Journal of Clinical Epidemiology 173 (2024) 111465

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Substantial discrepancies exist between registered protocol and published manuscript in trials on exercise interventions for chronic low back pain: a metaresearch study

Silvia Bargeri[a,1], Giacomo Basso[b,1], Ignazio Geraci[b], Greta Castellini[a], Alessandro Chiarotto[c], Silvia Gianola[a], Raymond Ostelo[d,e], Marco Testa[b], Tiziano Innocenti[b,d,f,*]

[a]*Unit of Clinical Epidemiology, IRCCS Istituto Ortopedico Galeazzi, Milan, Italy*
[b]*Department of Neurosciences, Rehabilitation, Ophthalmology, Genetic and Maternal Infantile Sciences (DINOGMI), University of Genova, Genova, Italy*
[c]*Department of General Practice, Erasmus MC, University Medical Center, Rotterdam, The Netherlands*
[d]*Department of Health Sciences, Faculty of Science, VU University Amsterdam, Amsterdam Movement Sciences, Amsterdam, The Netherlands*
[e]*Department of Epidemiology and Data Science, Amsterdam UMC Location Vrije Universiteit & Amsterdam Movement Sciences, Musculoskeletal Health, Amsterdam, The Netherlands*
[f]*GIMBE Foundation, Bologna, Italy*

Accepted 9 July 2024; Published online 15 July 2024

Abstract

Objectives: Reporting bias, prevalent in biomedical fields, can undermine evidence credibility. Our objective was to evaluate the proportion of discrepancies between registered protocols and published manuscripts in randomized controlled trials (RCTs) on exercise interventions for patients with chronic low back pain (CLBP).

Study Design and Setting: We conducted a cross-sectional meta-research study, starting from the 2021 ''Exercise therapy for CLBP'' Cochrane Review. We selected all RCTs reporting a protocol registration on a primary register of the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) or in ClinicalTrials.gov. We extracted data from both registered protocol and published manuscript of RCTs, collecting recruitment and administrative information (eg, record dates) and details of trial characteristics (eg, outcomes, arms, statistical analysis plan details [SAPs]). Independent pairs of reviewers assessed discrepancies between registered protocol and published manuscript for the reporting of primary and secondary outcomes domains, measurement instruments, timepoints, number of arms and SAPs(if attached). Outcome discrepancies were characterized as addition, omission, upgrade or downgrade.

Results: We included 116 RCTs reporting an available protocol registration. Overall, 100 RCTs (86.2%) distinguished between primary and secondary outcomes. Of these, 39 RCTs (39.0%) reported one or more discrepancies in primary outcomes, and 78 RCTs (78.0%) reported one or more discrepancies in secondary outcomes. Focusing on discrepancies for the primary outcome, 64.5% of added, upgraded or downgraded outcomes favored statistically significant effects. Few RCTs ($n = 6$) reported discrepancies in the number of arms. SAPs were poorly reported in the registered protocols ($n = 3$) for being compared to the publications.

Conclusion: We found substantial outcome discrepancies comparing registered protocols and published manuscripts in RCTs assessing exercise interventions for patients with CLBP, with some impacting the statistical significance of the effects. Readers are encouraged to approach RCTs results in this field with caution.

Keywords: Meta-epidemiology; Low back pain; Musculoskeletal; Exercise therapy; Selective reporting; Reporting bias

1. Introduction

Low back pain is one of the greatest contributors to years lived with disability and is the first cause of activity limitation, and absence from work [1], with the majority of the costs related to chronicity [2]. Among physical intervention options for chronic low back pain (CLBP), the first recommended is exercise therapy [3], whose effectiveness has been evaluated, and continues to be evaluated, in

Alphabetic order except for co-first, second and last authors.

Funding: Silvia Bargeri, Greta Castellini, and Silvia Gianola were supported and funded by the Italian Ministry of Health. The funder had no role in study design, data collection and analysis, decision to publish, or manuscript preparation.

[1] Co-first authors.

* Corresponding author. Faculty of Science Amsterdam, Department of Health Sciences, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.

E-mail addresses: tiziano.innocenti@gimbe.org; t.innocenti@vu.nl (T. Innocenti).

https://doi.org/10.1016/j.jclinepi.2024.111465

What is new?

Key Findings

- Our meta-research study reveals substantial outcome discrepancies between registered protocols and published manuscripts.

What this adds to what was known?

- Approximately one-third and three-quarters of RCTs showed discrepancies in primary and secondary outcomes, respectively, with half of them favoring statistically significant effects.

What is the implication, and what should change now?

- Researchers should adhere to prospectively registered protocols, documenting any deviations with justifications in both published manuscripts and registered protocols.
- Registries must reinforce the prospective registration of all trial information, including outcomes domains, measurement instruments, time points, arms, and SAPs.
- Clinicians, researchers, peer reviewers, and journal editors are encouraged to thoroughly consult the trial registry of the RCTs to assess reporting bias more deeply.

numerous randomized controlled trials (RCTs) [4]. However, bias in the design, conduct or reporting of RCTs can result in inaccuracies in systematic reviews or guidelines leading to subsequent errors in clinical practice, overestimating or underestimating the effects of an intervention [5].

A prominent issue is ''reporting bias'' [6,7], which may occur when there is a discrepancy between what was planned in the protocol and what is actually reported in the published manuscripts of the RCTs. Reporting bias can take various forms including outcome reporting bias [8,9], when authors select interesting or positive research findings while concealing results that contradict their hypothesis because of the statistical significance, magnitude or direction of the effect; selective reporting of analyses, when researchers conduct multiple analyses within a study but only report those that yield favorable results [10,11], publication bias [12], when studies with positive or statistically significant results are more likely to be published.

Discrepancies related to the reporting biases can occur in various manners, such as adding, omitting, switching data for a measured outcome domain, outcome measurement instrument, time-points or changing SAPs compared to what was reported in the registered protocols [13−16].

A methodological review assessing empirical studies of systematic reviews of RCTs for health care interventions revealed discrepancies in at least one primary outcome, such as additions, omissions, or switching (ie, upgrading from secondary to primary or downgrading from primary to secondary) in 4%−50% of publications compared to the registry entry or published protocol [17].

We therefore hypothesized that there may be discrepancies related to reporting biases between registered protocols and published manuscripts also in CLBP field. Previous study [18] showed that the quality and reporting standards of trials in the exercise for CLBP are suboptimal. Another [19] showed that most RCTs of the field have methodological limitations; however, they did not assess adherence to registered protocol. Therefore, we aimed to assess the proportion of RCTs with discrepancies between the registered protocol and the published manuscript in: primary and secondary outcomes domains, outcome measurement instruments, time-points, number of arms, and SAPs. Moreover, we investigated the characteristics of RCTs with and without outcome discrepancies.

2. Methods

2.1. Study design

We conducted a meta-research study [20,21] prospectively registered [22]. Deviations from protocol are reported in Supplementary File 1A. Since the reporting checklist for methods research studies is currently under development [23], we adapted the items from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) for reporting this metaresearch study [24].

2.2. Eligibility criteria and study selection

We started from the RCTs included in the largest Cochrane systematic review ''Exercise therapy for chronic low back pain'' [4] (n = 249) and in its ''awaiting assessment'' section (n = 172 RCTs for which data have not yet been extracted by Cochrane reviewers). To meet specific criteria for content, quality and validity, accessibility, unique identification, technical capacity and administration, protocols were considered eligible when they reported a registration in one of the international databases included in the primary registers of World Health Organization International Clinical Trial Registry Platform (ICTRP) [25] or in ClinicalTrials.gov, in line with the International Committee of Medical Journal Editors (ICMJE) [26] (Supplementary File 1B). First, a single reviewer (GB) screened the RCTs included in the Cochrane review [4] (n = 249), selecting those that declared a registered protocol (examining the ''characteristics of included studies" section). Next, the same reviewer (GB) assessed the eligibility of the 172 RCTs categorized as ''references to studies awaiting assessment''. Since information about these latter studies was not

extracted in the Cochrane review [4], a second reviewer (SB) cross-checked this phase. Any disagreements were resolved through discussion. More details are reported in Supplementary File 1C.

Published manuscripts that were either not found or not in English were excluded. We also excluded registered protocols (along with the related published manuscripts) under the following conditions: if they were not found, no longer available (ie, expanded access was available but is not currently available and will not be available in the future) [27], had an unknown status of recruitment (ie, a study whose last known status was recruiting, not yet recruiting, or active but not recruiting, and that has passed its completion date with the status not verified within the past 2 years) [27] and if they lacked information (eg, not reporting any outcome information at all).

2.3. Data extraction

Standardized data collection forms were used to record information from the registered protocol and the published manuscript. Two pairs of independent reviewers (SB, GB and IG, SG) extracted the following data from the registered protocol accessed on its registration platforms: registration date, study start date, primary outcome registration date, registered number of arms, description of interventions, SAPs (when attached in the register), primary and secondary outcomes domains (eg, pain), time-points (eg, 1 month follow-up) and outcome measurement instruments (eg, Visual Analoge Scale). We also collected information on how many versions existed.

For data extraction of the manuscript published, a single reviewer (IG) extracted the general characteristics of the RCTs available from the ''characteristics of included studies" section in the Cochrane review (n = 249) [4]. Next, the same reviewer (IG) extracted those characteristics from the full texts retrieved (''references to studies awaiting assessment'') with a second reviewer (SB) cross-checking this phase. Finally, two pairs of independent reviewers (SB, GB; IG, SG) extracted the following data (not available in the Cochrane review [4]): initial date of participant enrollment, involvement of an ethical committee (promoting transparency and accountability in research [28]), presence of a published protocol, journal impact factor, primary and secondary outcome domains, outcome measurement instruments, time points, number of arms, and SAPs. Disagreements were discussed with another reviewer (GC) to reach a final consensus. More details are reported in Supplementary File 1C.

We classified the trial status into 1) prospectively registered and 2) retrospectively registered according to its registration date. Prospective registration was defined as trial registration before or within a month of the first participant enrollment start date according to the protocol [29], the other cases were considered retrospective.

2.4. Detection of discrepancies between registered protocol and published manuscript

We defined discrepancies related to the reporting biases as differences between registered protocol (ie, from the last prospectively registered version released) and the published manuscript. Two pairs of independent reviewers (SB, GB; IG, SG) assessed discrepancies for primary and secondary outcome domains, outcomes measurement instruments, time-points, number of arms SAPs, when attached in the register (see Supplementary File 1B for SAPs contextualization). Disagreements were discussed with another reviewer (GC) to reach a final consensus. We adapted a previously published method [15,17] to classify discrepancies into: addition (eg, completely outcome domain/measurement instrument/time-point or arm added in the publication), omission (eg, completely outcome domain/measurement instrument/time-point or arm omitted in the publication), upgrade (registered secondary outcome changed to primary in the publication), downgrade (registered primary outcome changed to secondary in the publication) or change (eg, statistical analyses changed). If primary and secondary outcomes were not clearly distinguished in the published manuscript, we assessed the discrepancies irrespective of primary and secondary outcome and underlined this in results. More details in Supplementary File 1C.

When multiple versions of the same registered protocol were available, we also collected if discrepancies between the current (ie, latest version) and original (ie, first version) registered versions were present, according to the above-mentioned classification.

2.5. Impact of discrepant outcomes and statistically significant results

A discrepant outcome was considered to favor statistically significant results when: 1) a nonstatistically significant (eg, 95% confidence interval that crossed zero for continuous outcomes) primary outcome registered in the protocol was downgraded to a secondary in the manuscript published; 2) a statistically significant secondary outcome registered in the protocol was upgraded to a primary outcome in the manuscript published; and 3) addition of a nonregistered statistically significant primary outcome in the published manuscript [30].

2.6. Statistical analysis

We used descriptive statistics to assess the proportion of RCTs with discrepancies in primary and secondary outcome, arms and SAPs, as well as the proportion of discrepant outcomes favoring statistically significant effects. We also descriptively compared RCTs with and without primary outcome discrepancies for the following variables supported by literature [15,31]: protocol registration (prospective vs. retrospective), sample size ($>$100 vs

<100 participants), funding status (disclosed vs not disclosed), publication status of the protocol (published vs unpublished) and Journal Impact Factor (JIF) (as continuous variable). Specifically, JIF was assessed since higher JIF journals have typically more rigorous peer review processes [32], potentially preventing discrepancies [33]. Data analysis was performed with STATA 15 software.

3. Results

3.1. Study selection

Overall, we included 116 RCTs reporting an available protocol registration. The flow chart of study selection is reported in Figure. References of included and excluded studies with reasons are reported in Supplementary File 2.

3.2. General characteristics of the published manuscripts

Overall, the 116 RCTs included 11,613 patients with CLBP with a median age of 45 (IQR 41-51) and median percentage of female 63% (52%−74%). Most were two-arm studies (69.0%), comparing exercise to other active interventions (55.0%), published between 2016 and 2020 (69.8%). The median number of primary and secondary outcomes reported was two (IQR 1−2) and three (IQR 2−5), respectively. The majority disclosed the presence of funding (68.1%) while few (n = 24) cited a previously published protocol on indexed journals (2 0.7%). The median JIF was Three (IQR 2−4). Overall, 111 RCTs (95.7%) reported the involvement of an ethical committee (Table 1).

3.3. General characteristics of the registered protocols

Overall, most protocols were registered in ClinicalTrials.gov (n = 60; 51.7%) with half (n = 59; 50.9%) being prospectively registered. The remaining 57 retrospective protocols were registered with a median delay of 492 days (IQR 161−1202) from the first patient recruitment. The median number of primary and secondary outcomes registered in RCTs was two (IQR 1−2) and third (IQR 2−5), respectively. Only three protocols (2.6%) reported registered SAPs (Table 2).

3.4. Proportion of RCTs with outcome discrepancies

Overall, we found 95 RCTs (81.9%) out of 116 with one or more outcome (domains, measurement instruments, time-points) discrepancies. Out of the 116 RCTs included, 16 studies did not report clear distinction between primary and secondary outcomes, preventing us to assess them separately as primary or secondary outcomes (assessments in Supplementary File 3, Table S1). In the remaining 100 RCTs out of 116 (86.2%), we found 39 RCTs (39.0%) with one or more discrepancies in primary outcomes (domains, measurement instruments, time-points) and 78 RCTs (78.0%) with one or more discrepancies in secondary outcomes (domains, measurement instruments, time-points). The most frequent discrepancies were addition and

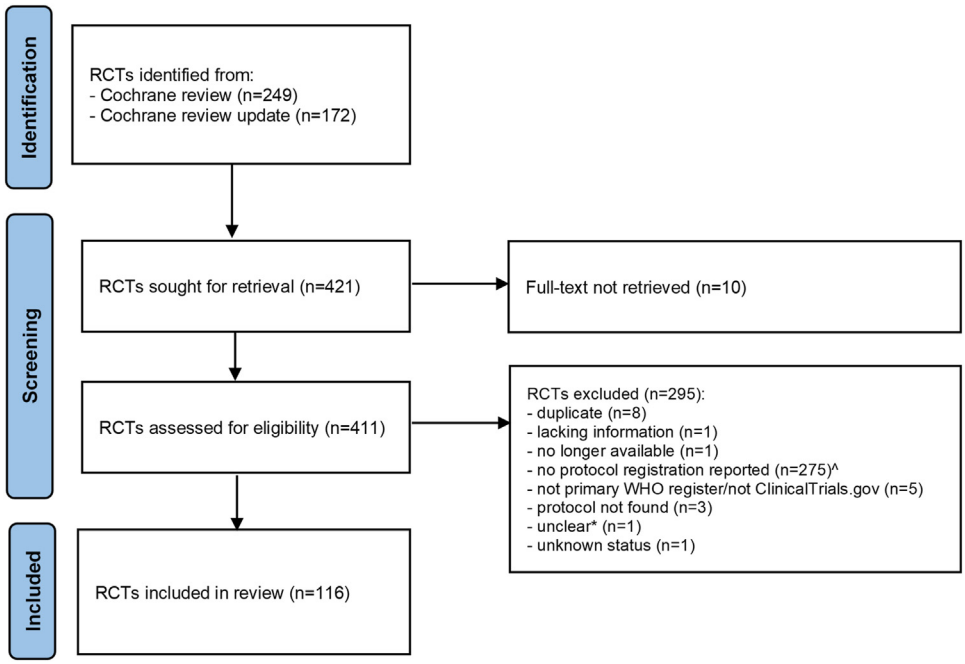

Figure. Flow chart of study selection^of which 186 have already been extracted by the Cochrane review the reported protocol, the reported protocol assesses the risk fall in older population which is ambiguous compared to the published manuscript.(For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1. Descriptive characteristics of the published manuscripts (n = 116)

RCTs characteristics	N	% Out of 116
Country		
Europe	39	33.6
Americas	31	26.7
Asia	34	29.3
Australia	12	10.3
Year		
2006–2010	9	7.8
2011–2015	26	22.4
2016–2020	81	69.8
Comparison		
Two-arm	80	69.0
Exercise vs active intervention	44	55.0
Exercise vs passive/usual care	36	45.0
Three-arm	36	31.0
Two active arms vs one passive	27	75.0
Three active arms	9	25.0
Funding		
Disclosed	79	68.1
Not funded	10	8.6
Not disclosed	27	23.3

RCT, Randomized Controlled Trial.

omission of secondary outcomes domains in 52 (52.0%) and 35 (35.0%) of RCTs, respectively (Table 3).

Three published manuscripts transparently reported reasons for differences between registered protocol and publication. We reported their acknowledged deviations in Supplementary File 3, Table S2.

3.5. Proportion of RCTs with discrepancies in intervention arms and statistical analyses

Overall, six RCTs (5.2%) added or omitted an intervention arm compared to protocol registrations. SAPs were too poorly reported in the registered protocols (n = 3) for being compared to the publications.

3.6. Relationship between discrepant primary outcomes and statistically significant effects

Overall, focusing our unit of analysis on the number of RCTs, among the 31 RCTs with added, upgraded or downgraded primary outcomes domains, measurement instruments or time-points, 20 (64.5%) reported discrepant outcomes with statistically significant effects.

Focusing on the number of outcome domains as unit of analysis, among the 28 added, upgraded or downgraded primary outcomes domains, 19 (67.9%) reported statistically significant effects in favor of exercise interventions (Table 4).

Table 2. General characteristics of the registered protocols (n = 116)

General characteristics	Prospective (n = 59)		Retrospective (n = 57)		Overall (n = 116)	
Registries						
Australian New Zealand Clinical Trials Registry	11	18.6	3	5.3	14	12.1
Brazilian Clinical Trial Registry (ReBeC)	0	0.0	2	3.5	2	1.7
Chinese Clinical Trial Registry	0	0.0	3	5.3	3	2.6
Clinical Research Information Service (CRiS). Republic of Korea	1	1.7	1	1.8	2	1.7
Clinical Trials Registry - India (CTRI)	2	3.4	1	1.8	3	2.6
ClinicalTrials.gov	31	52.5	29	50.9	60	51.7
German Clinical Trials Register	1	1.7	2	3.5	3	2.6
ISRCTN registry	5	8.5	9	15.8	14	12.1
Iranian Registry of Clinical Trials	5	8.5	6	10.5	11	9.5
Pan African Clinical Trials Registry	1	1.7	1	1.8	2	1.7
Thai Clinical Trials Registry	2	3.4	0	0.0	2	1.7
Statistical analysis plans						
Reported yes	3	5.1	0	0.0	0	2.6
Outcomes[a]						
Number of primary outcomes	2	1–2	1	1–3	2	1–2
Number of secondary outcomes	4	2–6	3	1–5	3	2–5
Arms[a]	2	2–3	2	2–2	2	2–3

[a] Median and interquartile range.

Table 3. Proportion of discrepancies for outcome reporting

	Primary outcome (n = 100 RCTs)								
	Domains			Measurement instruments[b]			Time-points[b]		
Discrepancies	*N* RCTs	%	*N* outcomes	*N* RCTs	%	*N* measurement instruments	*N* RCTs	%	*N* time-points
Addition	4	4.0	4	5	5.0	5	15	15.0	18
Omission	11	11.0	13	5	5.0	5	14	14.0	17
Upgrade	6	5.0	7	-	-	-	-	-	-
Downgrade	12	11.0	17	-	-	-	-	-	-
Total[a]	22	22.0	41	7	7.0	10	25	25.0	36
None	78	78.0	-	87	87.0	-	64	64.0	-
NA	0	0.0	-	6	6.0	-	11	11.0	-

N, number of studies with one or more outcome discrepancy in each category; NA, not assessed (ie, incomplete outcome registration as described in methods); RCTs, randomized controlled trials; % out of 100.

[a] We reported the total number of studies with one or more outcome discrepancy (ie, unit of analysis is the study) along with the total number of discrepant outcomes (ie, the unit of analysis is the outcome). Note that, to avoid double counting, the total number of RCTs is not the automatic sum of the number of studies reported in each column since more than one discrepancy type can be present in each study.

[b] In case of discrepancy already reported in outcome domains (eg, added pain and consequently the corresponding measurement instruments and time-point) we considered only as discrepancy the outcome domain and not the corresponding measurement instruments and time-points.

3.7. Characteristics of publications with discrepancies compared to those without discrepancies

Descriptively comparing RCTs with and without discrepancies in their primary outcomes, we observed that RCTs with discrepancies had a lower frequency of sample size exceeding 100 participants compared to those without discrepancies (20.5% vs. 47.5%). In Table 5 all other variables considered are reported.

3.8. Discrepancies between registered protocol versions

Overall, comparing current and original versions (when available in a register), 13 registered protocols out of 116 reported an outcome discrepancy in primary outcome and 16 in secondary outcomes. These discrepancies were performed after the original submitted date (Supplementary File 3, Table S3).

4. Discussion

4.1. Main findings

Our meta-research study found substantial discrepancies related to reporting bias between registered protocols and published manuscripts especially regarding outcomes (81.9%). Over one-third of RCTs exhibited discrepancies in the primary outcome, encompassing domains, measurement instruments, and time-points, with 64.5% of them favoring statistically significant effects. Discrepancies in primary outcome domains (eg, upgrading secondary outcomes to primary) can potentially lead to an overestimation of intervention effects or an underestimation of undesirable effects, given that the power calculation of sample size is based on the primary outcome [34,35]. Furthermore, we identified added or omitted secondary outcomes in three-quarters of the RCTs. The addition of more outcomes increases the likelihood of chance-driven statistically significant results [36], while the omission of outcomes may mask potential null effects.

Similar considerations apply to outcome measurement instruments and time-points. In the rehabilitation context, multiple measurement instruments (eg Roland-Morris Disability Questionnaire; Oswestry Disability Index) [37] are available for the same outcome domain, and authors may be tempted to select the most favorable results. Similarly, added or omitted time-points may be related to the statistical significance (eg, statistical power lost in longest follow-ups [38]). Nevertheless, logistical or limited financial factors should also be considered to justify any protocol deviation [39].

Discrepancies in the number of arms, such as transitioning from a three-armed RCT to a two-arm RCT (or vice versa), were infrequent (6% of the sample), and the assessment of SAPs was challenging since they were poorly reported in the registered protocols. However, SAPs have an important role in reducing the occurrence and facilitating the detection of bias, particularly in relation to selective analysis and reporting [10,11,16]. While SAPs are not mandatory in WHO primary registries [40], some platforms have specific sections or fields where researchers can upload additional documents [41–46]. Current regulations for ClinicalTrials.gov [47] require a copy of the protocol and SAPs to be submitted as part of clinical trial results information for those applicable trials [48] with a Primary Completion Date on or after January 18, 2017 (our sample

Secondary outcome (n = 100 RCTs)								
Domains			Measurement instruments[b]			Time-points[b]		
N RCTs	%	*N* outcomes	*N*	%	*N* measurement instruments	*N* RCTs	%	*N* time-points
52	52.0	102	9	9.0	10	17	17.0	20
35	35.0	79	8	8.0	9	16	16.0	20
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
66	66.0	181	11	11.0	19	26	26.0	40
34	34.0	-	84	84.0	-	63	63.0	-
0	0.0	-	5	5.0	-	11	11.0	-

was published in median in 2018, IQR 2014–2019). To support the reproducibility of research and avoid concerns of misconducting, a clear comprehensive and transparent account of preplanned SAPs must be available. To help researchers in their reporting, some recommendations are available for a minimum set of items that should be addressed [49].

Discrepancies were frequently observed among different versions of the registered protocols. Finally, examining RCT characteristics, prospective registration,

Table 4. Proportion of discrepant primary outcomes favoring statistically significant effects

Discrepant outcomes[a]	*N* outcomes domains favoring statistically significant effects	*N* measurement instruments favoring statistically significant effects	*N* Time-points favoring statistically significant effects
Addition			
Addition of a nonregistered statistically significant primary outcome in the manuscript published	4/4 (100%)	2/5 (40%)	8/18 (44.4%)
Upgrade			
A statistically significant secondary outcome registered in the protocol was upgraded to a primary outcome in the manuscript published	6/7 (85.7%)	-	-
Downgrade			
A nonstatistically significant primary outcome registered in the protocol was downgraded to a secondary in the manuscript published	9/17 (52.9%)	-	-
Total	19/28 (67.9%)	2/5 (40%)	8/18 (44.4%)

The unit of analysis is the number of outcomes.
[a] Omission of outcomes was not reported since statistically significant effects cannot be assessed.

Table 5. Publications with and without discrepancies in the primary outcome

	Discrepancies yes (n = 39)		Discrepancies no (n = 61)	
Dichotomous variables	**N°**	**%**	**N°**	**%**
Prospective registration	19	48.7	32	52.5
Sample size >100	8	20.5	29	47.5
Funding present	24	61.5	47	77.0
Protocol published	19	48.7	31	50.8
Continuous variables	**Median**	**IQR**	**Median**	**IQR**
Journal impact factor	3.3	2.5–5	3	2.3–4.4

IQR, interquartile range.

disclosure of funding sources, JIF and the existence of a published protocol were described similarly in RCTs with and without primary outcome discrepancies. Conversely, smaller studies (eg, less 100 participants) displayed to have more outcome discrepancies. This may be attributed to the fact that low-powered RCTs may struggle to obtain statistically significant and clinically relevant effects [34,35].

4.2. Comparison with previous studies

Our findings are consistent with other studies that have explored discrepancies between registered protocols and published manuscripts [13−16,50]. In the osteoarthritis field, 48% of trials had discrepancies in number, type or time point of primary efficacy outcome between protocol and manuscript [31]. Another cross-sectional study [30] examining reporting bias in exercise oncology trials found evidence of widespread selective outcome reporting and nonreporting bias (outcome switching, omitted preregistered outcomes, and silently introduced novel outcome) in 84% of outcomes.

4.3. Research implications

Preserving the integrity and credibility of RCTs requires strict adherence to prespecified protocols, avoidance of posthoc changes with comprehensive reporting of all outcome details. Despite the ICMJE and WHO [25,29] recommendations, over half of registered protocols were submitted with a median delay of 492 days from the first participants enrollment. Several factors might contribute to these delays, such as lack of awareness of requirements, poor organization and supervision of a research project, or not paying sufficient attention to accuracy [51]. In addition, some primary registries undertake retrospective registration [52]. However, complete accessibility, accountability and transparency can only be achieved when the trial is registered prospectively. Prospectively planning outcomes starting from validated core outcome sets can help to avoid discrepancies between protocols and published manuscript. For LBP, the recommended core outcome domains currently include physical functioning (eg, assessed by Oswestry Disability Index or 24-item Roland-Morris Disability Questionnaire), pain intensity (eg, Numerical Rating Scale), and Health-Related Quality of Life (eg, Short Form Health Survey 12) [53].

Simultaneously, researchers must diligently adhere to prespecified registered protocol, avoiding reporting biases and transparently document any deviations or amendments, along with their reasons, in both the published manuscripts and the updated registered protocols. In our sample, only three RCTs documented protocol deviations. While registration platforms can play a supportive role by encouraging authors to report any deviations and providing clear guidelines or fields for authors to document such changes during the registration process, ultimately is the responsibility of the authors to ensure that such deviations are appropriately reported and disclosed [54]. Current guidelines exist such as the Standard Protocol Item: Recommendations for Interventional Trials (SPIRIT) statement [55] for protocols of clinical trials and the CONSORT for RCTs studies [56], with forthcoming updates in 2024 [57]. The SPIRIT guidelines provide detailed instructions for the content of clinical trial protocols, including how to document and report deviations from preregistered plans [55]. As well, the forthcoming CONSORT 2024, in harmonization to the SPIRIT, is adding a new item regarding protocol deviations [57].

Lastly, peer reviewers and editors play a crucial role in validating the scientific validity of RCTs by checking trial registries and asking authors to report reasons for deviations. In fact, some journals require the registration in public trials registry as condition for publication at or before the onset of patient enrollment [58]. Additionally, adequate training for peer reviewers could enhance their ability to access and assess registered protocols and compare them to the publications. Furthermore, a lack of time is often perceived as a barrier to thoroughly check study protocols; this issue could be addressed by providing some incentives to busy peer reviewers [59,60].

4.4. Strength and limitations

We registered our planned protocol in a preprint repository and followed published reporting standard. We determined that our sample would be large enough to comment on the generalizability of findings to the CLBP field. In fact, among other Cochrane reviews for low back pain [61−67], we started from the largest sample of RCTs collected. However, some limitations should be mentioned. We involved one reviewer (cross-checked by a second reviewer) in selection and extraction of data already extracted by Cochrane reviewers. However, for outcome data and assessment of discrepancies, two pairs of independent reviewers were involved, as recommended [68]. We judged the prospective and retrospective registration considering the gap of days between the first submitted date and the

study start date available in registered protocols. Therefore, we cannot exclude that the study start date could be different than the one registered as few manuscripts reported the actual period of recruitment. To determine if a trial is prospectively registered, it would be more meaningful to consider when the collection of primary outcome data begins [69]. However, the start data collection date for the primary outcome is usually not reported in the trial registration, as it is not a designated field in the major trial registries. In addition, even if half protocols were retrospective, nearly all (95.7%) reported an ethics committee approval which has the role to evaluate, approve and monitor clinical trial protocols according to ethics declarations at International level [28]. For instance, some registries require ethics committee as a mandatory document to register a clinical trial [52]. Then, we planned to assess discrepancies in SAPs, but they were poorly reported, since they were not mandatory in WHO primary registers [40]. In addition, we did not contact authors in case of missing protocol registration or information. We cannot exclude to have missed some protocols or related documents, but we believe that the inclusion of these studies would not alter our main findings.

Finally, we did not assess if nondiscrepant outcomes also favored statistically significant effects to assure a positive association between statistical significance and discrepant outcomes. Likewise, in cases where a primary discrepant outcome favored statistically significant effects, we did not assess whether this influenced the final conclusion of the trial (eg, the trial could report positive conclusion on other outcomes).

5. Conclusion

Substantial outcome discrepancies were found between registered protocols and published manuscripts, threatening the validity of some RCTs on exercise interventions for patients with CLBP. To improve consistency between trial registry data and publications, we firstly recommend trial authors, and then peer reviewers and journal editors to check the protocol registry ensuring that results are consistent and transparently reported. At the same time, registries should mandate the prospectively and complete registration of outcomes and SAPs.

CRediT authorship contribution statement

Silvia Bargeri: Writing − review & editing, Methodology, Investigation, Data curation, Conceptualization. **Giacomo Basso:** Writing − review & editing, Methodology, Investigation, Data curation, Conceptualization. **Ignazio Geraci:** Writing − review & editing, Methodology, Investigation, Data curation, Conceptualization. **Greta Castellini:** Writing − review & editing, Methodology, Investigation, Data curation, Conceptualization. **Alessandro Chiarotto:** Writing − review & editing, Methodology, Conceptualization. **Silvia Gianola:** Writing − review & editing, Methodology, Investigation, Data curation, Conceptualization. **Raymond Ostelo:** Writing − review & editing, Methodology, Conceptualization. **Marco Testa:** Writing − review & editing, Conceptualization. **Tiziano Innocenti:** Writing − review & editing, Methodology, Data curation, Conceptualization.

Data availability

All data generated or analyzed during this study are included in this published article and stored at https://osf.io/yz63c/.

Declaration of competing interest

The authors declare that they have no competing interests.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2024.111465.

References

[1] Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:968−74.
[2] Collaborators GBDLBP. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2023;5:e316−29.
[3] Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CC, Chenot JF, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. Eur Spine J 2018;27:2791−803.
[4] Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. Cochrane Database Syst Rev 2021;9:CD009790.
[5] Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. BMJ 2012;344:d7202.
[6] Page MJ, Higgins JPT, Sterne JAC. Chapter 13: assessing risk of bias due to missing results in a synthesis updated August 2023. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.4. Cochrane; 2023. Available at: www.training.cochrane.org/handbook. Accessed June 3, 2024.
[7] Boutron I, Page MJ, Higgins JPT, Altman DG, Lundh A, Hróbjartsson A. Chapter 7: considering bias and conflicts of interest among the included studies updated August 2023. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.4. Cochrane; 2023. Available at: www.training.cochrane.org/handbook. Accessed June 4, 2024.
[8] Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ 2010;340:c365.

[9] Kirkham JJ, Altman DG, Chan AW, Gamble C, Dwan KM, Williamson PR. Outcome reporting bias in trials: a methodological approach for assessment and adjustment in systematic reviews. BMJ 2018;362:k3802.
[10] Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev 2009;2009:Mr000006.
[11] Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews. J Clin Epidemiol 2013;66:524−37.
[12] Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA 1990;263:1385−9.
[13] 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:e019831.
[14] Serpas VJ, Raghav KP, Halperin DM, Yao J, Overman MJ. Discrepancies in endpoints between clinical trial protocols and clinical trial registration in randomized trials in oncology. BMC Med Res Methodol 2018;18:169.
[15] Wong EK, Lachance CC, Page MJ, Watt J, Veroniki A, Straus SE, et al. Selective reporting bias in randomised controlled trials from two network meta-analyses: comparison of clinical trial registrations and their respective publications. BMJ Open 2019;9:e031138.
[16] Dwan K, Altman DG, Clarke M, Gamble C, Higgins JP, Sterne JA, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. PLoS Med 2014;11:e1001666.
[17] Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. Cochrane Database Syst Rev 2014;2014:Mr000035.
[18] Hayden JA, Ellis J, Ogilvie R, Boulos L, Stanojevic S. Meta-epidemiological study of publication integrity, and quality of conduct and reporting of randomized trials included in a systematic review of low back pain. J Clin Epidemiol 2021;134:65−78.
[19] Cashin AG, Lee H, Bagg MK, O'Hagan E, Traeger AC, Kamper SJ, et al. A systematic review highlights the need to improve the quality and applicability of trials of physical therapy interventions for low back pain. J Clin Epidemiol 2020;126:106−15.
[20] Puljak L, Makaric ZL, Buljan I, Pieper D. What is a meta-epidemiological study? Analysis of published literature indicated heterogeneous study designs and definitions. J Comp Eff Res 2020;9:497−508.
[21] Puljak L. Methodological research: open questions, the need for 'research on research' and its implications for evidence-based health care and reducing research waste. Int J Evid Base Healthc 2019;17:145−6.
[22] Bargeri S, Basso G, Geraci I, Castellini G, Chiarotto A, Gianola S, et al. Discrepancies between registered protocol and final publication in exercise interventions for chronic low back pain: a meta-research study. medRxiv 2023. https://doi.org/10.1101/2023.02.27.23286399.
[23] Lawson DO, Puljak L, Pieper D, Schandelmaier S, Collins GS, Brignardello-Petersen R, et al. Reporting of methodological studies in health research: a protocol for the development of the Methodological STudy reportIng Checklist (MISTIC). BMJ Open 2020;10:e040478.
[24] Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. Evid Base Med 2017;22:139−42.
[25] World Health Organization. Primary registries in the WHO registry network. https://www.who.int/clinical-trials-registry-platform/network/primary-registries. Accessed June 5, 2024.
[26] ICMJE. Available at: https://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html. Accessed June 5, 2024.
[27] ClinicalTrials.gov glossary terms. Available at: https://www.clinicaltrials.gov/study-basics/glossary. Accessed June 3, 2024.
[28] National Patient Safety Agency (NHS). The ethical review process for clinical trials in the European Union. Available at: https://www.ema.europa.eu/en/documents/presentation/ethical-review-process-perspectives-ethics-committee-david-neal_en.pdf. Accessed June 3, 2024.
[29] International Committee of Medical Journal Editors. Clinical trials registration. Available at: http://www.icmje.org/about-icmje/faqs/clinical-trials-registration/. Accessed June 3, 2024.
[30] Singh B, Fairman CM, Christensen JF, Bolam KA, Twomey R, Nunan D, et al. Outcome Reporting bias in Exercise Oncology trials (OREO): a cross-sectional study. medRxiv 2021. https://doi.org/10.1101/2021.03.12.21253378.
[31] Krsticevic M, Dosenovic S, Dimcea DA, Jedrzejewska D, Marques Lameirão AC, Almeida ES, et al. Outcome domains, outcome measures, and characteristics of randomized controlled trials testing nonsurgical interventions for osteoarthritis. J Rheumatol 2020;47:126−31.
[32] Severin A, Strinzel M, Egger M, Barros T, Sokolov A, Mouatt JV, et al. Relationship between journal impact factor and the thoroughness and helpfulness of peer reviews. PLoS Biol 2023;21:e3002238.
[33] Hopewell S, Witt CM, Linde K, Icke K, Adedire O, Kirtley S, et al. Influence of peer review on the reporting of primary outcome(s) and statistical analyses of randomised trials. Trials 2018;19:30.
[34] Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J 2003;20:453−8.
[35] Kadam P, Bhalerao S. Sample size calculation. Int J Ayurveda Res 2010;1:55−7.
[36] Andrade C. Multiple testing and protection against a type 1 (false positive) error using the bonferroni and hochberg corrections. Indian J Psychol Med 2019;41:99−100.
[37] Chiarotto A, Terwee CB, Ostelo RW. Choosing the right outcome measurement instruments for patients with low back pain. Best Pract Res Clin Rheumatol 2016;30:1003−20.
[38] Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. Br Med J 2012;344:e2809.
[39] Fitzpatrick T, Perrier L, Shakik S, Cairncross Z, Tricco AC, Lix L, et al. Assessment of long-term follow-up of randomized trial participants by linkage to routinely collected data: a scoping review and analysis. JAMA Netw Open 2018;1:e186019.
[40] International Clinical Trials Registry Platform (ICTRP)/ICRTP Registry Network/WHO data set. Available at: https://www.who.int/clinical-trials-registry-platform/network/who-data-set. Accessed May 3, 2024.
[41] ANZCTR Data field definitions. Available at: https://www.anzctr.org.au/docs/ANZCTR%20Data%20field%20explanation.pdf. Accessed May 3, 2024.
[42] Deutches register. German clinical trials register. Available at: https://drks.de/search/en. Accessed May 3, 2024.
[43] IRCT. Iranian registry of clinical trial. Available at: https://irct.behdasht.gov.ir/. Accessed May 3, 2024.
[44] ReBEC Frequently asked questions. Available at: https://ensaiosclinicos.gov.br/faq. Accessed May 3, 2024.
[45] EU clinical trials register. Available at: https://www.clinicaltrialsregister.eu/. Accessed June 4, 2024.
[46] ISRCTN registry. Available at: https://www.isrctn.com/. Accessed May 3, 2024.
[47] ClinicalTrial.gov. Frequently asked questions. https://clinicaltrials.gov/policy/faq#fr_8. Accessed June 4, 2024.
[48] Checklist for evaluating whether a clinical trial or study is an applicable clinical trial (ACT) under 42 CFR 11.22(b) for clinical trials initiated on or after January 18, 2017. Available at: https://cdn.clinicaltrials.gov/documents/ACT_Checklist.pdf. Accessed June 3, 2024.

[49] Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017;318:2337−43.
[50] Komukai K, Sugita S, Fujimoto S. Publication bias and selective outcome reporting in randomized controlled trials related to rehabilitation: a literature review. Arch Phys Med Rehabil 2024;105:150−6.
[51] Loder E, Loder S, Cook S. Characteristics and publication fate of unregistered and retrospectively registered clinical trials submitted to the BMJ over 4 years. BMJ Open 2018;8:e020037.
[52] Juneja A, Gupta J, Yadav N, Sharma S, Panchal Y, Adhikari T, et al. An overview of primary registries of WHO's international clinical trial registry platform. Ayu 2019;40:141−6.
[53] Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in nonspecific low back pain. Eur Spine J 2015;24:1127−42.
[54] Tse T, Fain KM, Zarin DA. How to avoid common problems when using ClinicalTrials.gov in research: 10 issues to consider. BMJ 2018;361:k1452.
[55] Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013;346:e7586.
[56] Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.
[57] Tunn R, Boutron I, Chan AW, Collins GS, Hróbjartsson A, Moher D, et al. Methods used to develop the SPIRIT 2024 and CONSORT 2024 Statements. J Clin Epidemiol 2024;169:111309.
[58] Journal of Physiotherapy. Guides for authors. Available at: https://www.sciencedirect.com/journal/journal-of-physiotherapy/publish/guide-for-authors. Accessed March 21, 2024.
[59] Willis JV, Ramos J, Cobey KD, Ng JY, Khan H, Albert MA, et al. Knowledge and motivations of training in peer review: an international cross-sectional survey. PLoS One 2023;18:e0287660.
[60] Innocenti T, Ostelo R, Verhagen A, Pinto RZ, Salvioli S, Giagio S, et al. Rehabilitation journal editors recognize the need for interventions targeted to improve the completeness of reporting, but there is heterogeneity in terms of strategies actually adopted: a cross-sectional web-based survey. J Evid Base Med 2023;16:111−5.
[61] Furlan AD, Giraldo M, Baskwill A, Irvin E, Imamura M. Massage for low-back pain. Cochrane Database Syst Rev 2015;2015:Cd001929.
[62] Mu J, Furlan AD, Lam WY, Hsu MY, Ning Z, Lao L. Acupuncture for chronic nonspecific low back pain. Cochrane Database Syst Rev 2020;12:Cd013814.
[63] van der Gaag WH, Roelofs PD, Enthoven WT, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for acute low back pain. Cochrane Database Syst Rev 2020;4:Cd013581.
[64] Poquet N, Lin CW, Heymans MW, van Tulder MW, Esmail R, Koes BW, et al. Back schools for acute and subacute non-specific low-back pain. Cochrane Database Syst Rev 2016;4:Cd008325.
[65] Marin TJ, Van Eerd D, Irvin E, Couban R, Koes BW, Malmivaara A, et al. Multidisciplinary biopsychosocial rehabilitation for subacute low back pain. Cochrane Database Syst Rev 2017;6:Cd002193.
[66] Rubinstein SM, Terwee CB, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low-back pain. Cochrane Database Syst Rev 2012;2012:Cd008880.
[67] Wegner I, Widyahening IS, van Tulder MW, Blomberg SE, de Vet HC, Brønfort G, et al. Traction for low-back pain with or without sciatica. Cochrane Database Syst Rev 2013;2013:Cd003010.
[68] MECIR Manual. Collecting data from included studies (C43-C51). Available at: https://community.cochrane.org/mecir-manual/standards-conduct-new-cochrane-intervention-reviews-c1-c75/performing-review-c24-c75/collecting-data-included-studies-c43-c51. Accessed June 3, 2024.
[69] Al-Durra M, Nolan RP, Seto E, Cafazzo JA. Prospective registration and reporting of trial number in randomised clinical trials: global cross sectional study of the adoption of ICMJE and Declaration of Helsinki recommendations. BMJ 2020;369:m982.