

Registration and reporting characteristics of trials investigating exercise therapy following total knee arthroplasty: a systematic review



Birk M GRØNFELDT ^{1,2}, Rasmus S HUSTED ^{1,2}, Rasmus H BRØDSGAARD ^{2,3},
Line HOLST ^{1,2}, Thomas KALLEMOSE ¹, Cathy M CHAPPLE ⁴, Carsten B JUHL ^{5,6},
and Thomas BANDHOLM ^{1,2,7,8}

¹ Department of Clinical Research, Copenhagen University Hospital Amager-Hvidovre, Hvidovre, Denmark; ² Physical Medicine & Rehabilitation Research-Copenhagen, Department of Physical and Occupational Therapy, Copenhagen University Hospital Amager-Hvidovre, Hvidovre, Denmark; ³ Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Aalborg, Denmark; ⁴ Centre for Health, Activity and Rehabilitation Research (CHARR), School of Physiotherapy, University of Otago, Dunedin, Otago, New Zealand; ⁵ Department of Sport Science and Clinical Biomechanics, University of Southern Denmark, Odense; ⁶ Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Herlev & Gentofte, Denmark; ⁷ Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; ⁸ Department of Orthopaedic Surgery, Copenhagen University Hospital Amager-Hvidovre, Hvidovre, Denmark.
ORCIDs, if available, can be found on the article page (<https://www.actaorthop.org/actao/>)
Correspondence: birk.mygind.groenfeldt@regionh.dk
Submitted 2025-12-10. Accepted 2026-04-30.

Background and purpose — Prospectively registering the primary trial outcome is important to reduce selective outcome reporting and increase the trustworthiness of findings, which guide clinical practice. The objectives of our systematic review were to explore and compare the reporting characteristics of prospectively and non-prospectively registered trials investigating exercise therapy following total knee arthroplasty.

Methods — Randomized trials comparing effects of exercise therapy after total knee arthroplasty for osteoarthritis were identified in 4 databases from 2000 to August 12, 2024. One primary outcome per trial was extracted, using a pre-specified hierarchical algorithm, irrespective of outcome domain. Pooled standardized mean differences (SMDs) were calculated on pre-specified outcome domains, and risk-of-bias assessed using the Cochrane Risk-of-Bias tool v2.

Results — 94 trials, comprising 9,396 participants, were included, of which 13 were prospectively registered, 33 retrospectively registered, and 48 unregistered. A single primary outcome was defined in 44% of the 94 trials, and 4 trials reported a primary outcome consistent with a prospective registration. The pooled SMD of primary outcomes was 0.06 (95% confidence interval [CI] −0.03 to 0.16) for prospectively registered trials, 0.67 (CI 0.22–1.11) for retrospectively registered trials, and 0.59 (CI 0.32–0.86) for unregistered trials. Lower risk-of-bias ratings and higher proportions of intention-to-treat adherence, dropout reporting, and adverse event reporting were observed among prospectively registered trials.

Conclusion — Among prospectively registered trials we showed smaller effect size estimates between interventions with lower risk-of-bias ratings, and higher proportions of intention-to-treat adherence, dropout reporting, and adverse event reporting in contrast to trials without prospective registration; furthermore, clear specification of a single primary outcome was uncommon among trials evaluating exercise therapy after total knee arthroplasty.

Numerous recent systematic reviews have evaluated exercise therapy interventions following total knee arthroplasty (TKA) [1–10], yet many of these reviews identify limitations due to trial heterogeneity, particularly in intervention content, outcome selection, and reporting practices [2–4,6,7,9,10].

Randomized controlled trials (RCTs) underpin clinical guideline recommendations because they are considered the gold standard for estimating causal intervention effects when randomization, pre-specified analysis plans, and prospective registration are in place [11–13]. Prospective trial registration involves registering details of a trial's objectives, interventions, and outcomes before enrolment of the first participant [14] and is required by organizations and medical journal editors [15–18]. Prospective registration, including specification of a primary outcome [19], is intended to limit selective outcome reporting and other questionable research practices [20,21]. Despite these requirements for prospective registration [15–17], selective outcome reporting is common [22,23], even in trials

Population	Intervention	Comparator	Outcome	Timeframe
Patients after TKA for knee osteoarthritis	Either exposure or a control group includes exercise therapy, with a different exercise protocol		No limit	No limit

Figure 1. PICOT table. Description of components relating to Population, Intervention, Comparator, Outcome, and Timeframe of interest. TKA: total knee arthroplasty.

published in medical journals considered high-impact [24]. A 2016 meta-epidemiological analysis of 64 Cochrane reviews found unregistered or retrospectively registered trials reporting larger treatment effect estimates than prospectively registered trials [25]. Clinical practice guidelines, and by extension routine treatment decisions, are informed by evidence from RCTs. Hence, a descriptive examination of trial registration characteristics and a quantitative assessment of differences in reported effect sizes by registration status are warranted.

Our review aimed to (i) descriptively explore trial registration characteristics (primary) and (ii) quantitatively assess differences in reported effect sizes (secondary) among prospectively registered and non-prospectively registered trials comparing exercise therapy interventions following TKA.

Methods

Eligibility criteria

We used the PICO(T) framework to define eligibility for the review (Figure 1). Participants were human adults who had undergone TKA followed by an exercise therapy intervention. This may have been the active intervention under investigation, or additional to the active intervention if the effects of exercise were reported separately. Comparators were no exercise, or a type of exercise clearly different from the active group. Any trial outcomes or durations were acceptable for inclusion for the primary aim. Eligible studies were RCTs published after 2000 (launch of clinicaltrials.org) with an accessible full-text publication. Exclusion criteria were studies of TKA for reasons other than osteoarthritis, any RCT with multiple published reports of the same dataset, and any report labelled “secondary analyses” to avoid double-counting participants. Trials were grouped based on whether the registration was prospective, retrospective, or unregistered.

Information sources

Medline, EMBASE, CINAHL, and the Cochrane Central Register of Randomized Controlled Trials (CENTRAL) were last searched on August 12, 2024. Reference lists of included trials and relevant systematic reviews (published within the last 5 years) were checked for eligible studies and forward citation tracking carried out through Web of Science. No data was sought from other sources mentioned in the PRISMA checklist (e.g., registries). Any inaccessible reports were sought from corresponding authors via mail or researchgate.net.

Search strategy

The search strategy targeted trial reports of knee arthroplasty and exercise therapy. The Cochrane highly sensitive search strings (adapted to not target drug therapy or animal studies) were applied to primarily identify RCTs in MEDLINE and EMBASE databases (Cochrane Handbook 6.5 chapter 4.4.7). No limits were applied with built-in filters. Search strings are available in Appendix 2.

Selection process

Before screening, all search results were combined and duplicates removed using the Covidence.org automated duplicate screener and manually by 2 individual reviewers who cross-checked remaining trials as well as group values collected for participant data (BMI, age, n, sex distribution, etc.). Two reviewers independently assessed each record by title and abstract initially, followed by full-text screening. Reviewer disagreements were discussed until consensus. We used Google Translate and DeepL.com to translate trials in languages unfamiliar to reviewers. There was no further use of automation tools, AI, or crowdsourcing for eligibility screening.

Data extraction

Data extraction involved multiple authors, with each study reviewed by a pair of authors independently. Disagreements were discussed until consensus. No missing data was sought from authors of included trials as we wished to synthesize publicly available data. Data points only available from figures were extracted and quantified using plotdigitizer.com.

Data items (outcomes)

Data extracted included: trial design (registration, multicenter, number of study arms), publication details (author, year, location), population (age, sex, BMI, type of surgery), intervention and control details (description, initiation, timing, duration, differentiation), and reporting completeness (sample size calculation, adherence to intention-to-treat principles [ITT], adverse events, dropouts). For each included trial, a single outcome was selected for quantitative synthesis using a prespecified hierarchical decision rule (Figure 2). This was done to prioritize protocol-level registration data, support transparent reporting, and reduce the risk of post hoc outcome selection (see details on changes to the selected outcome in Appendix 9 under 1st amendment). Primary outcome selection was carried out independently by 2 authors. Disagreements were discussed till consensus. Data for primary endpoints was extracted (mean and standard deviation [SD] at baseline and follow-up), in the same way by 2 authors and discussed till consensus. Details of each extracted variable are given in the key table available in the data repository available through the osf.io registration [dataset] [26].

For trials with defined primary outcomes not within the domains of pain, disability, performance-based function, or a combination thereof, the primary outcome score was not

1	The primary outcome reported in the online registration (if only one is reported)
2	The primary outcome reported in the trial report (if only one is reported)
3	The reported outcome with a sample size calculation (if only one is reported)
4	The outcome with the greatest emphasis in the abstract/objectives of the trial
5	The outcome with the greatest emphasis in the results/discussion paragraphs
6	The outcome mentioned first in the result section

Figure 2. Primary outcome hierarchy. Description of the hierarchy used to select a primary outcome for inclusion. Selection would start at level 1 and if conditions were not met then level 2, checking the level below, and so forth. For each included trial, a single outcome was selected for quantitative synthesis using this prespecified hierarchical decision rule. First, the single primary outcome explicitly specified in the trial's online registration was selected. If multiple or no primary outcomes were registered, the single clearly stated primary outcome reported in the trial publication was selected. Again, if multiple or no primary outcomes were stated, the outcome used for sample size calculation was selected, provided it was uniquely identifiable. If not, the outcome most emphasized in the abstract or objectives or, when ambiguity remained, the one most emphasized in the results or discussion sections was selected. Ultimately, when no other criteria were met, the first outcome reported in the results section was selected.

extracted. In studies with multiple outcome assessment time-points (baseline and/or follow-up) or with both intention-to-treat and naive analyses, the one specified for the trial's primary analysis was chosen (if available). When these were not specified after using the hierarchy (see Figure 2), the following extraction rules were used. If not specified, baseline was always the timepoint closest to initiation of intervention, and the follow-up timepoint, if not specified, was always the one closest to the completion of intervention. Intention-to-treat analysis outcomes were prioritized, unless another analysis was clearly specified as the primary outcome. We used conversion methods for trials reporting standard errors (SE) or 95% confidence intervals (CI) with no SDs and small sample size correction for trials with fewer than 60 participants (Chapter 6.5.2.2 in the Cochrane Handbook 6.5) [27].

Study risk of bias assessment

Risk of bias of the primary outcome was assessed for the effect of assignment to the intervention (ITT) using the Cochrane Collaborations Risk of Bias 2 tool (RoB2) [28]. The RoB2 assessment was deemed adequate to assess both baseline and follow-up values (i.e., the comparison between the 2). Each study was assessed independently by 2 assessors and disagreements were discussed until consensus.

Effect measures and covariates

To evaluate the influence of registration status on primary outcomes, continuous data (means and SDs) extracted from available outcome measures was used for statistical analyses. Study characteristics and the RoB2 evaluation are reported according to registration status. Analysis was performed evaluating registration status and the study characteristics listed in Table 1, excluding age, sex, BMI, and geographical location.

Statistics

The statistical analysis plan (SAP) was made publicly available on March 6, 2025 (after latest search, before any analysis) at <https://osf.io/kwqcb/files/osfstorage>. For analysis, the primary outcomes across the following domains were combined: self-reported pain, disability, performance-based function, and composites of these domains (such as total scores from WOMAC or KOOS questionnaires). The primary outcomes were pooled to form the analytical estimand (pooled SMD) for the statistical analyses of the current review, rather than a prespecified clinical outcome domain. The effect of exercise therapy compared with the comparator is expressed as the Hedges' g standardized mean difference (SMD), a bias corrected adjustment of Cohens d, suggested by the Cochrane Collaboration, and analyzed by meta-analysis using random effect modelling (Restricted Maximum Likelihood [REML]).

SMD was applied to enable comparison across trials as the focus was on relative effect sizes rather than clinical interpretation per domain. Meta-regression analysis was used to estimate the effect of registration status on the pooled SMD, as well as the effect of the secondary covariates. Additionally, possible effect modification from the study characteristics on registration status was analyzed by extending the meta-regressions to include an interaction term with registration status. All models were adjusted for baseline estimates. All extracted outcomes were "score inverted/normalized" so higher scores indicated improvement/benefit). The overall pooled effects (SMD) were presented so positive outcomes are in favor of the intervention group. Differences in effect between studies were expressed as inconsistency (I^2) as proposed by Higgins et al. (i.e., percentage of total variation across studies due to heterogeneity, rather than chance) [29]. The magnitudes of the effect sizes were interpreted using the guidelines from Cohen (0.2 = small, 0.5 = moderate, and 0.8 = large) [30]. The impact of the secondary covariates in the meta regression was evaluated by their ability to reduce the between study variance, rather than chance, expressed as τ^2 . All P values from analysis of the secondary variables were presented with correction for multiple testing by Bonferroni correction, with P values multiplied by the number of secondary variables. No missing data was imputed.

Sensitivity analysis

The ICJME formulated specific requirements for prospective registration of trials for medical journals in 2004 [17]. As it might not be reasonable to include studies before this requirement was published, a sensitivity analysis (not pre-specified in the protocol) was performed, using only studies after July 1, 2005 (end of the ICMJE grace period). All sensitivity analyses were performed as described in the statistics section.

Table 1. Trial reporting characteristics

Item	Prospective (n = 13)	Retrospective (n = 33)	Unregistered (n = 48)	Total (n = 94)
Demographics				
Age, median (IQR)	66.4 (64.7–68.0)	68.4 (66.9–70.9)	68.7 (67.3–69.9)	68.5 (66.6–70.0)
Sex, n (%)				
Female	1,348 (59)	1,744 (62)	3,246 (76)	6,378 (68)
Male	932 (41)	1,050 (38)	1,036 (24)	3,018 (32)
Body mass index, median (IQR)	32.1 (30.9–33.4)	29.9 (28.4–31.2)	30.1 (28.1–31.5)	30.3 (28.4–31.8)
Participants in trial, median (IQR) (range)	105 (60–334) (30–422)	60 (40–108) (12–290)	49 (35–91) (10–345)	60 (38.5–109.5) (10–422)
Geographic location, n by country	AU = 5 TW, UK = 2 FI, KR, US, DK = 1	US = 6 AU = 4 FI = 3 KR, CA, TR, DE, CN, JP, NO, ES, DK = 2 IT, GR = 1	US = 7 DE = 7 TR = 5 KR, TW, CA = 3 CN, JP, PK, AT, IR, FR = 2 IT, GR, PL, BR, UK, NL, AU, FI = 1	US = 14 AU = 10 DE = 9 TR = 7 KR = 6 TW, FI, CA = 5 CN, JP = 4 UK, DK = 3 PK, AT, IR, FR, IT, NO, ES, GR = 2 PL, BR, NL = 1
Design				
Multicenter study (Y/N), n	7/6	11/22	5/43	23/71
Single primary outcome definition (Y/N), n	8/5	18/15	15/33	41/53
Study arms (4/3/2 trial arms), n	0/2/11	1/4/28	0/4/44	1/10/83
ROB 2 assessment summary, n				
1. Bias arising from the randomization process (high/some concerns/low)	1/1/11	1/8/24	9/14/25	11/23/60
2. Bias due to deviations from intended interventions (high/some concerns/low)	1/5/7	11/16/6	23/21/4	35/42/17
3. Bias due to missing outcome data (high/some concerns/low)	2/2/9	6/5/22	9/16/23	17/23/54
4. Bias in measurement of the outcome (high/some concerns/low)	0/4/9	7/12/14	18/18/12	25/34/35
5. Bias in selection of the reported result (high/some concerns/low)	4/4/5	14/17/2	23/25/0	41/46/7
Overall bias (high/some concerns/low)	5/5/3	22/10/1	39/9/0	66/24/4
Reporting				
Primary outcome extracted from, n				
1: Online registration	7	13	0	20
2: Trial report	3	6	11	20
3: Sample size calculation	2	5	7	14
4: Emphasis in abstract/objectives	0	3	10	13
5: Emphasis in results/discussion	0	0	4	4
6: First outcome reported in results	1	6	16	23
Sample size calculations reported (Y/N) ^b , n	12/1	28/5	23/25	63/31
Dropouts reporting (yes/partial/no), n	13/0/0	27/2/4	34/8/6	74/10/10
Adverse events reporting (yes/partial/no), n	8/0/5	17/3/13	22/2/24	47/5/42
Data analysis adherent to intention-to-treat, n principles (yes/no/unclear)	8/5/0	10/14/9	10/31/7	28/50/16
Interventions				
Intervention initiation timing, median (IQR) in days after surgery	21 (3.5–42.1)	11 (3.4–40.5) 1 NA	7.2 (1.0–16.5) 8 unclear	8.0 (2.0–31.5) 1 NA, 8 unclear
“Inactive” comparator, n	2	1	5	8
Baseline measurement timing, n (preop./postop./unclear)	5/6/2	16/14/3	21/26/1	42/46/6
Follow-up timepoint, days, median (IQR)	180 (104.5–365)	70 (38.5–105)	60 (28–91)	70 (33.8–127.5)
Days from surgery (extracted outcome) ^a	1 NA	2 NA	3 NA	6 NA

Presentation of trial population, design, risk-of-bias, and intervention characteristics. Intention-to-treat definition: all randomized participants clearly included in assigned groups for analysis. NA = not available

^a Extracted either at primary outcome timepoint if specified, otherwise selected directly following end of planned exercise therapy intervention.

^b Post-hoc calculations are counted as sample size not being reported

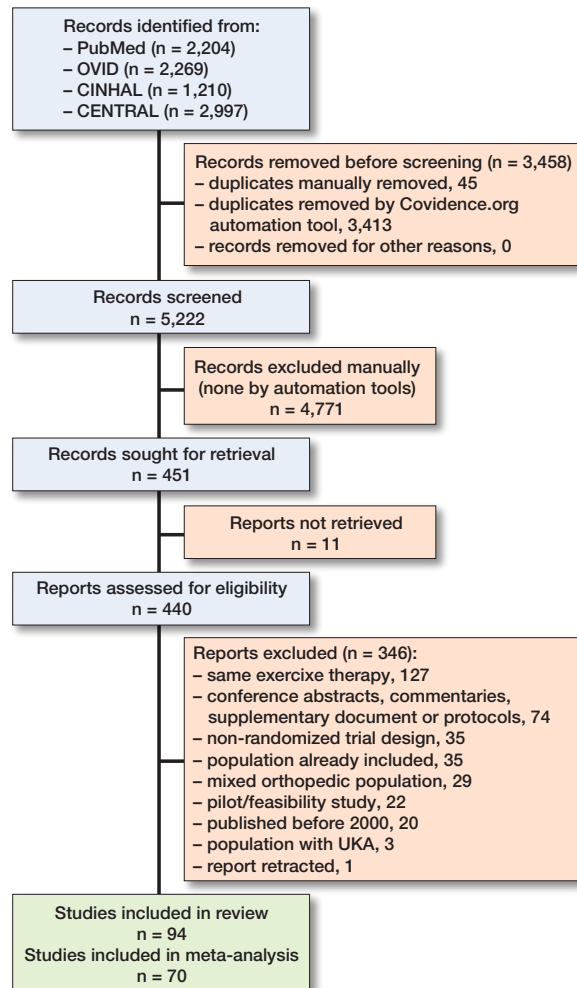


Figure 3. PRISMA flow diagram. Presentation of the process of identifying, screening, and including trials. Reports excluded are cited in Appendix 3

Ethics, registrations, data sharing plan, use of AI, funding, and disclosures

Ethical approval is not required for this study, as it is based solely on analyses of publicly available, aggregated data from trials and registrations. This prospectively registered systematic review (October 25, 2019, osf.io/6wzre) is reported in accordance with PRISMA 2020 [31] (Appendix 1). The protocol (October 4, 2019), Statistical Analysis Plan (March 6, 2025), data, and code, are publicly available: <https://osf.io/kwqcb/files/osfstorage> [dataset] [26] under the license CC-BY 4.0 International. None of the authors declare any financial or personal conflicts of interest. This work was not commissioned; no funding was received specifically for its conduct. AI was used for grammar and spelling and the output was controlled by the main author. Complete disclosure of interest forms according to ICMJE are available on the article page, doi: 10.2340/17453674.2026.46047

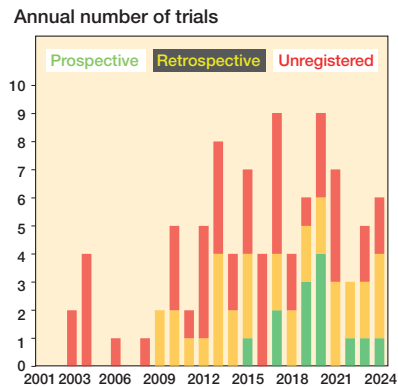


Figure 4. Distribution of included trials over time. X-axis = publication year, Y-axis = number of published trials.

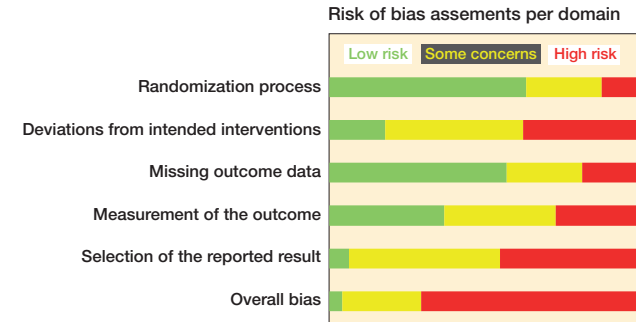


Figure 5. Risk of bias domains. Summary of risk of bias assessments for the included primary outcome using the Cochrane Risk of Bias 2 (RoB2) tool. N = 94.

Results

A total of 8,680 records were identified through database searches (Figure 3). After duplicates were removed, 5,222 records were screened, 440 full-text documents assessed, and 94 (13 prospective, 33 retrospective, 48 unregistered) reports were included for the main analysis (Table 1), comprising 9,396 participants (Figure 4—references in Appendix 10). Data from 70 (9 prospective, 25 retrospective, 36 unregistered) reports were included in secondary analyses (forest plot and meta-regression analyses—Figure 6 and Appendices 6 and 7). Reasons for exclusion by full-text screening were most often that interventions compared the same exercise therapy, or were conference abstracts, commentaries, etc. (i.e., not full trial reports) (Appendix 3). No additional reports were found by searching references or cited-by lists.

Trial design differences

Included trials with prospective registration presented with a higher median of participants, longer time until follow-up (for the extracted primary outcome), and more frequently used multicenter designs than non-prospectively registered trials

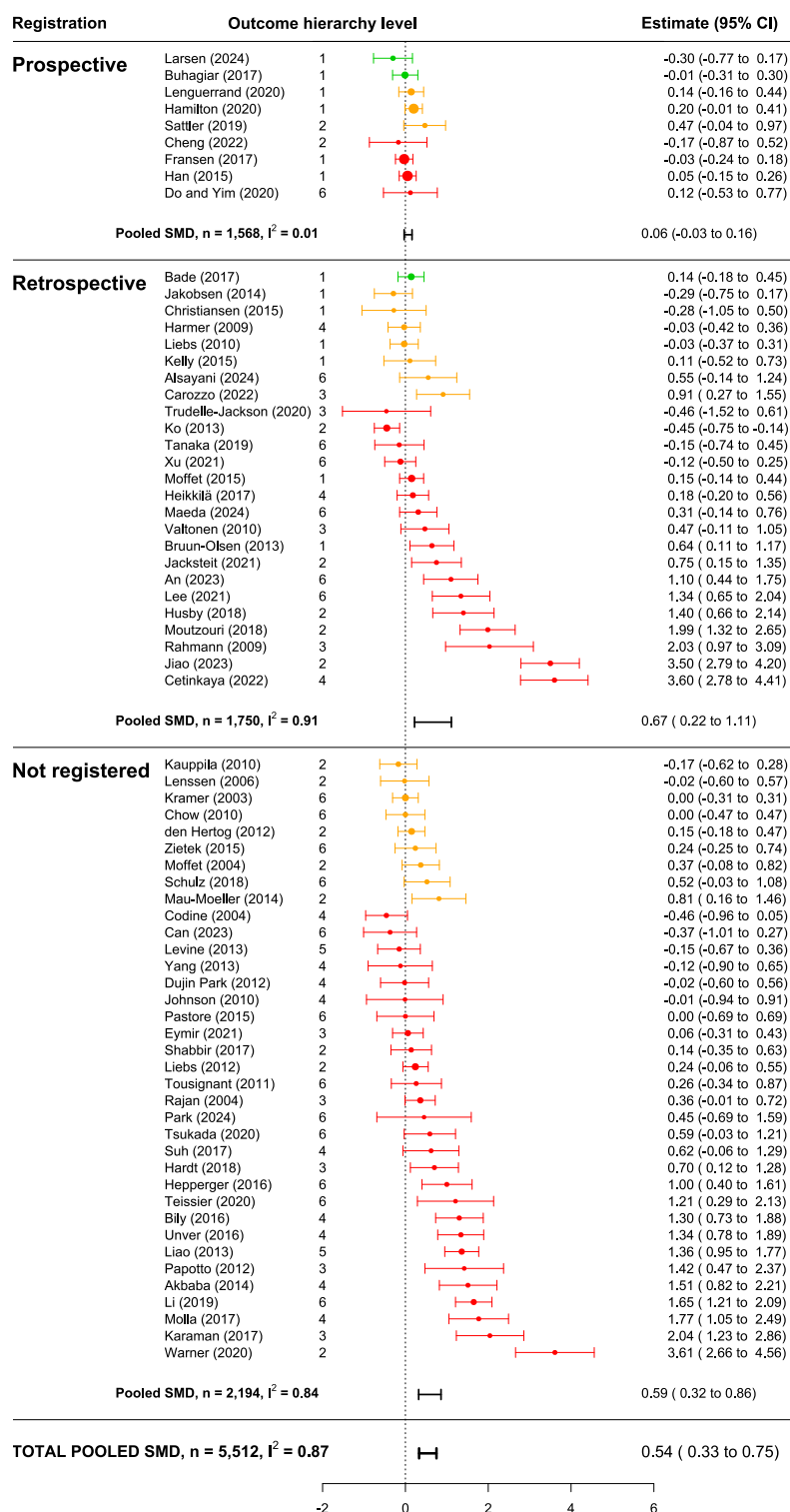


Figure 6. Forest plot. Presentation of pooled standardized mean differences of prospective ($n = 9$), retrospective ($n = 25$), and unregistered ($n = 36$) RCTs with extractable trial data within the domains of pain, disability, performance-based function, or a combination of those 3. A standardized mean difference favoring the exposure group (the group with more exercise) is represented with values above 0, values below 0 favor the control group (the group with less exercise). Green represents outcomes judged at low risk of bias, yellow outcomes judged at some concerns for bias, and red outcomes judged at high risk of bias. I^2 represents total variation across studies due to heterogeneity, rather than chance.

(see Table 1). For a more detailed overview of each trial, see the raw data extraction sheet and risk of bias assessments available in the data repository available through the registration [dataset] [26].

A brief description of the intervention and comparator groups, the selected outcome, and risk of bias domain judgments is available for all 94 included trials in Appendix 4.

Trial reporting differences

Across reporting domains (see Table 1), the proportion of trials meeting reporting criteria was consistently highest among prospectively registered trials, intermediate among retrospectively registered trials, and lowest among unregistered trials. This pattern was observed for: definition of a single primary outcome (62%/55%/31%), reporting of sample size calculations (92%/85%/48%), dropout reporting (100%/82%/71%), adverse event reporting (62%/52%/46%), and adherence to intention-to-treat principles (62%/30%/21%) (numerators and denominators available in Table 1).

In the 13 prospectively registered trials, a single primary outcome consistent with a single prospectively registered, unaltered, primary outcome was reported in 4 of the 13 trials. The primary outcome was changed or switched in 2 of the 13 trials. Multiple primary outcomes were registered and/or reported in 7 of the 13 trials (see Appendix 10 for references).

Risk of bias in trials

Outcomes in 4 of 94 (4.3%) trials were judged at low risk of bias, 24 of 94 (26%) at some concern of bias, and 66 of 94 (70%) at high risk of bias (Figure 5). Prospectively registered trials represented up to 3 of the 4 trial outcomes judged at low risk of bias, 5 of the 24 (21%) at some concerns for risk bias, and 5 of the 66 (7.6%) at high risk of bias. An overview of the risk of bias is available in Table 1, and an extended risk of bias table presenting experimental and control exercise therapy interventions is available in Appendix 4. The complete raw assessments are available in the data repository [dataset] [26].

Study effect differences

The pooled SMD of prospectively registered trials was 0.06 (CI -0.03 to 0.16), not favoring either group, whereas the pooled SMD of retrospectively trials was 0.67 (CI 0.22–1.11) and unregistered trials 0.59 (CI 0.32–0.86), both favoring the intervention group (Figure 6). Very low I^2 heterogeneity was present between prospectively

registered trials, and high I^2 heterogeneity presented for the retrospectively registered and unregistered trials. Funnel plots are available in Appendix 5.

The meta-regression analysis (Appendix 6) of the included trials observed the following differences: in retrospectively and unregistered trials the pooled SMD was 0.53 (CI –0.03 to 1.11) and 0.37 (CI –0.19 to 0.92) larger than for the prospectively registered trials. The covariates: outcome hierarchy levels, risk of bias, ITT adherence, dropout reporting, and primary outcome definition showed some association with the reported effect size estimate and explained notable between-study variance (τ^2). None of the differences for the covariates, except ITT adherence, were significant and generally presented large confidence intervals.

The interaction meta-regression analysis (Appendix 7) showed some estimates with interaction effects, those most pronounced being: description of trial dropouts, primary outcome hierarchy level, and adherence to ITT principles. However, none of the interactions were statistically significant after correction for multiple testing, and confidence intervals were generally wide, likely due to few observations being available within some levels.

Sensitivity analysis

Excluding trials not required to be prospectively registered, according to the July 1, 2005 ICMJE grace period, did not change the interpretation of outcomes (see Appendix 8). Only minor changes were observed: compared with registered trials, the SMD estimate for non-registered trials favored the intervention group slightly more, with slightly larger confidence intervals and τ^2 values. Furthermore, the Eggers test in the funnel plot for non-registered trials changed from being statistically significant to showing weak/borderline evidence of asymmetry.

Discussion

We aimed to explore and compare the reporting characteristics of prospectively and non-prospectively registered trials investigating exercise therapy following total knee arthroplasty and observed that, compared with non-prospectively registered trials, the outcomes of prospectively registered trials had smaller effect size differences between interventions with overall lower risk-of-bias ratings, and higher proportions of intention-to-treat adherence, dropout reporting, and adverse event reporting.

The core observation, that non-prospectively registered trials were associated with larger effect sizes than prospectively registered ones, mirrors the finding of the previously referred meta-epidemiological study of Cochrane reviews [25]. Fewer than half of the included trials clearly defined a single primary outcome, and among the 13 trials with pre-registered outcomes, but 4 of the 13 consistently reported the same preregistered

single primary outcome in the final report (Appendix 10). Riley et al. found 49 of 138 (36%) trials on musculoskeletal physical therapy were prospectively registered [32], compared with 13 of 94 (14%) in the current review; this is likely a result of older trials being included in the current review. Riley et al. also found that among prospectively registered trials, 42 of 49 (86%) reported outcomes consistent with registered intent, but post-randomization bias could not be ruled out for the remaining 96 trials [32]. Changing the pre-registered primary outcome after a trial begins is not recommendable, but can happen. It is less of concern if it happens shortly after the trial begins, when little data has been collected, and no analyses could be performed—provided the change is transparently reported. This was the reason why 1 retrospectively registered trial was considered at low risk of bias, as seen in Figure 6. Many of the included trials listed multiple “primary” outcomes without the accompanying methodological specifications, which are often recommended when more than 1 primary endpoint is intended (e.g., an explicit success criterion such as co-primary vs at least 1 prespecified multiplicity control, and sample-size planning aligned with that choice) [33]. Our observation is not intended as a critique of legitimate co-primary or dual-primary frameworks per se, but rather of unclear pre-specification in trials that have labelled several outcomes as “primary.” Such practice may complicate consistent outcome reporting and interpretation and could be associated with variable or larger effect estimates when combined with retrospective or absent registration. Considering the differences in effect sizes between prospectively registered and non-prospectively registered trials in the current sample (see Figure 6), and that just 3 of 94 trials reported a statistical analysis plan, post-randomization bias cannot be excluded in the current sample of RCTs comparing exercise therapy interventions after total knee replacement. While the current review illustrates how timely outcome registration enables readers and peer-reviewers to assess coherence between registration and reporting, timely registration alone does not necessarily ensure higher methodological quality, as that also depends on adherence to other core principles of RCT design and conduct.

A 2023 synthesis of meta-epidemiological studies of rehabilitation trials concluded that inadequate or unclear randomization sequence generation and allocation concealment were likely to produce exaggerated effect estimates for continuous outcomes, and higher risk of reporting bias appeared to be associated with an overestimation of effects [34]. In the current review, most of the extracted outcomes were rated at low risk of bias in the “randomization process” domain, indicating that the differences in effect size estimates were not clearly attributable to randomization and allocation concealment biases. In contrast, the “reporting bias” domain presented with the highest proportion of outcomes at some concern or high risk of bias in retrospectively registered and unregistered trials. This pattern may be consistent with an association between higher risk of reporting bias and larger pooled SMD estimates in those trials.

Limitations

The statistical analysis plan was registered after completion of data extraction but prior to any analysis. Owing to the approach chosen to select the primary outcome from each trial, the pooled estimates presented in the current meta-analyses do not represent the overall clinical effect of exercise therapy after TKA. Rather, they reflect pooled standardized mean differences based on trial-specific primary outcomes selected using a prespecified hierarchy, across different outcome domains. Consequently, the analyses aggregate outcomes representing different constructs of pain, disability, and performance-based function. While standardized mean differences are typically recommended for combining results within the same construct (Cochrane handbook 6.5, section 15.5.3), they were used here to facilitate comparison of effect size estimates across trials with differing primary outcomes. Accordingly, the analyses were intended to examine whether effect size estimates varied by trial characteristics, rather than to inform clinical effectiveness.

Our review findings show that non-prospectively registered RCTs of exercise therapy interventions after TKA tend to report larger effect size estimates compared with prospectively registered trials. These observations warrant cautious interpretation, as differences in effect size estimates, related to registration practices, may influence evidence synthesis used to inform clinical practice guidelines if not appropriately considered.

Conclusion

We found that prospective trial registration and clear specification of a single primary outcome were uncommon among trials evaluating exercise therapy after TKA. Prospectively registered trials were observed to differ from retrospectively registered and unregistered trials with respect to larger sample sizes, clearer reporting practices, lower risk-of-bias ratings, and smaller effect size differences between interventions.

Supplementary data

Appendices 1–10 are available as supplementary data on the article page, doi: 10.2340/17453674.2026.46047

Conception and design: RH, CC, TB, BG, CJ. Trial selection: BG, LH, RB, TB. Collection and assembly of data: BG, TB, CJ, RH, LH, CC, RB. Analysis and interpretation of the data: BG, TB, CJ, TK. Statistical expertise: TK, CJ. Drafting of the article: BG, TB. Critical revision of the article for important intellectual content: TB, CJ, TK, RH, LH, CC, RB. All authors approved the final decision to submit the manuscript for publication.

Handling co-editors: Bart Pijls and Robin Christensen

Acta thanks Justine M Naylor and Henricus Vermeulen for help with peer review of this manuscript.

1. Artz N, Elvers K T, Lowe C M, Sackley C, Jepson P, Beswick A D. Effectiveness of physiotherapy exercise following total knee replacement: systematic review and meta-analysis. *BMC Musculoskelet Disord* 2015; 16(1): 15. doi: 10.1186/s12891-015-0469-6.
2. Skoffler B, Dalgas U, Mechlenburg I. Progressive resistance training before and after total hip and knee arthroplasty: a systematic review. *Clin Rehabil* 2015; 29(1): 14–29. doi: 10.1177/0269215514537093.
3. Florez-García M, García-Pérez F, Curbelo R, Pérez-Porta I, Nishishinya B, Rosario Lozano MP, et al. Efficacy and safety of home-based exercises versus individualized supervised outpatient physical therapy programs after total knee arthroplasty: a systematic review and meta-analysis. *Knee Surg Sports Traumatol Arthrosc* 2017; 25(11): 3340–53. doi: 10.1007/s00167-016-4231-x.
4. Sattler L N, Hing W A, Vertullo C J. What is the evidence to support early supervised exercise therapy after primary total knee replacement? A systematic review and meta-analysis. *BMC Musculoskelet Disord* 2019; 20(1): 1–11. doi: 10.1186/s12891-019-2415-5.
5. Buhagiar M A, Naylor J M, Harris I A, Xuan W, Adie S, Lewin A. Assessment of outcomes of inpatient or clinic-based vs home-based rehabilitation after total knee arthroplasty: a systematic review and meta-analysis. *JAMA Netw Open* 2019; 2(4): e192810–e192810. doi: 10.1001/jamanetworkopen.2019.2810.
6. Mark-Christensen T, Juhl C, Thorborg K, Bandholm T. Is physical rehabilitation superior to no physical rehabilitation following total knee arthroplasty? A systematic review and meta-analysis. *medRxiv* 2020; 2020.04.29.20084392. doi: 10.1101/2020.04.29.20084392.
7. Alrawashdeh W, Eschweiler J, Migliorini F, El Mansy Y, Tingart M, Rath B. Effectiveness of total knee arthroplasty rehabilitation programmes: a systematic review and meta-analysis. *J Rehabil Med* 2021; 53(6): jrm00200. doi: 10.2340/16501977-2827 PubMed PMID: 33846757.
8. Chaudhry Y P, Hayes H, Wells Z, Papadelis E, Khanuja H S, Deirmengian C. Not all patients need supervised physical therapy after primary total knee arthroplasty: a systematic review and meta-analysis. *Cureus* 2023; 15(2): e35232. doi: 10.7759/cureus.35232 PubMed PMID: 36968854; PubMed Central PMCID: PMC10032555.
9. Konnyu K J, Thoma L M, Cao W, Aaron R K, Panagiotou O A, Bhumra M R, et al. Rehabilitation for total knee arthroplasty: a systematic review. *Am J Phys Med Rehabil* 2023; 102(1): 19–33. doi: 10.1097/PHM.0000000000002008 PubMed PMID: 35302953.
10. Koster A, Stevens M, van Keeken H, Westerveld S, Seeber G H. Effectiveness and therapeutic validity of physiotherapeutic exercise starting within one year following total and unicompartmental knee arthroplasty for osteoarthritis: a systematic review. *Eur Rev Aging Phys Act* 2023; 20(1): 8. doi: 10.1186/s11556-023-00317-4.
11. Frieden T R. Evidence for health decision making—beyond randomized, controlled trials. *New Eng J Med* 2017; 377(5): 465–75. doi: 10.1056/NEJMr1614394.
12. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med* 2018; 210: 2–21. doi: 10.1016/j.socscimed.2017.12.005.
13. Reito A, Overgaard S. Editorial: Why the randomized controlled trial is still at the apex and the gold standard for evaluating new medical and surgical interventions. *Acta Orthop* 2025; 96: 873–4. doi: 10.2340/17453674.2025.44947.
14. Bandholm T, Christensen R, Thorborg K, Treweek S, Henriksen M. Preparing for what the reporting checklists will not tell you: the PREPARE Trial guide for planning clinical research to avoid research waste. *B J Sports Med* 2017; 51(20): 1494–501. doi: 10.1136/bjsports-2017-097527.
15. World Medical Association. WMA Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. [Internet] 2017. Available from: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> [accessed November 15, 2017].
16. WHO. Trial registration [Internet]. [cited 2025 May 15]. Available from: <https://www.who.int/tools/clinical-trials-registry-platform/network/trial-registration>
17. ICMJE | About ICMJE | Clinical Trials Registration [Internet]. [cited 2025 May 15]. Available from: <https://www.icmje.org/about-icmje/faqs/clinical-trials-registration/>

18. **Christensen R, Ranstam J, Overgaard S, Wagner P.** Guidelines for a structured manuscript: statistical methods and reporting in biomedical research journals. *Acta Orthop* 2023; 94: 243-9. doi: 10.2340/17453674.2023.11656.
19. **Andrade C.** The primary outcome measure and its importance in clinical trials. *J Clin Psychiatry* 2015; 76(10). doi: 10.4088/JCP.15f10377.
20. **Andrade C.** HARKing, cherry-picking, P-hacking, fishing expeditions, and data dredging and mining as questionable research practices. *J Clin Psychiatry* 2021; 82(1): 20f13804. doi: 10.4088/JCP.20f13804.
21. **Cook C, Garcia A N.** Post-randomization bias. *J Man Manip Ther* 2020; 28(2): 69-71. doi: 10.1080/10669817.2020.1739153.
22. **Dwan K, Altman D G, Arnaiz J A, Bloom J, Chan A W, Cronin E, et al.** Systematic review of the empirical evidence of study publication bias and outcome reporting bias. Siegfried N, editor. *PLoS ONE* 2008; 3(8): e3081. doi: 10.1371/journal.pone.0003081.
23. **Dwan K, Altman D G, Clarke M, Gamble C, Higgins J P T, Sterne J A C, 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(6): e1001666. doi: 10.1371/journal.pmed.1001666.
24. **Goldacre B, Drysdale H, Dale A, Milosevic I, Slade E, Hartley P, et al.** COMPare: a prospective cohort study correcting and monitoring 58 mis-reported trials in real time. *Trials* 2019; 20(1): 118. doi: 10.1186/s13063-019-3173-2.
25. **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): 100. doi: 10.1186/s12916-016-0639-x.
26. **Grønfeldt B, Husted R, Brødsgaard R, Holst L, Chapple C, Kallelemose T, et al.** Trial registration and reporting characteristics of trials investigating exercise-based rehabilitation following total knee arthroplasty: a systematic review [Internet]. OSF; 2019 [cited 2025 Jun 26]. Available from: <https://osf.io/kwqcb/> doi: 10.17605/OSF.IO/KWQCB.
27. **Higgins J, Li T, Deeks J.** Chapter 6: Choosing effect measures and computing estimates of effect [last updated August 2023]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page M J, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5. Cochrane, 2024. [Internet]. [cited 2025 Jun 5]. Available from: [cochrane.org/handbook](https://www.cochrane.org/handbook).
28. **Sterne J A C, Savović J, Page M J, Elbers R G, Blencowe N S, Boutron I, et al.** RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; 14898. doi: 10.1136/bmj.14898.
29. **Higgins J P T.** Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414): 557-60. doi: 10.1136/bmj.327.7414.557.
30. **Cohen J.** *Statistical Power Analysis for the Behavioral Sciences* 2. Hillsdale, NJ: Lawrence Erlbaum; 1988. 579 p.
31. **Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D, et al.** The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372:n71. doi: 10.1136/bmj.n71.
32. **Riley S P, Swanson B T, Shaffer S M, Sawyer S F, Cleland J A.** The unknown prevalence of postrandomization bias in 15 physical therapy journals: a methods review. *J Orthop Sports Phys Ther* 2021; 51(11): 542_50. doi: 10.2519/jospt.2021.10491.
33. **Großhennig A, Thomas N H, Brannath W, Koch A.** How to avoid concerns with the interpretation of two primary endpoints if significant superiority in one is sufficient for formal proof of efficacy. *Pharm Stat* 2023; 22(5): 836-45. doi: 10.1002/pst.2314.
34. **Arienti C, Armijo-Olivo S, Ferriero G, Feys P, Hoogeboom T, Kiekens C, et al.** The influence of bias in randomized controlled trials on rehabilitation intervention effect estimates: what we have learned from meta-epidemiological studies. *Eur J Phys Rehabil Med* 2023; 60(1): 135-44. doi: 10.23736/S1973-9087.23.08310-7.