Journal of Clinical Epidemiology 176 (2024) 111568

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Many randomized trials in a large systematic review were not registered and had evidence of selective outcome reporting: a metaepidemiological study

Samuel Silva[a,b,*], Sareen Singh[a,c], Shazia Kashif[a], Rachel Ogilvie[a], Rafael Z. Pinto[b,d,e], Jill A. Hayden[a]

[a]*Department of Community Health and Epidemiology, Dalhousie University, Halifax, Canada*
[b]*Department of Physical Therapy, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil*
[c]*Faculty of Medicine, Dalhousie University, Halifax, Canada*
[d]*School of Health Sciences, University of New South Wales, Sydney, Australia*
[e]*Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia*

Accepted 9 October 2024; Published online 16 October 2024

Abstract

Objectives: The primary objectives were to describe characteristics of trial registration in the chronic low back pain (CLBP) field and assess the association of trial registration status (registered vs unregistered, prospectively registered vs retrospectively registered) with risk of bias, sufficient sample size, quality of reporting, and treatment effect estimates. Secondary objectives were to describe trial registration consistency with the final report and assess its association with risk of bias, sufficient sample size, and treatment effect estimates.

Study Design and Setting: A cross-sectional metaepidemiological study of trials included in a large Cochrane review on exercise treatments for CLBP. We extracted relevant trial and registration information and assessed trials' risk of bias using the Cochrane Risk of Bias 1 tool. We performed descriptive analyses, logistic regressions, and subgroup meta-analyses.

Results: We included 361 trials, of which 23.3% were prospectively registered. Registered trials had lower risk of bias (odds ratio [OR] 0.6; 95% confidence interval [CI] 0.5, 0.7) and higher reporting quality (OR 1.6; 95% CI 1.4, 1.8) than unregistered trials. Prospectively registered trials were more likely to have low risk of reporting bias (OR 2.7; 95% CI 1.2, 6.5) and higher quality of reporting (OR 1.3; 95% CI 1.1, 1.6) than retrospectively registered trials. Trial registration status was not associated with effect estimates. Among prospectively registered trials, 64.3% clearly defined primary outcome(s) in their registration, 58.3% had consistent sample sizes, and 22.6% had no evidence of selective outcome reporting. Trials that clearly defined primary outcome(s) were more likely to report larger effect estimates for pain intensity (mean difference −15.8; 95% CI −22.7, −8.9 vs −6.0; 95% CI −10.6, −1.5; $Q = 6.7$, $P = .01$), although the difference was small, the 95% CIs overlapped, and no difference was found for functional limitations.

Conclusion: A small proportion of trials in the CLBP field were registered prospectively and many presented registration inconsistencies. Registered trials tend to have lower risk of bias and higher quality of reporting. Policies are needed to improve prospective registration and registration consistency in the field. © 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Randomized controlled trial; Registries; Low back pain; Risk of bias; Selective outcome reporting; Systematic review

Funding: The Canadian Institutes of Health Research provided funding for the 'Exercise for the treatment of chronic low back pain' project that supported the overarching Collaborative Review (Project Grant Competition, PJT-173478). The Collaborative Review Central Team includes co-authors, Jill Hayden (review lead), Rachel Ogilvie (project coordinator), and Shazia Kashif (data analyst).

Preregistration: https://osf.io/zf82a/

* Corresponding author. Centre for Clinical Research, Room 231, 5790 University Ave, Halifax, Nova Scotia B3H 1V7, Canada.
E-mail addresses: ssilvaedf@gmail.com; s.silva@dal.ca (S. Silva).

https://doi.org/10.1016/j.jclinepi.2024.111568
0895-4356/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Plain Language Summary

Prospective trial registration is the practice of documenting the planned methods of a randomized controlled trial on a publicly available online platform (ie, website) before enrolling participants. Medical journals require trialists to prospectively register their trials to encourage the conduct of high-quality research and reduce the chance of trialists changing their research plan to report only positive or significant results (known as selective outcome reporting). We investigated whether trialists within the chronic low back pain field were registering their trials, and whether they followed their registered research plan.We used data from a large systematic review of 456 trials that tested the effectiveness of exercise as a treatment for chronic low back pain. We assessed each trials' registration status and whether prospectively registered trials had inconsistencies between their registered research plan and their research conduct (eg, evidence of selective outcome reporting). We also looked at the association among trial registration with trials' quality of reporting (ie, a marker of research transparency), risk of bias (ie, a marker of research quality), and the amount of low back pain improvement reported by the trials (ie, effect estimates).We found that less than 25% of trials were prospectively registered, and many had inconsistencies between their registered research plan and their research conduct. Overall, registered trials had lower risk of bias and higher quality of reporting. However, trial registration status and selective outcome reporting were not associated with effect estimates (the amount of back pain improvement reported by trials). Our findings highlight the need for trialists and journals to better follow trial registration guidelines and policies in the chronic low back pain field. Knowledge users should be cautious when consuming information from unregistered trials as they appear to be more likely to have quality concerns.

1. Introduction

According to the International Committee of Medical Journal Editors, any trial that began enrollment on or after July 1, 2005 should be registered prospectively [1]. The aim of this statement was to improve research transparency and mitigate selective reporting, publication bias, and unnecessary duplication of studies [2]. Despite this, a considerable proportion of trials from several medical fields remains unregistered or registered retrospectively [3].

There is evidence of increasing unnecessary duplication of trials on exercise for chronic low back pain (CLBP) [4,5]. This highlights the need to investigate current trial registration practices in the CLBP field to guide further strategies to mitigate research waste. Prospective registration can prevent unnecessary duplication by making ongoing projects publicly available to peers and funding agencies [2]. Furthermore, lack of registration has been reported to be a marker of poor transparency and increased risk of bias [6−8]. However, further exploration that considers the specificities of the CLBP field context is needed [4]. Ultimately, the lack of prospective registration may facilitate selective outcome reporting, and it has been suggested that selective outcome reporting is often driven by statistically significant results [9]. Thus, it is important to investigate whether nonprospectively registered trials or trials with registration inconsistencies (eg, registered but with evidence of selective outcome reporting) may be inflating the small effect sizes of exercise treatments for CLBP [5].

Our primary objectives were to (1) describe characteristics of trial registration in the CLBP field and (2) assess the associations of trial registration status (registered vs unregistered, prospectively registered vs retrospectively registered) with risk of bias, sufficient sample size, quality of reporting, and treatment effect estimates. Our secondary objectives were to (1) describe trial registration consistency with the final report and (2) assess the associations of trial registration consistency with risk of bias, sufficient sample size, and treatment effect estimates.

2. Materials and methods

2.1. Study design and data source

We conducted a cross-sectional metaepidemiological study of randomized controlled trials (RCTs) included in a single review. The protocol was prospectively registered (https://osf.io/zf82a/). We have followed guidelines for reporting metaepidemiological research (Appendix 1) [10]. Deviations from the registered protocol are described in Appendix 2.

We considered all trials that were eligible for the update of the 'Exercise treatments for chronic low back pain' Cochrane review [5,11]. The electronic literature searches for the review update were completed on May 19, 2022.

2.2. Eligibility criteria

Complete eligibility criteria for RCTs included in the Cochrane review can be found in the review protocol [11]. For this study, we restricted inclusion to trials with

What is new?

Key findings

- Only one-quarter of the trials on exercise treatments for chronic low back pain were registered prospectively. Registered trials had higher quality of conduct and reporting than unregistered trials.
- Nearly half of prospectively registered trials in the chronic low back pain field had evidence of selective outcome reporting.

What this adds to what is known?

- This study included a large sample of trials and addressed a more specific field than previous studies on trial registration practices, allowing for context-specific interpretations/recommendations.

What is the implication and what should change now?

- Implementation strategies and journal policies to increase prospective registration and consistency/quality of registration are needed in the chronic low back pain field.
- Knowledge users should use caution when consuming evidence from unregistered trials as they may be prone to methodological and transparency concerns.

a listed start date on or after July 1, 2005, based on the dates of data collection or conduct reported in the full trial publication. This date was selected as it reflects the start date of International Committee of Medical Journal Editors's requirements for trial registration [1]. Additional processes for assessing the eligibility of trials that did not report a start date are described in our protocol.

2.3. Data extraction

For this study, we extracted detailed registration information, such as registered outcomes, outcome measures, follow-up time points of assessment, and target sample sizes, using DistillerSR v.2.35 (https://www.distillersr.com/) (see Appendix 3 for details). For trial registrations with multiple versions available, we assessed the version saved prior to, and closest to, the participant enrollment start date. All data were extracted by one author and confirmed (or corrected with consensus) by a second author. Additional detail regarding data extraction for all other variables can be found in the associated Cochrane review publications [5,11] and in Appendix 3.

2.4. Trial registration status

Trials were categorized as prospectively registered, retrospectively registered, or unregistered (Fig 1). Prospective registration was defined as registration of the trial before or within the same calendar month that participant enrollment began. Retrospective registration was defined as registration after the calendar month in which participant enrollment began. Unregistered was defined as those trials with no evidence of a trial registration in the trial publication or registry.

2.5. Trial registration consistency

Trial registration consistency was assessed among prospectively registered trials. We assessed (1) whether

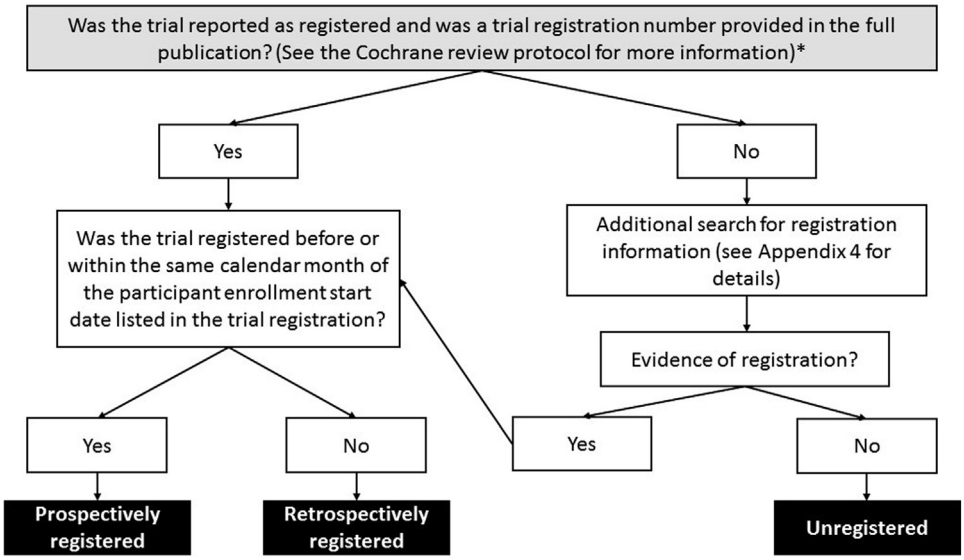

Figure 1. Framework for assessing trial registration status. *Hayden JA, Ogilvie R, Kashif S, Singh S, Boulos L, Stewart S, et al. Exercise treatments for chronic low back pain: a network meta-analysis (Protocol). Cochrane Database Syst Rev. 2023; 6:CD015608.

primary outcomes were clearly defined in the registration; (2) whether the sample size was consistent, by comparing the registered target sample size to the actual number of participants enrolled in the trial; and (3) whether there was evidence of selective outcome reporting. See Table 1 for definitions and assessment criteria.

2.6. Risk of bias and sufficient sample size

To evaluate trial conduct, we assessed risk of bias and assessed the sufficiency of sample size for finding a realistic treatment effect. See Table 2 for details.

2.7. Quality of reporting

Quality of reporting was assessed using a count of 24 items (adapted from [4]), including descriptions of study design (7 items), population (8 items), exercise treatment (7 items), and outcome reporting (2 items). Higher scores indicated higher quality of reporting.

2.8. Treatment effectiveness outcomes

We assessed the effectiveness of exercise treatments compared to no treatment or other conservative treatment, using measures of pain intensity and functional limitations at the follow-up closest to the end of treatment (see our protocol for description of measurement tools included). The individual trial outcomes were rescaled and reported on a 0−100 scale, with negative mean effect sizes indicating improvement (eg, decreased pain intensity).

2.9. Statistical analysis

We described trial registration characteristics using frequencies with proportions, means with standard deviations, and medians with interquartile ranges. We described the proportion of prospectively registered, retrospectively registered, and unregistered trials. We further described the proportion of prospectively registered trials with registration inconsistencies.

We used multiple logistic regression models to estimate associations of trial registration status with risk of bias, sufficient sample size, and quality of reporting. We further investigated associations between trial registration status and each Cochrane Risk of Bias 1 tool bias domain (with blinding items removed from the performance and detection domains). We compared registered with unregistered trials, and prospectively registered with retrospectively registered trials. We also investigated associations of trial registration consistency variables with risk of bias and sufficient sample size. Models were adjusted for funding source (ie, funded or not funded). We presented adjusted estimates as odds ratios (ORs) and 95% confidence intervals (CIs).

We performed subgroup inverse variance, random-effects meta-analyses of treatment effect. Subgroups were defined based on the trial registration status and trial registration consistency variables. We analyzed comparisons of exercise versus no treatment and exercise versus other conservative treatment, where placebo or sham comparisons were grouped with a 'no treatment' group. We prioritized reporting results for exercise versus no treatment comparisons as we expected greater statistical power to find subgroup differences for these comparisons. We calculated mean differences (MDs) and 95% CIs, and reported the between-study variance (tau^2) and the proportion of the observed dispersion in effect sizes due to between-studies heterogeneity (I^2). The Q statistic was used to test for significant subgroup differences ($\alpha = 0.05$). We conducted metaregression analyses to explore potential moderation effects of funding status on the association of trial registration status and trial registration consistency with reported effect estimates. Regression coefficients and 95% CIs were generated. We reported the pooled effect size of the set of prospectively registered trials without registration inconsistencies (ie, primary outcomes clearly defined AND consistent sample size [or inconsistent but with larger sample size in the final report] AND no evidence of selective outcome reporting) for the pain intensity outcome, considering comparisons of exercise versus no treatment. The Hartung-Knapp-Sidik-Jonkman adjustment was applied to all meta-analysis and metaregression models [13]. We used the *R* software (version 4.3.1) for all analyses and the 'meta' package (version 6.5.0) for metaanalyses and metaregressions [14] (code is available in our protocol registration).

2.10. Sensitivity analysis

For the analyses regarding the 'consistent sample size' variable, we performed a sensitivity analysis classifying trials that had larger sample sizes in the publication than the registered target sample sizes as having consistent sample sizes. During data extraction and data cleaning processes for the Cochrane review [11], a 'fatal flaw flag' was added to trials in which serious data reliability concerns were identified (a detailed definition of the 'fatal flaw flag' is described in Appendix 2). Thus, we also conducted a sensitivity analysis omitting trials with a 'fatal flaw flag'.

3. Results

3.1. Characteristics of included trials

Of 456 RCTs eligible for the ongoing update of the Cochrane review, 361 (79.2%) met our study's eligibility criteria and were included (Fig 2). The list of included trials is available in our protocol registration. We were able to access prerecruitment registrations for 77/84 (91.7%) prospectively registered trials.

Table 3 shows the characteristics of included trials. Unregistered trials had a higher proportion of trials conducted

Table 1. Definition and criteria for evaluating the trial registration consistency variables

1. Primary outcome(s) clearly defined	The trial was considered to have clearly defined primary outcome(s) when the outcome construct (eg, pain intensity), measure (eg, Numeric Rating Scale), and time point(s) (eg, 6 weeks after randomization) were adequately described in the registration. In cases where the construct was only described as 'pain' and the measure used was the Visual Analog Scale or Numeric Rating Scale, we considered the construct as adequately described since we could clearly infer from the measure that 'pain intensity' was evaluated. When the construct was not described, but the measure and time point(s) of assessment were, we considered the outcome as 'clearly defined' when it was an established measure for a widely known construct in the field such as the Numeric Rating Scale for pain intensity, and the Roland-Morris Disability Questionnaire or Oswestry Disability Index for functional limitations. We considered time points as adequately described when a specific follow-up time (eg, 12 weeks) was described followed or not by a reference point (eg, after randomization). When only the follow-up time point of assessment was described, we assumed that the temporal reference used was after randomization. When the follow-up time was described using 'postintervention/post-treatment' as a temporal reference and it was not possible to identify the duration of the intervention in the registration, we considered these time points of assessment as not adequately described. In cases of multiple primary outcomes, we applied these criteria for each outcome. For example, if the trial had 3 primary outcomes and the measure was missing for one outcome, the trial was judged as 'primary outcome(s) not clearly defined' in the registration.
2. Sample size consistency	To assess sample size consistency, we compared the registered target sample size and the sample size reported in the full publication (ie, number of participants enrolled in the trial). Sample size was judged to be consistent when there was a difference of 10% or less between the registered sample size and the sample size reported in the full publication.
3. Selective outcome reporting	We considered there to be evidence of selective outcome reporting when at least one of the following 5 inconsistencies was present: (1) a registered primary outcome was reported as a nonprimary outcome in the published trial; (2) a registered primary outcome was omitted in the published trial; (3) a new primary outcome was introduced in the published trial; (4) a registered nonprimary outcome was reported as a primary outcome in the published trial; and (5) the timing of assessment of the registered and published primary outcomes differed (for this item, we accepted discrepancies of up to 20% from the registered time points to avoid capturing discrepancies due to changes in treatment duration or other preplanning issues). We defined primary outcomes in the registration or in the published trial when they were explicitly described as ''primary,'' ''key,'' ''main,'' ''principal,'' ''first,'' ''central,'' ''major,'' ''dominant,'' ''leading,'' or ''foremost'' outcomes or endpoints in the methods section or in the description of the trial objectives (full-text or abstract). If none were explicitly described, we considered the outcome(s) used for sample size calculation as primary. If the primary outcomes remained unclear, we considered the trial as not evaluable, and the trial was not included in the analysis of selective outcome reporting. Registrations with primary outcome(s) not clearly defined and those missing essential data to evaluate selective outcome reporting were not included in the analysis of selective outcome reporting.

in Asia (44.7% vs 13.1% and 18.3% for prospectively registered and retrospectively registered, respectively), a lower proportion of funded trials (29.6% vs 65.5% and 62.0% for prospectively registered and retrospectively registered trials, respectively), and tended to have smaller sample sizes (mean: 55.8 vs 96.5 and 72.0 for prospectively registered and retrospectively registered trials, respectively). Appendix 6 shows the exercise types that were evaluated in the included trials. Unregistered trials had a higher proportion of 'core strengthening' exercise groups tested as interventions (45.6% of exercise groups vs 30.8% and 33.6% for prospectively registered and retrospectively registered trials, respectively). Appendix 7 describes the proportion of trials meeting each of the 24 items of the quality of reporting rubric. Notably, fewer unregistered trials reported funding sources, conflict of interest statements, details about participant flow, randomization method, and adverse events.

3.2. Descriptive results

Among the 361 included trials, 155 (42.9%) were registered; 84 (23.3%) were prospectively registered, and 71 (19.7%) were retrospectively registered (Fig 2).

Among the 84 prospectively registered trials, 54 (64.3%) clearly defined their primary outcomes. Outcome constructs were clearly defined in 70 (83.3%), measurement tools in 68 (80.9%), and the time points of assessment in 69 (82.1%) trials. Forty-nine trials (58.3%) had consistent sample sizes, 11 (13.1%) reported larger sample sizes in the publication than in the registered target sample size, 22 (26.2%) reported smaller sample sizes in the publication

Table 2. Definition and criteria for evaluating the quality of conduct variables risk of bias and sufficient sample size for finding a realistic effect

1. Risk of bias	Risk of bias was assessed using 13 items recommended by the Cochrane Back and Neck Group guidelines [12]. The items were compiled to judge risk of bias according to the 6 Cochrane Risk of Bias 1 tool bias domains (ie, selection bias, performance bias, detection bias, reporting bias, attrition bias, and other biases) [12]. We treated the assessment of the overall risk of bias as a count of the number of Cochrane Risk of Bias 1 tool bias domains rated as some concerns or high risk of bias (traditional overall risk of bias classification would classify nearly all trials as having high risk of bias due to unfeasibility of blinding participants and assessors [5]). Two assessors independently conducted risk of bias assessments, with consensus by a third assessor.
2. Sufficient sample size for realistic effect	We assessed trials to determine if the sample size (ie, participants enrolled in each group) was sufficiently large to detect a realistic effect. Trials were categorized as having a sufficient or insufficient sample size. We defined the smallest sample size necessary to detect (80% power) a mean difference for each outcome (ie, pain intensity and functional limitations), considering each comparison group, based on the pooled effect sizes from the 'Exercise treatments for chronic low back pain' Cochrane review update (results yet to be published) [11] (see Appendix 5 for values). The *R* software (version 4.3.1) was used to perform the power calculations.

than in the registered target sample size, and 2 (2.4%) did not report the target sample size in their registrations. Forty-one trials (48.8%) showed evidence of selective outcome reporting. The proportion of prospectively registered trials meeting each of the 5 criteria for selective outcome reporting ranged from 9.5% to 33.3% (Table 4).

3.3. Trial registration status and consistency versus risk of bias, sufficient sample size, and quality of reporting

Compared to unregistered trials, registered trials had a lower overall risk of bias (OR 0.6; 95% CI 0.5, 0.7) and were more likely to have a low risk of the individual Cochrane Risk of Bias 1 tool domains of selection bias (OR 3.6; 95% CI 2.2, 5.8), attrition bias (OR 2.4; 95% CI 1.4, 4.1), and other biases (OR 2.3; 95% CI 1.3, 4.0). Registered trials also had higher quality of reporting (OR 1.6; 95% CI 1.4, 1.8). There was no difference between prospectively and retrospectively registered trials for overall risk of bias (OR 0.9; 95% CI 0.7, 1.2); however, prospectively registered trials were more likely to have low risk of reporting bias (OR 2.7; 95% CI 1.2, 6.5). Prospectively registered trials were more likely to have higher quality of reporting than retrospectively registered trials (OR 1.3; 95% CI 1.1, 1.6) (Fig 3). Registration consistency variables were not associated with overall risk of bias and sufficient sample size (Table 5). See Appendix 8 for all outputs from the logistic regression models.

3.4. Trial registration status and consistency versus effect estimates

3.4.1. "Exercise versus no treatment" comparisons

We found that trials that clearly defined their primary outcomes reported slightly greater improvements in the pain intensity outcome than trials that did not clearly define their primary outcomes (MD −15.8; 95% CI −22.7, −8.9 vs −6.0; 95% CI −10.6, −1.5, respectively; $Q = 6.7$, $P = .01$) (Fig 4); however, there was an overlap of 95% CIs and no association was found for functional limitations (MD −5.7; 95% CI −7.8, −3.7 vs −9.2; 95% CI −20.1, 1.7, respectively; $Q = 0.5$, $P = .48$) (Fig 5). We found no subgroup differences for effect estimates when comparing registered and unregistered trials, prospectively registered and retrospectively registered trials, trials with consistent sample sizes and trials with inconsistent sample sizes, and trials with evidence of selective outcome reporting and trials without evidence of selective outcome reporting (Fig 4 and Fig 5). The metaregression models showed no moderation effect of funding status for these comparisons (Appendix 9). See Appendix 10 for all outputs from the subgroup meta-analyses models. The pooled effect size of prospectively registered trials without registration inconsistencies for the pain intensity outcome, considering 'exercise versus no treatment' comparisons was MD −8.9; 95% CI −18.0, 0.25; six studies (see Appendix 11 for the forest plot).

3.4.2. "Exercise versus other conservative treatment" comparisons

We found no subgroup differences considering 'exercise versus other conservative treatment' comparisons (Appendix 12).

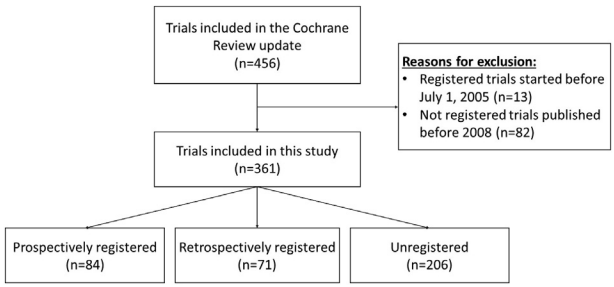

Figure 2. Flowchart of the trial selection process.

Table 3. Characteristics of included studies stratified by trial registration status ($n = 361$)

Characteristic	Registered ($n = 155$)		Unregistered ($n = 206$)
	Prospectively registered ($n = 84$)	Retrospectively registered ($n = 71$)	
Year of publication			
2008–2010	4 (4.8)	4 (5.6)	22 (10.7)
2011–2015	14 (16.7)	10 (14.1)	71 (34.5)
2016–2020	44 (52.4)	42 (59.2)	98 (47.6)
≥2021	22 (26.2)	15 (21.1)	15 (7.3)
Geographic region of trial conduct			
Asia	11 (13.1)	13 (18.3)	92 (44.7)
Africa	3 (3.6)	3 (4.2)	7 (3.4)
Europe	26 (31.0)	16 (22.5)	45 (21.8)
Middle East	10 (11.9)	21 (29.6)	42 (20.4)
North America	12 (14.3)	8 (11.3)	8 (3.9)
South America	11 (13.1)	8 (11.3)	10 (4.9)
Oceania	10 (11.9)	2 (2.8)	1 (0.5)
Mixed regions	1 (1.2)	0 (0.0)	0 (0.0)
Not reported	0 (0.0)	0 (0.0)	1 (0.5)
Population source			
Clinical	42 (50.0)	42 (59.2)	128 (62.1)
Occupational	6 (7.1)	3 (4.2)	12 (5.8)
General or mixed population	31 (36.9)	18 (25.4)	32 (15.5)
Other/not specified	5 (6.0)	8 (11.3)	34 (16.5)
Number of treatment groups in trial			
2	61 (72.6)	56 (78.9)	160 (77.7)
3	20 (23.8)	13 (18.3)	41 (19.9)
4	3 (3.6)	2 (2.8)	4 (1.9)
5	0 (0.0)	0 (0.0)	1 (0.5)
Comparison groups			
Placebo/sham	4 (4.8)	0 (0.0)	5 (2.4)
No treatment	31 (36.9)	20 (28.2)	72 (35.0)
Other conservative treatment	13 (15.5)	11 (15.5)	34 (16.5)
Follow-up periods available[a]			
Close to treatment duration	82 (97.6)	69 (97.2)	198 (96.1)
Moderate term	37 (44.0)	18 (25.4)	43 (20.9)
Long term	20 (23.8)	11 (15.5)	8 (3.9)
Funding status			
Funded	55 (65.5)	44 (62.0)	61 (29.6)
Not funded	29 (34.5)	27 (38.0)	145 (79.4)
Number of subjects in study (mean ± SD)	96.5 ± 77.8	72.0 ± 45.2	55.8 ± 46.4
Number of subjects in study (median [IQR])	67.5 [44.7–129.2]	60.0 [40.5–91.0]	41.5 [30.0–63.0]
Number of subjects per group (mean ± SD)	92.4 ± 73.3	69.7 ± 45.9	55.0 ± 44.8
Number of subjects per group (median [IQR])	65.0 [44.0–116.5]	60.0 [39.0–85.0]	41.5 [30.0–63.0]

Data are presented as mean ± standard deviation, median (interquartile range), or frequency (percentage).
IQR, interquartile range; SD, standard deviation.
[a] Close to treatment duration = follow-up measured within 1 week before or after the end of treatment; moderate term = follow-up measured at more than 14 weeks to 47 weeks after the end of treatment, closest to 6 months; long term = follow-up measured at least 48 weeks after the end of treatment, closest to 12 months.

Table 4. Number and percentage of prospectively registered trials (n = 84) that met, did not meet, or were not evaluable for each criterion for selective outcome reporting evaluation

Selective reporting evaluation criteria	Yes	No	Not evaluable[a]
1. A registered primary outcome was reported as a nonprimary outcome in the published trial	11 (13.1)	56 (66.7)	17 (20.2)
2. A registered primary outcome was omitted in the published trial	19 (22.6)	64 (76.2)	1 (1.2)
3. A new primary outcome was introduced in the published trial	17 (20.2)	50 (59.5)	17 (20.2)
4. A registered nonprimary outcome was reported as a primary outcome in the published trial	8 (9.5)	59 (70.2)	17 (20.2)
5. The timing of assessment of the registered and published primary outcomes differed	28 (33.3)	32 (38.1)	24 (28.6)
Overall evidence of selective outcome reporting	41 (48.8)	19 (22.6)	24 (28.6)

The trial was considered to show evidence of selective outcome reporting when at least one of the five criteria was met.
[a] Due to unclear definition of primary outcomes in the publication or unclear time point of assessment in the registration.

3.5. Sensitivity analysis

There were no changes in the interpretation of our main findings after conducting sensitivity analyses.

4. Discussion

4.1. Main findings and interpretations

4.1.1. Trial registration status

Less than a quarter of published trials on exercise treatments for CLBP were prospectively registered (23.3%), a smaller proportion than those found for other medical disciplines [7,15]. Registered trials had lower overall risk of bias, corroborating previous studies [7,8], and higher quality of reporting than unregistered trials. Trial registration was also associated with lower risk of bias for specific domains, with risk of selection bias showing the strongest association. Trial registration practices seem to be more established among higher-quality, larger, and funded trials. We hypothesize that these trials are more likely to have been conducted by experienced research teams, more familiar with current research methods and reporting standards. Strategies to increase awareness of the importance of

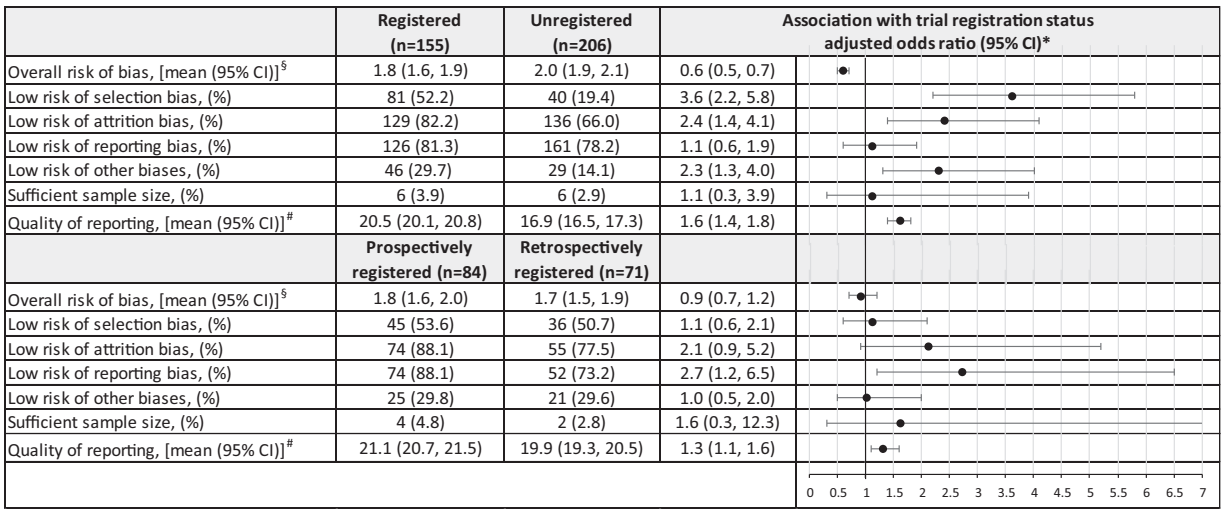

Figure 3. Effect estimates from the logistic regression models on the association between trial registration status with quality of conduct and reporting variables (odds ratios and 95% confidence intervals). We compared registered trials with unregistered trials and prospectively registered trials with retrospectively registered trials. *Abbreviations*: CI, confidence interval *Models were adjusted for funding status. §Overall risk of bias was defined as a count of the number of Cochrane Risk of Bias 1 tool domains rated as some concerns or high risk of bias (0–six domains). Higher scores mean higher risk of bias. #Quality of reporting was assessed using a 0–24 scale, with higher scores meaning higher quality of reporting.

Table 5. Association of trial registration consistency variables with overall risk of bias and sufficient sample size

Trial registration consistency variables	Overall risk of bias[b]	Sufficient sample size for detecting a realistic effect[c]
Primary outcomes clearly defined (n = 84)[a]	1.0 (0.7, 1.5)	0.2 (0.0, 1.8)
Sample size consistency (n = 84)[a]	1.0 (0.7, 1.4)	0.7 (0.1, 6.2)
Selective outcome reporting (n = 60)[a]	1.0 (0.7, 1.6)	4.1 (0.4, 94.3)

Estimates are presented as odds ratios and 95% confidence intervals. Models were adjusted for funding status.

[a] Trials with primary outcomes not clearly defined, inconsistent sample size, and evidence of selective outcome reporting as references, respectively.

[b] Overall risk of bias was defined as a count of the number of Cochrane Risk of Bias 1 tool domains rated as some concerns or high risk of bias (0–6 domains). Higher scores mean higher risk of bias.

[c] Insufficient as reference (sufficient sample size, n = 4; insufficient sample size, n = 80).

trial registration must be targeted beyond the regions where the dominant research teams in the field are concentrated. We found no significant difference between prospectively and retrospectively registered trials for overall risk of bias, which corroborates a previous study with trials from diverse medical fields [7]. However, prospective registration was associated with lower risk of reporting bias and higher quality of reporting. This is indicative of greater transparency and more complete reporting of data from prospectively registered trials.

We found no association between trial registration status and treatment effect estimates. The literature on this topic seems to be conflicting; some studies found that registered trials report smaller effects estimates and are less likely to find positive findings [16,17] while other studies failed to find any differences [18,19]. Our findings somewhat contradict a previous study by members of our team that looked at integrity concerns in trials on exercise for CLBP published up to 2018 and found, in unadjusted analyses, that unregistered trials tended to report slightly greater improvements in pain than registered trials [4]. In the previous study, publication date was not employed as an eligibility criterion, unlike this study, where we excluded trials with a start date before July 2005 or published before 2008. Therefore, the previous study included a set of older trials (more commonly unregistered) compared to the set of trials included in this study. Notably, 137 new trials were added in this study. Older trials may be more prone to novelty bias and may be more likely to report larger effect estimates, which may help explain the different findings [20].

4.1.2. Registration consistency

A considerable proportion of trials included unclear definitions of primary outcomes in their registrations (35.7%). According to the World Health Organization guidelines [21], trialists should clearly indicate in the registrations the construct they are interested in measuring, the measurement tool used to measure the construct, and the planned

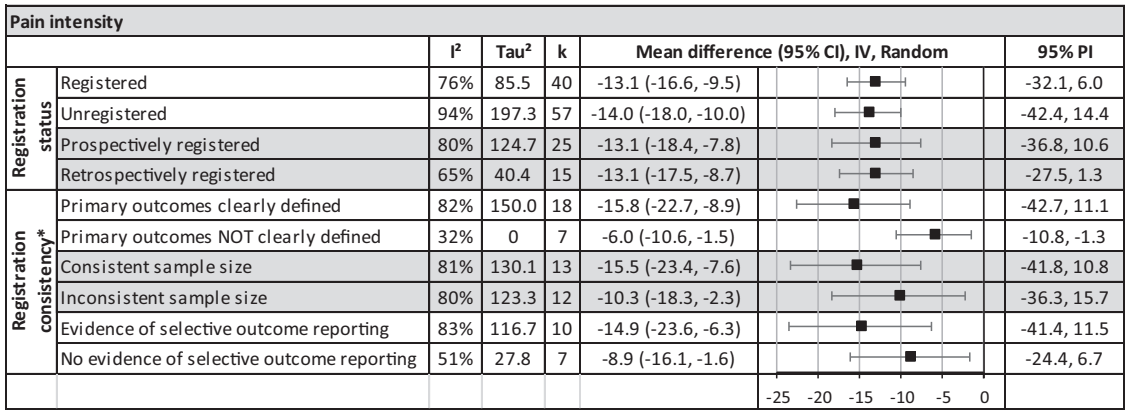

Figure 4. Summary of parameters from subgroup meta-analyses for the pain intensity outcome (exercise vs no treatment). Pain intensity measures were rescaled to a 0–100 scale, with negative mean effect sizes indicating improvement. The Q statistic was used to test for significant differences between subgroups (α = 0.05). We performed the following comparisons: registered trials and unregistered trials (Q = 0.12, P = .72), prospectively registered trials and retrospectively registered trials (Q = 0.00, P = .99), trials that clearly defined their primary outcomes in the registration and trials that did not clearly define their primary outcomes (Q = 6.67, P = .01), trials with consistent sample sizes and trials with inconsistent sample sizes (Q = 1.03, P = .31), and trials with evidence of selective outcome reporting and trials with no evidence of selective outcome reporting (Q = 1.58, P = .21). * Trial registration consistency variables were assessed only among prospectively registered trials. *Abbreviations*: CI, confidence interval; IV, inverse-variance method; k, number of eligible trials; PI, prediction interval.

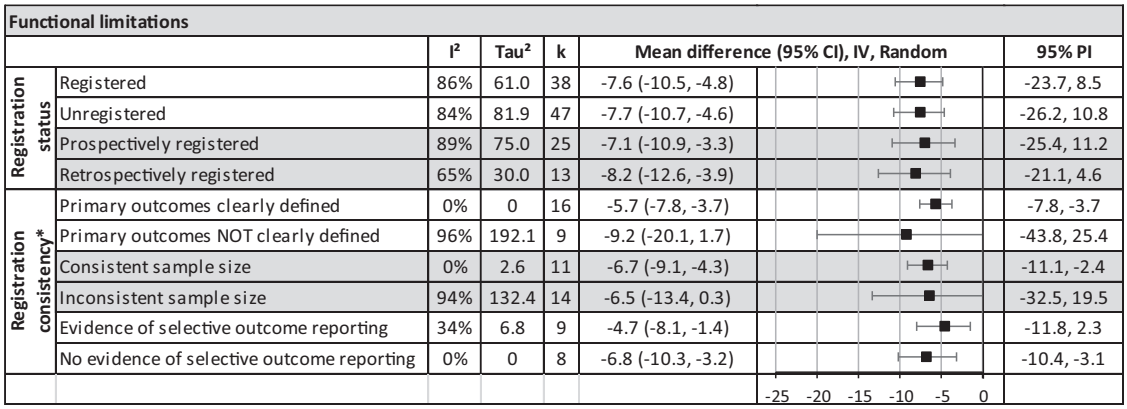

Figure 5. Summary of parameters from subgroup meta-analyses for the functional limitations outcome (exercise vs no treatment). Functional limitations measures were rescaled to a 0−100 scale, with negative mean effect sizes indicating improvement. The *Q* statistic was used to test for significant differences between subgroups ($\alpha = 0.05$). We performed the following comparisons: registered trials and unregistered trials ($Q = 0.00$, $P = .98$), prospectively registered trials and retrospectively registered trials ($Q = 0.19$, $P = .66$), trials that clearly defined their primary outcomes in the registration and trials that did not clearly define their primary outcomes ($Q = 0.51$, $P = .48$), trials with consistent sample sizes and trials with inconsistent sample sizes ($Q = 0.00$, $P = .95$), and trials with evidence of selective outcome reporting and trials with no evidence of selective outcome reporting ($Q = 0.97$, $P = .32$). * Trial registration consistency variables were assessed only among prospectively registered trials. *Abbreviations*: CI, confidence interval; IV, inverse-variance method; k, number of eligible trials; PI, prediction interval.

time points for follow-up assessments. Lack of clear and complete description of primary outcomes can lead to selective outcome reporting and make it difficult for peer-reviewers to verify whether the authors have followed the research protocol. We encourage trial registration platforms to accept only registrations with descriptions of primary outcomes that meet the World Health Organization guidelines.

Trials with primary outcomes clearly defined tended to report larger treatment effect estimates for the pain intensity outcome. Sensibly, previous evidence suggests that when outcomes are preplanned to match the outcomes that the exercise intervention under investigation actually targets, trials are more likely to find positive results for CLBP-related outcomes [22]. We speculate that trialists that clearly defined their primary outcomes may have better planned their trials and selected outcomes based on a stronger rationale, which may be associated with larger effect estimates. However, interpretation of this finding must be made with caution as its clinical relevance is uncertain (small difference and overlapping 95% CIs) and no association was found for functional limitations.

Nearly a quarter of prospectively registered trials reported smaller sample sizes in the publication than the registered target sample sizes. Due to the complexity of conducting a trial, challenges related to recruitment might occur. However, trialists must perform power calculations a priori and carefully consider the feasibility of recruiting the required number of participants to detect a realistic effect. The CLBP evidence base is currently oversaturated with trials of exercise for CLBP; further 'underpowered' trials are not needed and are unlikely to bring significant contributions [5].

Nearly half of prospectively registered trials showed evidence of selective outcome reporting. Overall, roughly 10% of included trials were registered prospectively and had no evidence of selective outcome reporting. A study with trials from broad physiotherapy subdisciplines found a larger prevalence of selective outcome reporting (48% vs 73%) [15]. This difference may be explained by the more flexible criteria that we used to define selective outcome reporting, which considered the CLBP field context. We encourage trialists to report outcomes as described in their registered protocol, with any deviations described clearly within the publication, as recommended in the consolidated standards of reporting trials checklist [23].

Although we found a tendency for trials with selective outcome reporting to report larger treatment effect estimates for pain intensity, the difference was not statistically significant. Only 17 trials were included in the subgroup meta-analyses; therefore, we acknowledge the possibility of being underpowered to perform this comparison.

4.2. Strengths and limitations

This study adds to the evidence on the importance of prospective trial registration by including a larger sample of trials than previous studies [15,17,24]. By restricting our study to only trials of CLBP, we were able to explore a homogeneous sample including all available trials in the field. We acknowledge that our findings might not be generalizable to other fields. In addition, we drew on widely used and established criteria to assess trial registration consistency variables [15,21,25−27], adapting the criteria to fit the context of the CLBP field.

The relatively low proportion of prospectively registered trials limited our power to investigate potential associations of trial registration consistency with the variables investigated. Furthermore, it is worth noting that only 12 trials had sufficient samples to detect a realistic effect, which made it difficult to find any association between sample size and trial registration status.

5. Conclusion

While prospective trial registration does not ensure a high-quality trial, it does provide evidence-consumers with sufficient information to make an educated assessment. Although prospective registration practices and the use of reporting guidelines (eg, consolidated standards of reporting trials checklist) have been widely endorsed by medical journals, this has not been successful in preventing the publication of unregistered trials and trials with unjustified registration/reporting inconsistencies. We recommend the implementation of stricter journal policies to increase prospective registration and registration consistency within the CLBP field; enforcement of registration and reporting guidelines would need to take place at the publication level.

Evidence from prospectively registered trials should be prioritized; we encourage knowledge users to be cautious when consuming unregistered trials, as they were found to lack transparency and be of lower overall quality. When possible, we encourage evidence consumers to cross-reference trial registrations against trial reports to look for evidence of selective outcome reporting. Trial registration assessment could be part of a multicomponent assessment of research integrity in evidence syntheses. This is being implemented in the Cochrane review that provided data for this study [11], where problematic trials have been excluded from the review. Such a criterion could help reduce the impact of poor research transparency and malpractice when interpreting the body of evidence.

Ultimately, trial registration status and selective outcome reporting were not associated with reported treatment effect estimates. We presume that (lack of) registration and selective outcome reporting may not have a significant impact on the pooled effect sizes of exercise treatments for CLBP.

CRediT authorship contribution statement

Samuel Silva: Writing − original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Sareen Singh:** Writing − review & editing, Methodology, Investigation, Conceptualization. **Shazia Kashif:** Methodology, Formal analysis, Data curation. **Rachel Ogilvie:** Writing − review & editing, Investigation, Data curation, Conceptualization. **Rafael Z. Pinto:** Writing − review & editing, Conceptualization. **Jill A. Hayden:** Writing − review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

None.

Acknowledgments

The authors acknowledge Leah Boulos and Kristy Hancock, Evidence Synthesis Coordinators at the Maritime SPOR SUPPORT Unit for assistance with the literature search for the collaborative review. The authors acknowledge the contributions of Collaborative Review Leads: Lisandra Almeida de Oliveira, Geronimo Bejarano, Kasper Bülow, Carol Cancelliere, Annemarie de Zoete, Fabianna Jesus-Moraleida, Tiê Parma Yamato, Bruno Saragiotto, Lisa Susan Wieland, and additional central and subreview team members contributing to screening and data extraction: Nora Bakaa, Jennifer Cartwright, Gaelan Connell, Cristiano Costa, Ben Csiernik, Stephanie Di Pelino, Junior Vitorino Fandim, Shireen Harbin, Wilhelmina IJzelenberg, Carsten Bogh Juhl, Mariana Leite, Alanna MacDonald, Luciana Macedo, Devin Manning, Diego Roger-Silva, Pedro Isaac Santos Chaves, Heather Shearer, Daniele Sirineu Pereira, Danielle Southerst, Maria N Wilson, Jessica Wong, Leslie Verville, and Hainan Yu. The authors thank Somayyeh Mohammadi for support with Persian language trials and Hainan Yu for support with Chinese language trials. The authors would like to thank their patient advisor, Heather Taylor. The authors appreciate advice and guidance from their Collaborative Review Working Group members: Rachelle Buchbinder, Manuela Ferreira, Andrea Furlan, Jan Hartvigsen, Toby Lasserson, Chris Maher, Amir Qaseem, Peter Tugwell, and Maurits van Tulder.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2024.111568.

Data availability

We have shared the code and the dataset containing trial registration data in our OSF registration. We will share an updated dataset upon request when the Cochrane review that contributed the data is published.

References

[1] Angelis C De, 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. N Engl J Med 2004;351(12): 1250−1.

[2] Organization WH. Trial registration [Internet]. Why is trial registration Important? 2023. Available at: https://www.who.int/clinical-trials-registry-platform/network/trial-registration. Accessed September 25, 2023.

[3] Lindsley K, Fusco N, Teeuw H, Mooij E, Scholten R, Hooft L. Poor compliance of clinical trial registration among trials included in systematic reviews: a cohort study. J Clin Epidemiol 2021;132:79−87.

[4] 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.

[5] Hayden J, Ellis J, Ogilvie R, Malmivaara A, van Tulder M. Exercise therapy for chronic low back pain (Review). Cochrane Database Syst Rev 2021;9(9):CD009790.

[6] Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E. The knowledge system underpinning healthcare is not fit for purpose and must change. BMJ 2015;350:h2463.

[7] Tan AC, Jiang I, Askie L, Hunter K, Simes RJ, Seidler AL. Prevalence of trial registration varies by study characteristics and risk of bias. J Clin Epidemiol 2019;113:64−74.

[8] Lindsley K, Fusco N, Li T, Scholten R, Hooft L. Clinical trial registration was associated with lower risk of bias compared with non-registered trials among trials included in systematic reviews. J Clin Epidemiol 2022;145:164−73.

[9] Hopewell S, Loudon K, Clarke M, Oxman A, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results (Review). Cochrane Database Syst Rev 2009;2009(1): MR000006.

[10] Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. Evid Based Med 2017;22(4):139−42.

[11] Hayden JA, Ogilvie R, Kashif S, Singh S, Boulos L, Stewart S, et al. Exercise treatments for chronic low back pain: a network meta-analysis (Protocol). Cochrane Database Syst Rev 2023;6:CD015608.

[12] Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 updated method guideline for systematic reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976) 2015;40(21):1660−73.

[13] Inthout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol 2014;14(1):1−12.

[14] Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 2019;22(4): 153−60.

[15] Santos N, Elkins MR, Lemes IR, Stubbs PW, Franco R, Pinto RZ. Clinical trial registration has become more prevalent in physical therapy but it is still inadequate: a meta-research study. Musculoskelet Sci Pract 2023;67:102854.

[16] Dechartres A, Ravaud P, Atal I, Riveros C, Boutron I. Association between trial registration and treatment effect estimates: a meta-epidemiological study. BMC Med 2016;14(1):1−9.

[17] Papageorgiou SN, Xavier GM, Cobourne MT, Eliades T. Registered trials report less beneficial treatment effects than unregistered ones: a meta-epidemiological study in orthodontics. J Clin Epidemiol 2018;100:44−52.

[18] Odutayo A, Emdin CA, Hsiao AJ, Shakir M, Copsey B, Dutton S, et al. Association between trial registration and positive study findings: cross sectional study (Epidemiological Study of Randomized Trials - ESORT). BMJ 2017;356:1−11.

[19] Haring R, Ghannad M, Bertizzolo L, Page MJ. No evidence found for an association between trial characteristics and treatment effects in randomized trials of testosterone therapy in men: a meta-epidemiological study. J Clin Epidemiol 2020;122:12−9.

[20] Luo Y, Heneghan C, Persaud N. Catalogue of bias: novelty bias. BMJ Evidence-Based Med 2023;28(6):410−1.

[21] Organization WH. WHO trial registration data set (version 1.3.1) [Internet]. Available at: https://www.who.int/clinical-trials-registry-platform/network/who-data-set. Accessed December 18, 2023.

[22] Wood L, Foster NE, Lewis M, Bishop A. Exercise interventions for persistent non-specific low back pain - does matching outcomes to treatment targets make a difference? A systematic review and meta-analysis. J Pain 2021;22(2):107−26.

[23] 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.

[24] Taylor NJ, Gorman DM. Registration and primary outcome reporting in behavioral health trials. BMC Med Res Methodol 2022;22(1):1−8.

[25] Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparasion of registered and published primary outcomes in randomized controlled trials. JAMA 2009;302(9):977−84.

[26] Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291(20):2457−65.

[27] Sun LW, Lee DJ, Collins JA, Carll TC, Ramahi K, Sandy SJ, et al. Assessment of consistency between peer-reviewed publications and clinical trial registries. JAMA Ophthalmol 2019;53226(5):552−6.