UNIVERSITÀ DEGLI STUDI
DI GENOVA

campus savona

Università degli Studi di Genova

Scuola di Scienze Mediche e Farmaceutiche

Dipartimento di Neuroscienze, Riabilitazione, Oftalmologia, Genetica e Scienze Materno-Infantili

Master in Riabilitazione dei Disordini Muscoloscheletrici

A.A. 2022/2023

Campus Universitario di Savona

Discrepancies between registered protocol and final publication in exercise interventions for chronic low back pain: a meta-research study

Candidato:

Dott. Giacomo Basso

Relatore:

Dott. Tiziano Innocenti

Index

Abstract 3
1. Introduction 4
2. Methods 5
2.1 Study design 5
2.2 Eligibility criteria 5
2.3 Data extraction 5
2.4 Detection of discrepancies between registered protocol and final publication 6
2.5 Comparison between discrepant outcomes and statistically significant results 7
2.6 Statistical analysis 7
3. Results 8
3.1 Study selection 8
3.2 Final publications general characteristics 8
3.3 Registered protocol general characteristics 9
3.4 Discrepancies between registered protocol and final publication in outcomes 10
3.5 Discrepancies between registered protocol and final publication in arms and statistical analyses 11
3.6 Relationship between discrepancies in primary outcomes definition and p value 11
3.7 Association between the characteristics of publications with any kind of discrepancies compared to those without discrepancies in the primary outcome 11
4. Discussion 13
4.1 Main findings 13
4.2 Comparison with previous studies 14
4.3 Research implications 14
4.4 Strength and Limitations 14
5. Conclusion 15
Appendix A. Protocol 18

Abstract

Background: Discrepancies between registered protocol and final publication are common in randomized controlled trials (RCTs) of various medical fields, undermining their integrity and credibility. However, this has not yet been assessed in exercise RCTs for chronic low back pain (cLBP).

Objectives: We aimed to assess the prevalence of discrepancies between the registered protocol and final publication for primary and secondary outcomes definitions, outcomes measurement, timepoints, number of arms and statistical analysis plans in cLBP exercise RCTs.

Methods: We performed a meta-research study, prospectively registered (MEDRXIV/2023/286399). We started from the RCTs included in the 2021 Cochrane review "Exercise therapy for chronic low back pain" to select 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. Eligible registered protocol and final publication were compared to identify discrepancies.

Results: We included 80 RCTs reporting an available protocol registration. Less than half protocols (n=38) were prospectively registered. Overall, we found 142 different discrepancies in outcomes between protocol registrations and publications in 66 RCTs (82.5%): 53 for primary outcomes (n=18 in outcome definitions, n=9 in measurements, n=26 in timepoints) in 37 RCTs (46.3%) and 89 for secondary outcomes (n=55 in outcome definitions, n=12 in measurements, n=22 in timepoints) in 62 RCTs (77.5%). Focusing on discrepancies for primary outcome definition, 53.3% favored p-value. Very few discrepancies (n=5) were found in the number of arms. Statistical analysis plans were poorly reported by registers (n=2) to being compared to publication.

Conclusion: Our findings offer evidence of common outcome discrepancies between registered protocols and final publications with some favoring positive conclusions in cLBP exercise RCTs. We recommend clinicians, researchers, peer reviewers and journal editors to consult the trial registry of the RCT to ensure that the results are consistent with the prospective registration and transparently reported.

1. Introduction

Low back pain is one the greatest contributors to years lived with disability and is the first cause of activity limitation, and absence from work[13]. One widely used intervention for chronic low back pain (cLBP) is exercise therapy, which has been examined in numerous randomized controlled trials (RCTs)[11]. Due to their important effect on clinical practice, there is a necessity to have transparent reporting of RCT results[15]. However, bias in the design, conduct or reporting of RCTs can result in inaccuracies in systematic reviews or guidelines and subsequent errors in clinical practice[9].

Several meta-research studies in the medical field[6, 8, 9, 16, 29, 31] have shown that discrepancies between what is reported in the registered protocol and what is reported in the final publication are common. This can lead to selective reporting bias and refers to a publication practice where study authors preferentially publish interesting or positive research findings while concealing results that do not confirm their hypothesis because of the statistical significance, magnitude or direction of the effect[23 19 24]. Despite some improvement over time, it has been shown that the study quality and reporting of trials in the exercise for cLBP field continue to be lacking[10]. However, it is still unclear what is the prevalence of discrepancies between the registered protocol and final publication in these trials.

This could strongly affect the conclusions of systematic review, overestimating the effects of an intervention or underestimating its undesirable effect, compromising the credibility of the evidence synthesis itself.

Starting from the largest updated Cochrane review on the effectiveness of exercise intervention in cLBP[11] we aimed to assess:

- The prevalence of RCTs with a discrepancy between the registered protocol and final publication in primary and secondary outcomes according to outcomes measures, timepoints, number of arms, and statistical analysis plans

- The characteristics of RCTs with and without discrepancies between the registered protocol and final publication

2. Methods

2.1 Study design

We conducted a meta-research study [27, 28] prospectively registered (MEDRXIV/2023/286399)[5]. Since the reporting checklist for methods research studies is currently under development[20], we adapted items from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) for reporting meta-research studies[22].

2.2 Eligibility criteria

We started from the RCTs included in the 2021 Cochrane review"[11] (n=249) and identified in its update (n=172) to select all RCTs reporting a protocol registration. Hayden et al. included RCTs that compared exercise to no treatment, usual care, placebo or another conservative intervention among adults with cLBP. Trials could include interventions provided to participants in any setting (e.g., healthcare, occupational, general and mixed populations). The intervention could have been combined with or without the addition of other components (eg, education, manual therapy).

Protocols were considered when they were registered to a primary register of the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) or in ClinicalTrials.gov according to the International Committee of Medical Journal Editors (ICMJE)[14].

If no information about protocol registration is reported, if the protocol is not available or not in English language, we excluded it.

2.3 Data extraction

Standardized data collection forms were used to record information from registered protocol and the final publication of the trial.

For data extraction of the registered protocol, we collected: registration date, study start date, primary outcome registration date, primary completion date (i.e., date of final collection of data for the primary outcome), registered number of arms, description of interventions, statistical analysis plans, nature and number of primary and secondary outcomes (e.g., pain), time points (e.g., 1 month follow up) and outcome measurements (e.g., visual analogue scale). We also collected how many versions of the registered protocol exists.

For data extraction of the final publication, we used the dataset of the Cochrane review to extract RCTs' general characteristics (e.g., author, year, ID number of the protocol registration and/or reference of protocol publication, initial date of participant enrollment, setting, sample size, cLBP duration (e.g., months), radicular symptoms (e.g. leg pain and/or neurological symptoms), mean age, sex, conflict of interests, funding (non-industry/industry-sponsored), journal of publication, journal impact factor (JIF), number of arms, description of interventions, statistical analysis, nature and number of primary and secondary outcomes, time points and outcome measurements).
We classified the trial status into 1) prospectively registered; 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 [1].

2.4 Detection of discrepancies between registered protocol and final publication

We defined discrepancies as differences between registered protocol (i.e., from the last prospectively registered version released) and final publication. To ensure a comprehensive assessment, we checked related documents for each RCT (e.g., published protocol, statistical analysis plans, supplementary materials).
Two pairs of two independent reviewers (SB, GB; IG, SG) detected discrepancies for primary and secondary outcomes, outcomes measures, time points, number of arms and statistical analysis plans. We adapted a previously published method[25] to classify discrepancies into: change in definition (e.g., outcome proposed) or measure (e.g., VAS instead of NPRS), addition (e.g., completely outcome measure or arm added, new timepoint added), omission (e.g., excluded primary outcome, excluded arm). In case of switching between primary and secondary outcome we classified it into upgrade (secondary outcome changed to primary) and downgrade (primary outcome changed to secondary).

If no primary outcome was explicitly defined within the manuscript, we considered the outcome used for the power calculation to be the primary published outcome. In case of multiple outcomes/time points are planned in the registered protocol, but not reported in the final publication, we checked related publications referring to the same protocol.

We distinguished between discrepancies reported and not reported in the final publication (i.e., deviation transparently reported in the manuscript), checking the final publication for an explanation

of any deviation from the protocol. If deviations are transparently declared and likely to be justified, we did not consider them as discrepancies.

Before starting the assessment, a calibration phase was performed by the four reviewers (SB, GB, IG, SG) piloting a small sample of 4 RCTs with protocols posted in different registries. Disagreements were discussed during a debrief meeting with another reviewer (GC) to reach a final consensus.

2.5 Comparison between discrepant outcomes and statistically significant results

According to a previous study[30], a discrepancy was considered to favor statistically significant results when: 1) a non-statistically significant (p-value > 0.05 or a confidence interval that crossed zero for continuous outcomes) primary outcome registered in the protocol was downgraded to a secondary in the final publication; 2) a statistically significant secondary outcome registered in the protocol was upgraded to a primary outcome in the final publication; and 3) addition of a non-registered statistically significant primary outcome in the final publication. We prioritized results of between-group comparisons. If more time points are available, we collected any comparison favoring the exercise intervention. If between-groups comparison is not available, we collected within-group results favoring the exercise group versus control.

2.6 Statistical analysis

We used descriptive statistics to assess the prevalence of discrepancies. As a post-hoc analysis, we used logistic regression to assess if some variables (i.e., prospective registration, sample size >100, funding, protocol published) are associated with the presence of any discrepancy in the primary outcome. Statistical significance was set at $P < 0.05$. Data analysis were performed with STATA software.

3. Results

3.1 Study selection

Overall, we included 80 RCTs reporting an available protocol registration. The flow chart of study selection is reported in **Figure 1.**

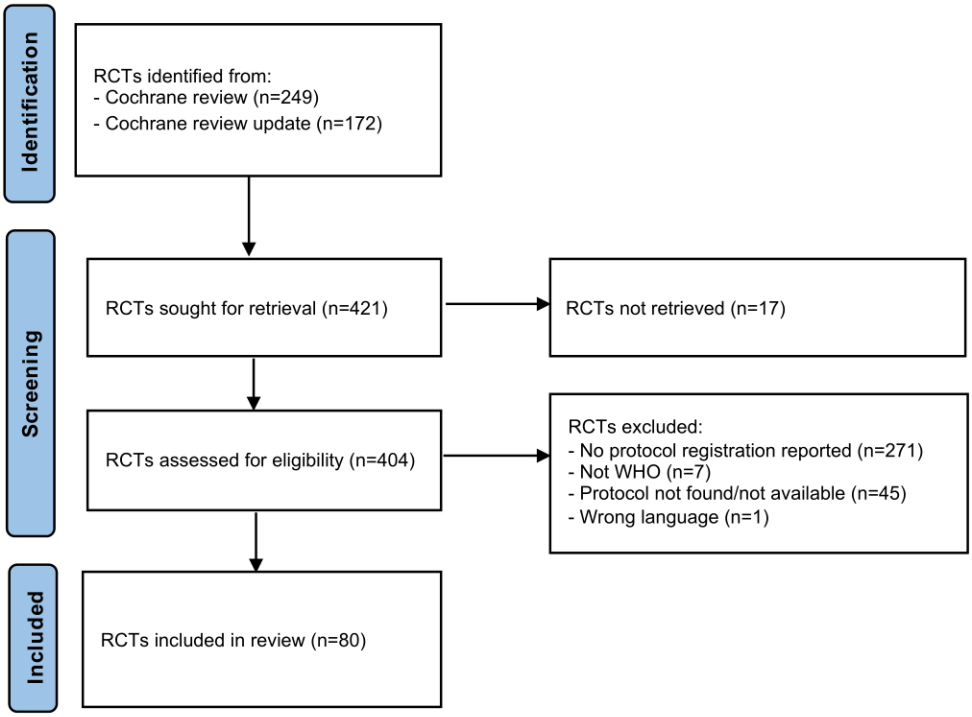

Figure 1. *Flow chart of study selection*

3.2 Final publications general characteristics

Overall, the 80 RCTs included 7427 cLBP patients. The median year of publication was 2018 (IQR 2016-2020). Most RCTs were funded by a non-industry sponsor (n=48; 60%). The median number of arms was 2 (IQR 2-3) assessing different type of exercises versus other exercises or other conservative treatments. The median primary and secondary outcomes reported was 2 (IQR 1-3) and 3 (IQR 2-5), respectively. Fourteen studies reported a published protocol (17.5%). All general characteristics are reported in **Table 1**.

Table 1. *General characteristics of final publications*

	All studies (n=80)	
Number of participants*	67	42 to 119
Age*	44.2	39 to 53
Proportion of female*	63.5%	50% to 76%
Leg pain or neurologic symptoms		
Yes	1	1.2%
Some participants	20	25.0%
No	27	33.8%
NR	32	40%%
Symptom duration		
< 3 years	17	21.3%
> 3 years	25	31.2%
NR	38	47.5%
Funding source reported		
Industry-sponsored	0	0.0%
Non industry sponsored	48	60.0%
No funding	11	13.7%
Not reported	21	26.3%
Outcomes*		
Number of primary outcomes	2	1 to 3
Number of secondary outcomes	3	2 to 5
Arms*		
Number of arms	2	2 to 3
Published protocol	14	17.5%
Ethic commetee yes	75	93.8%

*median and IQR

3.3 Registered protocol general characteristics

Overall, most protocols were registered on ClinicalTrials.gov (n=40; 50,0%) with almost half (n=38; 47,5%) prospectively registered. The remaining 42 retrospective protocols were registered with a median delay of 535 days (IQR 230-1197). The median number of primary and secondary outcomes registered was 2 (IQR 1-3) and 3 (IQR 1-6) for each RCT, respectively. Only two protocols reported statistical analysis plans. General characteristics of all protocols are reported in **Table 2.**

Table 2. *General characteristics of registered protocol*

	Prospective (n=38)		**Retrospective (n=42)**		**Overall (n=80)**	
Registries						
Australian New Zealand Clinical Trials Registry	7	18.4%	2	4.8%	9	11.3%
Brazilian Registry of Clinical Trials - ReBEC	0	0.0%	2	4.8%	2	2.5%
Chinese Clinical Trial Registry	0	0.0%	3	7.1%	3	3.8%
ClinicalTrials gov	19	50.0%	21	50.0%	40	50.0%
Clinical Trials Registry - India (CTRI)	1	2.6%	1	2.4%	2	2.5%
German Clinical Trials Registrer	1	2.6%	2	4.8%	3	3.8%
ISRCTN registry	4	10.5%	8	19.0%	12	15.0%
Iranian Registry of Clinical Trials	4	10.5%	3	7.1%	7	8.8%
Pan African Clinical Trials Registry	1	2.6%	0	0.0%	1	1.3%
Thai Clinical Trials Registry	1	2.6%	0	0.0%	1	1.3%
Statistical analysis plans						
Reported yes	1	2.6%	1	2.4%	2	2.5%
Outcomes*						
Number of primary outcomes	2	1 to 2	1	1 to 3	2	1 to 3
Number of secondary outcomes	4	2 to 6	3	1 to 6	3	1 to 6
Arms*	2	2 to 3	2	2 to 3	2	2 to 3

*median and IQR

3.4 Discrepancies between registered protocol and final publication in outcomes

Overall, we found 142 different discrepancies in reporting outcomes between protocol registrations and publications in 66 RCTs (82.5%): 53 for primary outcomes (n=18 in outcome definitions, n=9 in measurements, n=26 in timepoints) in 37 RCTs (46.3%) and 89 for secondary outcomes (n=55 in outcome definitions, n=12 in measurements, n=22 in timepoints) in 62 RCTs (77.5%). The most frequent type of discrepancy was change of primary outcome timepoint (n=12) or addition (n=18) and omission (n=18) of secondary outcomes (**Table 3**). More than one type of discrepancy was found in primary outcome definition (n=9); in secondary outcome definition (n=16) and secondary outcome time point (n=1).

Table 3. *Discrepancies between final publication and registered protocols*

Discrepancies		**Primary outcome**			**Secondary outcome**		
		Definition	*Measure*	*Time point*	*Definition*	*Measure*	*Time point*
Change in definition	*Change in outcome/measure/time point proposed*	0	4	12	3	9	8
Addition	*Completely new outcome/measure/timepoint added*	3	3	7	18	2	7
Omission	*Completely excluded outcome/measure/timepoint*	3	1	7	18	1	6
Upgrade	*Secondary outcome changed to primary*	2	0	0	0	0	0
Downgrade	*Primary outcome changed to secondary*	1	1	0	0	0	0
Mixed		9	0	0	16	0	1
Na		5	4	8	2	3	8
None		57	67	46	23	65	50
Total discrepancies		18	9	26	55	12	22
% Out of 80 studies		22.5	11.3	32.5	68.8	15	27.5

*sum of change in definition, addition, omission, upgrade, downgrade, mixed.

3.5 Discrepancies between registered protocol and final publication in arms and statistical analyses

Overall, we found five discrepancies (3 omission, 2 addition) in five RCTs in reporting the number of arms between protocol registrations and publications. Discrepancies in statistical analysis were not possible to assess since nearly all registered protocols did not report statistical analysis plans to compare with final publications.

3.6 Relationship between discrepancies in primary outcomes definition and p value

Overall, we found 18 discrepancies in the primary outcome definition. Of these, three were omitted outcomes, therefore we cannot assess the relationship with the p value. Of the remaining 15, 8 (53.3%) favored p-value.

3.7 Association between the characteristics of publications with any kind of discrepancies compared to those without discrepancies in the primary outcome

Table 4 shows that it is more likely that publications without discrepancies in the primary outcome had more than 100 participants. Prospective registration, reporting the source of funding and having a published protocol were not associated with discrepancies.

Table 4. *Publications with discrepancies compared to those without discrepancies in the primary outcome*

Variables	**Discrepancies yes (n=37)**		**Discrepancies no (n=43)**		**OR**
	N°	**%**	**N°**	**%**	
Prospective registration	19	51.4	19	44.2	1.3 (0.6-3.2)
Sample size >100	6	16.2	17	39.5	**0.3 (0.1-0.9)**
Funding reported	24	64.9	35	81.4	0.4 (0.2-1.2)
Protocol published	5	13.5	9	20.9	0.6 (0.2-2.0)

Legend: OR, odds ratio; *statistically significant (p-value < 0.05).

4. Discussion

4.1 Main findings

Our meta-research study aimed to evaluate discrepancies between registered protocol and final publication in primary and secondary outcomes, arms and statistical analysis plans in a sample of 80 RCTs of cLBP registered in ClinicalTrials.gov or in a primary register of the WHO ICTRP. Despite the ICMJE and WHO recommendations[1 3] do not allow period for registration after the recruitment of the first participant, more than half protocols were registered with a median of 535 days later.

Overall, we found discrepancies in primary and/or secondary outcome in the majority of RCTs. Particularly, one out of two of RCTs reported any discrepancy in the primary outcome (including measurements scale and time points) whereas focusing on discrepancies only in outcome definition, 22.5% RCTs upgraded, downgraded, added or omitted a primary outcome, favoring in more than half of cases the p value. Considering that sample size calculation is based on the sample size required to detect an effect on the primary outcome[18], a discrepancy in the primary outcome may lead to overestimate the effects of an intervention or underestimate its undesirable effects, threatening the validity of the study.

We also detected common discrepancies in secondary outcomes in three out of four RCTs, mainly related to the addition of secondary outcomes. Discrepancies in outcome measurements can occur in rehabilitation intervention where multiple and similar questionnaires can be available for the same domain (e.g., RMDQ; ODI), however, investigators should transparently report such deviations from protocols in the final publication. Discrepancies in number of arms (e.g., from three-arms RCT to two-arm RCT) were unusual (6% of the sample) whereas the assessment was not possible in statistical analysis plans since most registered protocols did not report them. However, when researchers do not provide a pre-registered statistical analysis plan, they can analyze data using statistical approaches and models that may favor positive results leading to selective reporting and publication bias.[12, 23 24]

Considering RCT characteristics, prospective registration, reporting the source of funding and having a published protocol did not seem to be associated with the presence of a discrepancy whereas smaller studies with less than 100 participants were more likely to report a discrepancy. This can be due to the fact that low-powered RCTs may fail to achieve statistically significant and relevant clinical effects[17] with more temptation for investigators to favor positive results. As reported by the

CONSORT statement[21], a study should be large enough to have a high probability (power) of detecting as statistically significant a clinically important difference of a given size if such a difference exists.

4.2 Comparison with previous studies

Our findings are consistent with other studies that have explored discrepancies between registry entries and published manuscripts[6, 8, 9, 16, 29, 31] A similar cross-sectional study[30] examining reporting bias in exercise oncology trials (OREO) found evidence of widespread selective outcome reporting and non-reporting bias (outcome switching, omitted preregistered outcomes, and silently introduced novel outcome) in 84% of outcomes. However, their unit of analysis was the number of outcomes whereas our unit of analysis was the type of discrepancy (each RCT categorized by a discrepancies).

4.3 Research implications

Higher quality protocols can enable rigorous trial implementation, enhance the quality and efficiency of protocol review, and reduce the burden of avoidable protocol amendments[7]. Current guidelines exist to prevent outcome discrepancies, such as CONSORT for research reporting[21] and the Standard Protocol Item: Recommendations for Interventional Trials (SPIRIT) statement for protocols of clinical trials[7]. Registries should mandate the prospective registration of primary and secondary outcomes with complete details (e.g., including outcome measurements and time points) as well as statistical analysis plans to make readers ascertaining which outcome and statistically analysis were originally planned.

When assessing the scientific validity of an RCTs, clinicians, peer reviewers and editors should check trial registries to ensure that the outcome of interest is present, described in sufficient detail, and prospectively registered by inspecting the history of revisions record in the clinical trial registration. As well, investigators should transparently report any declaration of a discrepancy compared with registered protocol.

4.4 Strength and Limitations

Our study has several strengths. We registered our planned protocol in a pre-print repository and followed published reporting standard. We determined that our sample would be large enough to provide a precise qualitative summary of the literature and enable us to comment on the generalizability of our findings to the cLBP field.[26] We also report some limitations. Firstly, we judged the prospective and retrospective registration considering the gap of days between the first

submitted date and the study start date. However, we cannot exclude that the study start date could be different than the one registered. Unfortunately, few final publications reported the actual period of recruitment. In addition, even if half protocols were retrospective, nearly all reported an ethics committee approval which has the role to evaluate, approve and monitor clinical trial protocols according to ethics declarations at International level[2]. We could not validate if the trials were registered before any primary outcome data are collected. The start data collection date for primary outcome measure is usually not reported in the trial registration, as it is not a designated field in the major trial registries. Perhaps, to determine if a trial is prospectively registered, it would be more meaningful to consider when the collection of primary outcome data begins. [4] Secondly, some judgements about discrepancies were not possible to be assessed since registered protocols did not report statistical analysis plans and not always a full outcome registration (e.g., outcome measure, time points). Incomplete outcome descriptions give researchers the possibility to selectively choose outcome measurement (e.g. prioritizing a specific subscale) and timepoints. Third, we tried to assess all documents related to the final publication however we cannot be sure that we checked all the related publications (e.g., long term time point). Fourth, registries different than Clinical trial gov did not report some information such as changes between versions. Changes between versions are fundamental for transparency and for assessing the correct version of the protocol, since some discrepancies present in a prospective version could be adjusted in the retrospective version. Lastly, as post-hoc analysis, we compared studies with and without discrepancies to assess if certain characteristics (e.g., prospective registration, number of participants, funding status) could be associated with the presence of a discrepancy; however, we are aware that some of these independent variables can be correlated in the regression model (e.g., prospective registration and protocols).

5. Conclusion

Our findings offer evidence of common outcome discrepancies between registered protocols and final publications threatening the validity of some cLBP exercise RCTs. To improve consistency between trial registry data and publications, we recommend clinicians, researchers, peer reviewers and journal editors to consult the trial registry of the RCT to ensure that the results are consistent and transparently reported. As well, registries should mandate the prospectively and complete registration of outcomes, arms and statistical analysis plans.

REFERENCES

1. International Committee of Medical Journal Editors. Clinical trials registration. http:// www.icmje.org/about-icmje/faqs/clinical-trials-registration/. .
2. National Patient Safety Agency (NHS). The ethical review process for clinical trials in the European Union. https://www.ema.europa.eu/en/documents/presentation/ethical-review-process-perspectives-ethics-committee-david-neal_en.pdf.
3. World Health Organization. WHO Registry criteria https://www.who.int/clinical-trials-registry-platform/network/registry-criteria.
4. 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.
5. Bargeri S, Basso G, Geraci I, et al. Discrepancies between registered protocol and final publication in exercise interventions for chronic low back pain: a meta-research study. *medRxiv*. 2023;2023.2002.2027.23286399.
6. Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med*. 2010;153:158-166.
7. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *Bmj*. 2013;346:e7586.
8. Dwan K, Altman DG, Clarke M, 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.
9. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ*. 2012;344:d7202.
10. 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.
11. 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.
12. 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.
13. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Annals of the rheumatic diseases*. 2014;73:968-974.
14. ICMJE. https://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html.
15. Ioannidis JP. Why Most Clinical Research Is Not Useful. *PLoS Med*. 2016;13:e1002049.
16. Jones CW, Misemer BS, Platts-Mills TF, 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.
17. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. *Emerg Med J*. 2003;20:453-458.
18. Kadam P, Bhalerao S. Sample size calculation. *Int J Ayurveda Res*. 2010;1:55-57.
19. 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.

20. Lawson DO, Puljak L, Pieper D, 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.
21. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
22. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med*. 2017;22:139-142.
23. Page MJ, Higgins JPT, Sterne JAC. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: eds. *Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane, 2022.* 2022:
24. 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;66:524-537.
25. Page MJ, McKenzie JE, Kirkham J, 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;MR000035.
26. Pandis N, Fleming PS, Worthington H, Salanti G. The Quality of the Evidence According to GRADE Is Predominantly Low or Very Low in Oral Health Systematic Reviews. *PLoS One*. 2015;10:e0131644.
27. 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 Based Healthc*. 2019;17:145-146.
28. 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.
29. 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.
30. Singh B, Fairman CM, Christensen JF, et al. Outcome Reporting bias in Exercise Oncology trials (OREO): a cross-sectional study. *medRxiv*. 2021;2021.2003.2012.21253378.
31. Wong EK, Lachance CC, Page MJ, 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.

Appendix A. Protocol

1. Introduction

Randomized controlled trials (RCTs) are the gold standard for assessing the effectiveness of health-related interventions (Guyatt, Haynes et al. 2000). Due to their important effect on clinical practice, there is a necessity to have transparent reporting of RCT results (Ioannidis 2016). Bias in the design, conduct or reporting of RCT can result in inaccuracies in systematic reviews or guidelines and subsequent errors in clinical practice (Hart, Lundh et al. 2012).

Several meta-research studies have been published in the medical field (Bourgeois, Murthy et al. 2010, Hart, Lundh et al. 2012, Dwan, Altman et al. 2014, Jones, Misemer et al. 2018, Serpas, Raghav et al. 2018, Wong, Lachance et al. 2019) showing that discrepancies between what is reported in the registered protocol and what is reported in the final publication are common. This can lead to selective reporting bias and refers to a publication practice where study authors preferentially publish interesting or positive research findings while concealing results that do not confirm their hypothesis because of the statistical significance, magnitude or direction of the effect (Page, Higgins et al. 2022) (Kirkham, Altman et al. 2018) (Page, McKenzie et al. 2013).

In the physiotherapy field, it is still unclear what is the prevalence and the impact of discrepant reporting between the registered protocol and final publication on chronic low back pain (CLBP), which is the main representative of musculoskeletal disorders due to high prevalence and high disease burden (Hoy, March et al. 2014).

Starting from the 2021 Cochrane review publication "Exercise therapy for chronic low back pain" (Hayden, Ellis et al. 2021) on the effectiveness of exercise therapy interventions in adults with CLBP we will aim to assess:

- The proportion of RCTs with a discrepancy between the registered protocol and final publication in objectives, primary and secondary outcomes, outcomes measures, timepoints, number of arms, interventions proposed, and statistical analysis plans
- The characteristics of RCTs with and without discrepancies between the registered protocol and final publication

2. Methods

2.1 Study design

This is a meta-research study (Puljak 2019, Puljak, Makaric et al. 2020). Since the reporting checklist for methods research studies is currently under development (Lawson, Puljak et al. 2020), we will adapt items from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) for reporting meta-research studies.

2.2 Eligibility criteria

We will start from the RCTs included in the 2021 Cochrane review publication "Exercise therapy for chronic low back pain" (Hayden, Ellis et al. 2021) (n=249) and identified in its update (n=172) to select a random sample of 100 RCTs with registered protocol. Hayden et al. included RCTs that compared exercise to no treatment, usual care, placebo or another conservative intervention among adults with CLBP. Trials could include interventions provided to participants in any setting (e.g., healthcare, occupational, general and mixed populations). The intervention could have been with or without the addition of other components (eg, education, manual therapy).

2.3 Registry search

For each RCT we will retrieve information about protocol registration and related publication reported in the Cochrane review (Hayden, Ellis et al. 2021). If no information will be reported, we will check availability of protocols in ClinicalTrials.gov, European Union Clinical Trials Register and International Standard Randomized Controlled Trial Number Register (via the WHO International Clinical Trials Registry Platform). Published article protocols will be reviewed only when registered protocol are missing.

In each database, the search strategy will include keywords such as title, principal investigator, country, exercise therapy name (e.g., Lumbar strengthening, stabilization, Qigong, yoga) and condition name (e.g., CLBP) with manual review of the trial entries. If we will not find a trial in these registries, we will search the national registry in the country where the trial will be conducted. We will also search in supplementary materials of the final publication for any additional information about protocol registration. If neither independent search identified a corresponding registry entry, we will consider the trial unregistered and exclude it.

2.4 Data extraction

Standardized data collection forms will be used to record information from registered protocol and final publication corresponding to each included trial. Two reviewers (SB, GB) will carry out the data extraction phase and the data will be independently checked by another author (IG). Any discrepancies will be resolved by discussion.

For data extraction of the registered protocol, we will collect: the date of initial trial registration, the date on which a primary outcome was first registered, the registered primary outcome(s), registered secondary outcomes and results posted on the registry. We classified the registration status into 1) prospectively registered; 2) retrospectively registered according to the 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.

For data extraction of the final publication, we will use the dataset of the Cochrane review to extract RCTs' general characteristics (e.g., author, year, ID number of the protocol registration, setting, number of arms, sample size, CLBP duration, neurological/radicular symptoms, mean age, sex, description of the intervention, outcomes information, follow up, conflict of interests, funding, journal of publication). In addition, we will extract journal impact factor, effect sizes and direction of the effect. The direction of the effect will be dichotomized as statistically significant or statistically nonsignificant results.

2.5 Detection of discrepancy between registered protocol and final publication

We will define discrepancies as differences between registered protocol and final publication. Discrepancies will be assessed by three reviewers (SB; GB; IG) and compared to the registered protocol for objectives, primary and secondary outcomes, outcomes measures, timepoints, number of arms, interventions proposed, and statistical analysis plans.

Before starting the assessment, a calibration phase will be performed by the three reviewers (SB, GB, IG) piloting a small sample of 10 RCTs and calculating the Fleiss' Kappa for multiple raters. Disagreements will be discussed during a debrief meeting with another reviewer (TI) to reach a final consensus. At the end of the assessment, as a quality assessment measure, 10% of the RCTs will be cross-checked by another reviewer.

We will adapt a previous published method (Page, McKenzie et al. 2014) to classify discrepancies into: change in definition (e.g., intervention proposed) or measure (e.g., VAS instead of NPRS), addition (e.g.., completely new objective, outcome measure or arm added, new timepoint added), omission (e.g., excluded primary outcome, excluded arm, objective omitted). In case of switching between primary and secondary outcome we will classify it into upgrade (secondary outcome changed to primary) and downgrade (primary outcome changed to secondary).

If no primary outcome will be explicitly defined within the manuscript or abstract, we will consider the outcome used for the power calculation to be the primary published outcome. If no primary outcome was defined and there was no power calculation, we will consider the published primary outcome to be undefined.

If inconsistencies will be found between trial registration and the publication for any of the above variables, we will check the final publication for an explanation of any deviation from the protocol. In case of multiple outcomes/time points planned in the registered protocol but not reported in the final publication, we will check related publications referring to the same protocol.

2.6 Statistical analysis

We will use descriptive statistics to assess the proportion of RCTs with and without a discrepancy between the registered protocol and final publication. Data analysis will be performed with STATA software.