What are the unsupervised exercise adherence rates in clinical trials for knee osteoarthritis? A systematic review

Kristin M. Smith a, *, B. James Massey a, b, Jodi L. Young a, Daniel I. Rhon a, c

a Science Program in Physical Therapy, Bellin College, Green Bay, WI, USA
b Department of Physical Therapy, Wingate University, Wingate, NC, USA
c Department of Rehabilitation Medicine, School of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

Received 12 June 2022; received in revised form 11 July 2023; accepted 4 August 2023
Available online 12 August 2023

Abstract

Background: Exercise is an effective intervention for knee osteoarthritis (OA), and unsupervised exercise programs should be a common adjunct to most treatments. However, it is unknown if current clinical trials are capturing information regarding adherence.

Objective: To summarize the extent and quality of reporting of unsupervised exercise adherence in clinical trials for knee OA.

Methods: Reviewers searched five databases (PubMed, CINAHL, Medline (OVID), EMBASE and Cochrane). Randomized controlled trials where participants with knee OA engaged in an unsupervised exercise program were included. The extent to which exercise adherence was monitored and reported was assessed and findings were subgrouped according to method for tracking adherence. The types of adherence measurement categories were synthesized. A quality assessment was completed using the Physiotherapy Evidence Database (PEDro) scores.

Results: Of 3622 abstracts screened, 176 studies met criteria for inclusion. PEDro scores for study quality ranged from two to ten (mean=6.3). Exercise adherence data was reported in 72 (40.9%) studies. Twenty-six (14.8%) studies only mentioned collection of adherence. Adherence rates ranged from 3.7 to 100% in trials that reported adherence. For 18 studies (10.2%) that tracked acceptable adherence, there was no clear superiority in treatment effect based on adherence rates.

Conclusions: Clinical trials for knee OA do not consistently collect or report adherence with unsupervised exercise programs. Slightly more than half of the studies reported collecting adherence data while only 40.9% reported findings with substantial heterogeneity in tracking methodology. The clinical relevance of these programs cannot be properly contextualized without this information.

Keywords

Adherence; Compliance; Exercise therapy; Home exercise; Knee osteoarthritis; Unsupervised exercise

* Corresponding author at: 111 Raley Blvd Ste 140, Chico, CA 95928, USA.
E-mail: kristin.smith@alumni.bellincollege.edu (K.M. Smith).

https://doi.org/10.1016/j.bjpt.2023.100533
1413-3555/© 2023 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier España, S.L.U. All rights reserved.
Introduction
Knee osteoarthritis (OA) is a common condition responsible for high healthcare use and expenditures. Upwards of 8.6% of lifetime medical costs can be attributed to knee OA ($19,600 ± 16,200 in 2013 US dollars).

Another average of $7104 (in 2007 US dollars) are lost in work productivity by employed individuals undergoing total knee arthroplasty.

Clinical guidelines recommend exercise as a core treatment intervention, emphasizing the importance of exercise prescription and use of strength training in both land and water-based settings as appropriate.

In a survey of 1064 Australian physical therapists, 99% of respondents reported the provision of strengthening exercises when serving patients with knee OA.

A recent meta-analysis concluded that no new research on the benefits of exercise for knee OA is needed due to the robustness of the current body of literature establishing its effectiveness in pain reduction and increase in function.

Exercise is a common intervention in clinical trials for knee OA, used in both supervised and unsupervised formats.

Adherence is commonly described in the medical literature as a proactive behavior such as completing unsupervised exercise; this differs from compliance which is considered a more passive behavior. Adherence plays an important role in the effectiveness of exercise.

One study of overweight patients with knee OA found that increased exercise adherence was associated with improved walking distance and disability levels. Because adherence could influence treatment effect, it is important to understand how adherence is measured in trials that prescribe unsupervised exercise programs. These rates have been studied for other diagnoses. A recent systematic review of rotator cuff related shoulder pain found adherence rates to unsupervised exercise to be between 20 and 100%.

When adherence to exercise programs designed to reduce falls was greater than 80%, the association with risk reduction for falls was highest.

There are published adherence rates for combined unsupervised and supervised exercise in individuals with spondyloarthritis (51–95%) and advanced cancer (44–95%), as well as those awaiting organ transplants (82.5–100%).

Additionally, the known adherence rates for individuals with intermittent claudication for walking is 77.6% and “alternative” exercise is 85.8%. Each of these studies used the term adherence, but included both unsupervised and supervised exercise. However, these rates are unknown in patients with knee OA, where exercise is a recommended core treatment that should be used often.

Currently, no consensus exists on optimal or standardized methods for reporting exercise adherence in clinical trials. There is substantial variation in the manner in which adherence to exercise interventions is identified, reported, and measured in clinical research.

In a systematic review that investigated measures of adherence to unsupervised exercise programs among populations with “long-term physical conditions,” 61 different measures of adherence within 58 studies were identified, but the majority did not have established reliability or validity. Bailey and colleagues assessed variation in measurement and definition of exercise adherence among studies investigating the effects of exercise on musculoskeletal pain, and only 49% of the 86 included studies provided parameters for determining satisfactory adherence to exercise interventions.

No studies have looked at the influence of adherence on exercise therapy treatment effect. Adherence rates and methods for tracking adherence are unknown in the knee OA population. Without known adherence rates, decision making is limited in its ability to generalize the information from the studies about exercise choice and dosing. Hence, the purpose of this review was to determine the impact of unsupervised exercise program adherence on exercise therapy treatment effect, as well as determine how adherence is tracked and reported in exercise trials for knee OA.

Methods
Information sources
In accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), this systematic review investigated clinical trials that included exercise programs for the management of individuals with knee OA as identified through database searches in PubMed, CINAHL, Medline (OVID), EMBASE and Cochrane, as well as reference lists from studies meeting the inclusion criteria, with publication dates of 1 January 2002 through 31 December 2021.

The protocol was registered (PROSPERO CRD42020187787).

Unsupervised exercise was defined as dosed therapeutic exercise that is completed without the direct observation of a healthcare provider, researcher, or fitness instructor. It is often referred to as home exercise program or home-based exercise. Supervised exercise is defined as dosed therapeutic exercise that is completed with the direct observation of a healthcare provider, researcher, or fitness instructor.

Operationally, we included both facility-based visits and telehealth visits that included video and audio.

Eligibility criteria
Randomized clinical trials were included if they consisted of adults 18 years or older with a diagnosis of knee OA where the primary intervention was supervised exercise therapy that also included a structured unsupervised exercise program (not simply instructions to exercise more or be physically active). Studies were excluded if they were not written or translated in English or included participants with a history of knee arthroplasty because surgery may influence exercise application (i.e., exclusion revision surgeries).

Search strategy
Searches included MeSH terms and keywords related to the knee using the Boolean operator “OR” including, but not limited to, variations of the terms tibiofemoral or patellofemoral. These were combined using the Boolean operator “AND” with MeSH
terms and keywords for OA variations. Lastly, we combined these using “AND” with keywords related to physical therapy or MeSH terms and keywords for exercise (including specific types of exercise such as walking, stretching, or tai chi). Supplementary Material A provides detailed search strategies.

The searches and full-text studies were organized within Covidence (Veritas Health Innovation, Melbourne, Australia, www.covidence.org), a review management software, where all study screening, data extraction and risk of bias assessments were conducted.

Selection and data collection process

Titles and abstracts and full-text studies were screened by two reviewers (KS and BM), with discrepancies resolved through consensus. If no consensus could be reached, a third reviewer would arbitrate (JY). Data were extracted from the final full text studies by one reviewer (KS or BM) and the other confirmed accuracy (unblinded).

Data items

The following data elements were extracted: the presence of unsupervised exercise adherence reports and the method for monitoring adherence. While adherence and compliance have different operational definitions, they are used interchangeably in research reports, and trials using either term were included as long as the intent of tracking fidelity of the unsupervised exercise program was present. For the intent of this review, adherence was defined as “the degree behavior corresponds with an agreed on recommendation.” Compliance was defined as yielding to or obeying clinician instructions.26 These usually occur on a spectrum or scale and are rarely dichotomous endstates. See Supplementary Material B for all operational definitions.

Study risk of bias and reporting bias assessment

Methodological quality and bias was extracted from the Physiotherapy Evidence Database (PEDro). For 21 studies, there was no score in the PEDro database and were therefore scored by the reviewers (KS and BM) with consensus reached for all 21 studies (no third reviewer required).27–47 The PEDro scale is considered a valid and reliable measure of bias.48 PEDro scoring includes 11 criteria regarding eligibility, randomization, blinding, baseline statistics, outcome data, missing values and intention to treat, between-group comparisons, and reporting of statistical data. The first item is not used as part of the final scoring because it does not affect internal validity; therefore, the final score is between 0 and 10. A score of five or greater is considered adequate quality.49

Synthesis methods

A synthesis of adherence tracking approaches was conducted. Extracted data were sub-grouped both by collection method and reporting properties. Based on similar systematic reviews, six broad categories of tools for tracking adherence were created and included: single-item self-report questionnaires, multi-item self-report questionnaires, home (unsupervised) exercise diaries, accelerometer-based direct monitoring devices including fitness watches, app/computer-assisted, and researcher check-in.10,19,50 Single-item self-report questionnaires consist of a single question that allows subjects to qualify adherence to exercise parameters for the particular study, whereas multi-item self-report questionnaires accomplish this with the inclusion of more than one question.10,19,50 Multi-item questionnaires were uniquely generated for a particular study or used previously established measures.10,19,50 Generally, home (unsupervised) exercise diaries track adherence with the use of ongoing record keeping of exercise performance throughout the course of a study.10,19,50 Accelerometer-based direct monitoring devices track a subject’s physical activity with the use of electronic devices that are sensitive to positional change in space. App/computer-assisted measures are electronic interfaces that promote adherence with the use of electronic reminders and adherence reporting. Researcher check-in utilizes ongoing monitoring of adherence through verbal or written report by the subjects at predetermined intervals.

As data were extracted, it became apparent that, in addition to different methods for tracking adherence, studies could also be grouped by the various methods for reporting adherence. These included categories such as percentage of exercises completed (mean by group) and percentage of individuals meeting a threshold predetermined by the researchers (repetitions per day, days exercised per week etc.). Observed commonalities allowed for eight different categories of methods for reporting adherence.

Effect measures

If the data could be homogenized adequately, a meta-analysis was planned to assess the influence of unsupervised exercise program adherence on exercise therapy treatment effect.

Certainty assessment

Certainty assessments were not conducted because the heterogeneity and lack of adherence reporting did not allow for effect estimates.

Results

From 3622 abstracts screened once duplicates were removed, 442 full-texts were reviewed, and 176 met the final eligibility criteria (Fig. 1). Mean PEDro scores are reported in Table 1. Individual item and total PEDro score are presented for each study in Supplementary Material C. The mean and median PEDro score for all studies was 6.3 and 7 respectively, with scores ranging from 2 to 9.

The mean exercise adherence rate was 67.9% and ranged from 3.7% to 100% but was only reported in 72 (40.9%) of the trials. Another 26 (14.8%) trials reported that they collected adherence data but did not report any values. Specific adherence rates for each study are shown in Supplementary Material D.

Measures of adherence were highly variable across studies. Forty-three (43.9%)22,35,42,51–97 studies used an unsupervised exercise diary, 23 (23.5%)38,46,91–111 used multiple collection methods, 12 (12.2%)9,112–122 used researcher check-in, 6 (6.1%)100,123–127 used a single-item questionnaire, 5 (5.1%)128–132 used a multi-item questionnaire, and 2 (2.0%)133
used an accelerometer or fitness watch. Another 7 (7.1%) studies used a method that did not fit into these categories or the collection method was unclear. All but one of the studies that used multiple collection methods used at minimum an unsupervised exercise diary as one of the methods. The most common combination of methods used in 13 studies (13.3%) was both an unsupervised exercise diary and researcher check-in. Pedometers were used by 3 (3.1%) studies that used multiple collection methods and 1 (1.0%) study a specific collection method.

Methods for monitoring and reporting unsupervised exercise adherence were grouped into eight categories determined by reporting properties. The most commonly used reporting method was the mean percentage of prescribed exercises completed by each group (n = 24, 33.3% of reporting studies). Other reporting methods included Likert scales or percentage of individuals who completed all exercises. Nine (12.5% of reporting studies) studies were included in a “unique” category because only one study used each method. One study used both adherence to home exercise three times weekly using a numeric-rating scale and Exercise Adherence Rating Scale (EARS) Section B.

---

**Table 1** Pooled descriptive data across studies that collected and reported adherence data (n = 72 studies).

<table>
<thead>
<tr>
<th>Descriptive data</th>
<th>Pooled results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of participants</td>
<td>9492 (132)</td>
</tr>
<tr>
<td>(mean per study)</td>
<td>64.0 [38.8 – 75.3]</td>
</tr>
<tr>
<td>Pooled mean age of participants [range]</td>
<td>Male: 30.6%</td>
</tr>
<tr>
<td>Pooled mean duration of studies [range]</td>
<td>33.5 weeks</td>
</tr>
<tr>
<td>Pooled mean PEDro scale score [range]</td>
<td>6.7 [2 – 9]</td>
</tr>
<tr>
<td>Pooled mean adherence rate [range]</td>
<td>67.9% [3.7% - 100%]</td>
</tr>
</tbody>
</table>
were based on a percent completion, and the rest met a minimum threshold of exercise (Table 3).

**Discussion**

Of 98 trials (55.7% of all studies) making a claim that adherence was measured, actual rates were reported in only 72 trials (40.9%). The mean adherence rate was 67.9% across these trials (range from 3.7% to 100%). Only 18 trials (10.2%) reported a predetermined threshold of acceptable adherence and based on these studies, there was no clear indication that adherence improved outcomes, although large heterogeneity in methodology limits any conclusions. The most commonly used tools to measure adherence were exercise diaries followed by researcher check-in. Multiple methods were used in 23 studies, in which authors used a combination of assessments. There was also high variability in the approaches taken to report adherence across studies, with eight distinct categories identified. It is notable that although we grouped adherence into eight categories based on similar characteristics, there was further variability within some of those categories. For example, there was wide variation among studies as to what researchers considered to represent being adherent. Overall, poorly measured adherence and/or reported adherence further confounds our understanding of treatment effects for exercise interventions for knee OA.

**Comparison to recent literature**

A recent systematic review studied adherence to many types of knee OA interventions and found that treatment frequency was the most commonly reported measure but few

---

**Table 2** Methods for reporting adherence to unsupervised home exercise programs.

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of studies (% of 98 studies that collected adherence data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of prescribed exercises completed - mean for group(s)</td>
<td>24 (24.5%)</td>
</tr>
<tr>
<td>Percentage of prescribed exercises completed - median for group(s)</td>
<td>1 (1.0%)</td>
</tr>
<tr>
<td>Percentage of individuals meeting thresholds of acceptable adherence</td>
<td>18 (18.4%)</td>
</tr>
<tr>
<td>predefined by researchers (reps/day and days exercised/week, daily</td>
<td></td>
</tr>
<tr>
<td>steps, etc.)</td>
<td></td>
</tr>
<tr>
<td>Likert or numeric rating scales</td>
<td>10 (10.2%)</td>
</tr>
<tr>
<td>Unique methods</td>
<td>9 (9.2%)</td>
</tr>
<tr>
<td>Percentage value without description of how value was calculated</td>
<td>4 (4.1%)</td>
</tr>
<tr>
<td>Percentage of individuals who completed all exercises</td>
<td>4 (4.1%)</td>
</tr>
<tr>
<td>Percentage of individuals crudely qualified into categories (e.g. good</td>
<td>2 (2.0%)</td>
</tr>
<tr>
<td>moderate, etc.)</td>
<td></td>
</tr>
<tr>
<td>No reporting method but indicated that collected adherence</td>
<td>26 (26.5%)</td>
</tr>
</tbody>
</table>

---

![Fig. 2](image-url) Method for collecting adherence data from studies that collected data.
studies quantified what an acceptable frequency was (i.e. level of adherence). In 2017, a systematic review reporting on adherence measures for unsupervised exercise programs used with musculoskeletal disorders found that out of 234 ways to measure exercise adherence, only 49 were reproducible and only seven had any validity and reliability metrics to support them. All seven of these were multi-item questionnaires, but only one, the Community Health Activities Model Program for Seniors (CHAMPS), has been validated and even this was of questionable quality for research

Table 3 Details of studies reporting predetermined thresholds of acceptable adherence.

<table>
<thead>
<tr>
<th>Study</th>
<th>Met-predetermined measure of adequate adherence (Rate -%)*</th>
<th>Positive outcome favoring exercise therapy with respect to comparator</th>
<th>Description of adequate adherence rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Abbott et al. 2013</td>
<td>42.8, 80.2, 65.2</td>
<td>Yes</td>
<td>Both groups had exercise; No difference in primary outcomes between groups.</td>
</tr>
<tr>
<td>3. Christensen et al. 2015</td>
<td>3.7</td>
<td>No</td>
<td>Unspecified</td>
</tr>
<tr>
<td>4. Crossley et al. 2015</td>
<td>77</td>
<td>Yes</td>
<td>Unspecified</td>
</tr>
<tr>
<td>5. de Rooij et al. 2017</td>
<td>86</td>
<td>Yes</td>
<td>Completed 2/3 of exercise sessions</td>
</tr>
<tr>
<td>6. Fitzgerald et al. 2011</td>
<td>52, 47</td>
<td>Both groups had exercise; no difference in primary outcomes between groups</td>
<td>Completed &gt;80% of sessions</td>
</tr>
<tr>
<td>7. Hinman et al. 2007</td>
<td>84</td>
<td>Yes</td>
<td>Unclear -% of participants who completed unsupervised exercised in any amount</td>
</tr>
<tr>
<td>8. Kloek et al. 2018</td>
<td>81.1</td>
<td>Both groups had exercise; no difference in primary outcomes between groups</td>
<td>Completed 2/3 of modules</td>
</tr>
<tr>
<td>9. Liao et al. 2021</td>
<td>82.6, 84.3</td>
<td>Both groups had exercise; group with higher adherence rate had better primary outcomes.</td>
<td>Used indicated color band for at least 1 session</td>
</tr>
<tr>
<td>10. Loew et al. 2017</td>
<td>62, 35.44, 38.44, 56.1, 43.1, 37.9, 76.4, 54.9, 43.4</td>
<td>All groups had exercise; adherence was the primary outcome measure. Preference for exercise appears to affect adherence.</td>
<td>Completed &gt;66% of sessions</td>
</tr>
<tr>
<td>11. Osteras et al. 2019</td>
<td>67, 45</td>
<td>Both groups had exercise; group with higher adherence rate had better primary outcomes.</td>
<td>Had specific exercise for intervention but measured adherence as those who met physical activity guidelines over trial period</td>
</tr>
<tr>
<td>12. Rattanachaiyanont and Kuptniratsaikul 2008</td>
<td>57.4, 22</td>
<td>Both groups had exercise; no difference in outcomes between groups</td>
<td>At least 50 repetitions per day for at least 5 days per week</td>
</tr>
<tr>
<td>13. Schlenk et al. 2021</td>
<td>77.4, 81.1</td>
<td>Yes</td>
<td>Completed &gt;75% of exercises</td>
</tr>
<tr>
<td>14. Sharma M, Singh A, Dhillon MS, Kaur S 2018</td>
<td>60†</td>
<td>Both groups had exercise; no difference in outcomes between groups</td>
<td>Unclear - “Completed exercises correctly”</td>
</tr>
<tr>
<td>15. Stensrud et al. 2015</td>
<td>62.2</td>
<td>Yes</td>
<td>Met-or exceed weekly step goal</td>
</tr>
<tr>
<td>16. Talbot et al. 2003</td>
<td>56, 35</td>
<td>Yes</td>
<td>Completed 70 +/- 10 min of weekly exercise</td>
</tr>
<tr>
<td>17. Wallis et al. 2017</td>
<td>70</td>
<td>No</td>
<td>Completiing exercises 3 times per week for one group; “continuing to walk” for comparator</td>
</tr>
<tr>
<td>18. Xiao et al. 2021</td>
<td>70.59, 35</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

* The multiple percentages per study are data points for different groups OR different time periods.
† Only tracked adherence in the intervention group.
‡ Only reported one value across both groups.
A more recent tool that has been validated in patients with chronic low back pain and appears to be useful for both clinical and research is the EARS. This is also currently being studied in patients with knee OA. In our review, only three studies used these validated tools (one used CHAMPS; two used EARS).

Data collection methods for adherence

The majority of studies that collected adherence data (n = 56; 57.1%) used self-reported exercise diaries, either alone or in combination with other collection methods. This is consistent with the findings by Duong et al., where 66.4% of studies used an unsupervised exercise diary. We found high variability for when and how these were filled out by patients. Some studies expected the exercise diary to be completed daily and others weekly. There was also variance in what variables were logged, such as simply whether a session was done or if more complex details were required, including reporting the number of minutes completed or exact number of repetitions.

There were also many different strategies that researchers used to calculate percentage values representing adherence across studies. While this may appear to be the most feasible strategy, there does not appear to be extensive research on the validity and reliability of exercise diaries for this purpose. Findings from a cohort study not in this review simultaneously using accelerometers and exercise diaries found that self-reported exercise adherence over a 12-week period was overestimated in the diary and had only a moderate correlation with accelerometer data. One study in our review used both exercise diaries and an accelerometer but did not correlate the data between the two methods.

Another study concluded that the design of the diary likely does not matter, but that individualized data from the diaries were highly variable and lacked validity. However, in another systematic review on the use of home-based rehabilitation interventions, the conclusion was that exercise diaries may be valid in certain populations. Therefore, due to conflicting conclusions, the adherence rates from exercise diaries may not be useful.

Clinical implications

What is known about the efficacy of exercise interventions for people with knee OA could be adversely affected by the limited number of studies that collect and report exercise adherence data. Reporting of exercise dosing in this population is not held to the standards applied in drug trials. Almost half of the studies (n = 78; 44.3%) did not monitor exercise adherence, limiting conclusions that can be made about its efficacy. Based on this information, it is difficult to fully understand exercise treatment effects reported in trials and make recommendations for exercise prescription in clinical care. While an individual patient data (IPD) meta-analysis could potentially show greater promise for answering this research question, it would likely still be of limited value until adherence measures are better harmonized and reported across exercise trials for knee OA.

This review highlights a critical research gap: the lack of standardization with collecting and reporting of exercise adherence and the consideration of adherence when analyzing treatment effect in exercise therapy trials. The inability to establish superiority between different types of exercise for knee OA could be influenced by a poor understanding of adherence rates. Further research to define optimal method(s) for collecting and reporting exercise adherence is an important need. These data are necessary to understand how adherence rates influence treatment effects for knee OA trials. Until this research is completed, researchers are most likely best served by using multiple methods for collection of data adherence including a home diary, a multi-item questionnaire (possibly the EARS), and an accelerometer.

Limitations

There are some limitations to discuss. Diagnostic criteria for knee OA could not be verified or validated, and therefore assessing heterogeneity and phenotypes of participants across the disease spectrum was not possible (e.g., low versus high severity). Exercise may have been used but was not the primary intervention in some trials, and therefore was missed by our search terms. Also, several studies in this review assessed “adherence” to the entire intervention, of which unsupervised exercise was only one part. Poor adherence in these cases might not have been specific to the exercise components.

Conclusion

Data regarding adherence to unsupervised exercise programs are not consistently collected (or reported) among studies investigating exercise interventions for knee OA. Approximately half of the studies (n = 98; 55.7%) claimed to collect these data, but there was no consistent method for collecting or reporting across studies. Clinical trials should adequately collect and report adherence to unsupervised exercise, as this factor is essential to developing a better understanding of treatment effect. A more consistent and validated approach would improve the ability to compare results across trials.

Disclaimer

The view(s) expressed herein are those of the author(s) and do not reflect the official policy or position of Brooke Army Medical Center, the U.S. Army Medical Department, the Department of the Army, the Uniformed Services University of the Health Sciences, the Defense Health Agency, the Department of Defense, or the U.S. Government.

Conflicts of interest

The authors have no financial disclosures to report and no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjpt.2023.100533.
References


