Saadat N, Sanchez T, Chiovitti M, Qi L, et al. Evaluation of journal registration policies and prospective registration of randomized clinical trials of nonregulated health care interventions. JAMA Intern Med 2019;179(5):624−32.
[13] Cybulski L, Mayo-Wilson E, Grant S. Improving transparency and reproducibility through registration: the status of intervention trials published in clinical psychology journals. J Consult Clin Psychol 2016;84(9):753−67.
[14] Scott A, Rucklidge JJ, Mulder RT. Is mandatory prospective trial registration working to prevent publication of unregistered trials and selective outcome reporting? An observational study of five psychiatry journals that mandate prospective clinical trial registration. PLoS One 2015;10:e0133718.
[15] Shinohara K, Tajika A, Imai H, Takeshima N, Hayasaka Y, Furukawa TA. Protocol registration and selective outcome reporting in recent psychiatry trials: new antidepressants and cognitive behavioural therapies. Acta Psychiatr Scand 2015;132(6):489−98.
[16] Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst Rev 2011;1:Mr000031.
[17] Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One 2013;8:e66844.
[18] Jones CW, Misemer BS, Platts-Mills TF, Ahn R, Woodbridge A, Abraham A, et al. Primary outcome switching among drug trials with and without principal investigator financial ties to industry: a cross-sectional study. BMJ Open 2018;8(2):e019831.
[19] Shepshelovich D, Yahav D, Tibau A, Amir E. Assessment of frequency and reporting of design changes among clinical drug trials published in influential medical journals. Eur J Intern Med 2020;71:45−9.
[20] Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA 2010;303:2058−64.
[21] Chiu K, Grundy Q, Bero L. 'Spin' in published biomedical literature: a methodological systematic review. PLoS Biol 2017;15(9):e2002173.
[22] Jellison S, Roberts W, Bowers A, Combs T, Beaman J, Wayant C, et al. Evaluation of spin in abstracts of papers in psychiatry and psychology journals. BMJ Evidence Based Med 2019. https://doi.org/10.1136/bmjebm-2019-111176.
[23] de Vries YA, Roest AM, de Jonge P, Cuijpers P, Munafò MR, Bastiaansen JA. The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression. Psychol Med 2018;48(15):2453−5.
[24] Leykin Y, DeRubeis RJ. Allegiance in psychotherapy outcome research: separating association from bias. Clin Psychol Sci Pract 2009;16(1):54−65.
[25] Munder T, Brutsch O, Leonhart R, Gerger H, Barth J. Researcher allegiance in psychotherapy outcome research: an overview of reviews. Clin Psychol Rev 2013;33(4):501−11.
[26] Cuijpers P, Karyotaki E, Ciharova M, Quero S, Pineda B, Munoz R, et al. A meta-analytic database of randomised trials on psychotherapies for depression. Open Science Foundation; 2020. Available at https://osf.io/825c6/. Accessed August 10, 2020.
[27] Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990−2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388(10053):1545−602.
[28] Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006;3(11):e442.
[29] Cuijpers P. Four decades of outcome research on psychotherapies for adult depression: an overview of a series of meta-analyses. Can Psychol Psychol Canad 2017;58(1):7−19.
[30] Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry 2008;165(10):1272−80.
[31] Cuijpers P. Meta-analyses in mental health research: a practical guide 2016. Available at https://indd.adobe.com/view/5fc8f9a0-bf1e-49d3-bf5f-a40bfe5409e0. Accessed August 10, 2020.
[32] Gewandter JS, McKeown A, McDermott MP, Dworkin JD, Smith SM, Gross RA, et al. Data interpretation in analgesic clinical trials with statistically nonsignificant primary analyses: an ACTTION systematic review. J Pain 2015;16(1):3−10.
[33] Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. Proc Natl Acad Sci U S A 2018;115(11):2613−9.
[34] Cristea IA, Naudet F. Increase value and reduce waste in research on psychological therapies. Behav Res Ther 2019;123:103479.
[35] Munder T, Fluckiger C, Gerger H, Wampold BE, Barth J. Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? a meta-analysis. J Couns Psychol 2012;59(4):631−7.
[36] Lieb K, von der Osten-Sacken J, Stoffers-Winterling J, Reiss N, Barth J. Conflicts of interest and spin in reviews of psychological therapies: a systematic review. BMJ Open 2016;6(4):e010606.
[37] Nosek BA, Lakens D. Registered reports: a method to increase the credibility of published results. Soc Psychol 2014;45(3):137−41.
[38] Nutu D, Gentili C, Naudet F, Cristea IA. Open science practices in clinical psychology journals: an audit study. J Abnorm Psychol 2019;128(6):510−6.
[39] Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.
[40] Center for Open Science: simple registered report protocol preregistration 2020. Available at http://osf.io/rr/. Accessed August 10, 2020.
[41] An innovative way to publish. Nature 2019;571:447.
[42] Hardwicke T, Ioannidis J. Mapping the universe of registered reports. Nat Hum Behav 2018;2(11):793−6.