The Value of Preoperative Exercise and Education for Patients Undergoing Total Hip and Knee Arthroplasty

A Systematic Review and Meta-Analysis

Rebecca Moyer, PT, PhD
Kathy Ikert, PT
Kristin Long, PT
Jacquelyn Marsh, PhD

Investigation performed at The University of Western Ontario, London, Ontario, Canada

Abstract

Background: Existing evidence regarding the value of preoperative education and/or exercise (prehabilitation) for patients undergoing total joint replacement is conflicting. The purpose of this study was to conduct an updated, comprehensive systematic review with meta-analyses to determine the longitudinal effects and efficacy of prehabilitation on postoperative outcomes in patients undergoing total hip arthroplasty (THA) or total knee arthroplasty (TKA).

Methods: We searched 11 electronic databases (MEDLINE, AMED, CINAHL, Embase, Scopus, ProQuest, PEDro, SportDiscus, PsycINFO, and Cochrane) from their inception to May 2016 for randomized controlled trials that compared changes in pain, function, strength, anxiety, and hospital length of stay following THA or TKA. Two reviewers independently determined study eligibility, rated study quality, and extracted data. There were no restrictions on study dates, patient characteristics, or the follow-up time point at which postoperative outcomes were measured. We excluded trials comparing 2 interventions. Methodological quality assessments were performed using the Cochrane risk-of-bias tool. We calculated pooled estimates, with 95% confidence intervals (CIs), of standardized mean differences (SMDs).

Results: Thirty-five studies with 2,956 patients were included. After a preoperative program, patients undergoing THA, but not TKA, had significantly less postoperative pain than controls (SMD = 0.15, 95% CI = 0.03 to 0.27, p = 0.017). Postoperative function was also significantly improved compared with controls, with similar improvement after THA (SMD = 0.32, 95% CI = 0.15 to 0.50, p < 0.001) and TKA (SMD = 0.32, 95% CI = 0.06 to 0.57, p = 0.015). Significantly greater quadriceps strength was observed after TKA (SMD = 0.42, 95% CI = 0.16 to 0.68, p = 0.002). No significant differences in hamstring strength were observed between groups after TKA (p = 0.132). Small-to-moderate but nonsignificant improvements in anxiety (SMD = 0.17, 95% CI = −0.05 to 0.39; p = 0.128) were observed after THA, and length of stay was significantly shorter after TKA (SMD = 0.54, 95% CI = 0.24 to 0.84, p < 0.001) and THA (p = 0.027).

Conclusions: Overall effect sizes for prehabilitation were small to moderate. In patients undergoing TKA, significant improvements were observed in function, quadriceps strength, and length of stay. In patients undergoing THA, significant improvements were observed in pain, function, and length of stay. Included studies were inconsistent with

Disclosure: No external funding was received for this study. The Disclosure of Potential Conflicts of Interest forms are provided with the online version of the article (http://links.lww.com/JBJSREV/A285).
regard to the types of outcome measures reported, and the quality of the interventions varied. A more standardized approach to reporting of clinical trial interventions and patient compliance is needed to thoroughly evaluate the effects of prehabilitation on postoperative outcomes.

**Level of Evidence**: Therapeutic Level II. See Instructions for Authors for a complete description of levels of evidence.

Osteoarthritis is a leading cause of pain, disability, and health-care resource use, with growing concern regarding its increasing incidence and socioeconomic burden. Total joint arthroplasty is a cost-effective intervention for hip and knee osteoarthritis, improving pain, function, and quality of life; however, an aging and increasingly obese population has resulted in higher demands for this procedure. There were 47,137 total hip arthroplasties (THAs) and 57,718 total knee arthroplasties (TKAs) performed in Canada from 2012 to 2013, representing an increase of 16.5% since 2008. In the United States, >7 million Americans have had a hip or knee replacement and by 2030, the demands for THA and for TKA are estimated to climb 174% and 673%, respectively, from 2005 levels. Internationally, the average rate of primary THA and TKA is 175 procedures per 100,000 individuals. Although policymakers have focused on reducing wait times worldwide for priority procedures such as THA and TKA, these wait times are still excessive for patients with pain and disability.

Evidence suggests that better preoperative health status (e.g., greater physical function and strength) is a strong predictor of good postoperative outcome following joint replacement.

Additional evidence has shown that health-related quality of life deteriorates during the preoperative period; therefore, longer wait times may have negative consequences for postoperative outcomes. Preoperative patient expectations may also impact function and quality of life after arthroplasty. The potential deterioration of function during the preoperative period, accompanied by the effects of greater disease severity on postoperative outcomes, highlights a critical opportunity to supplement current preoperative management. Importantly, both patient education and exercise prior to surgery (prehabilitation) may benefit mobility, improve function, and optimize surgical outcomes.

Although previous systematic reviews have investigated the effects of preoperative education and/or exercise on outcomes following THA and TKA, the evidence is conflicting regarding whether patients receive added benefit if they undergo prehabilitation, and whether this improvement is similar for patients undergoing THA and TKA. Additionally, the scientific rigor of the previous reviews is limited by the inclusion of outdated reports. There is a need to critically evaluate the most recent literature and to determine whether postoperative outcomes differ between the prehabilitation programs that are effective preoperatively and those that are not. There is also support for the use of targeted exercises to improve outcomes for patients with hip and knee osteoarthritis, given the rapidly growing incidence rates, the socioeconomic burden, and prolonged wait times. Therefore, assessment of the role of prehabilitation prior to THA and TKA and identification of which intervention is most effective are urgently needed. The purpose of this study was to conduct an updated, comprehensive systematic review with meta-analyses to determine the longitudinal effects and efficacy of preoperative rehabilitation on postoperative pain and function in patients undergoing THA or TKA (quantified using standardized mean differences [SMDs]). We also evaluated secondary outcomes of interest, including strength, anxiety, and hospital length of stay.

**Materials and Methods**

**Literature Sources**
The following electronic databases were searched from their inception to May 2016: MEDLINE, AMED, CINAHL, Embase, Scopus, ProQuest, PEDro, SportDiscus, PsycINFO, and Cochrane. Searches used combined and/or truncated key terms including rehabilitation, prehabilitation, preoperative, presurgical care, exercise, education, physical therapy, physiotherapy, total knee arthroplasty, total hip arthroplasty, total knee replacement, total hip replacement, joint arthroplasty, joint replacement, knee prosthesis, hip prosthesis, and joint prosthesis. The terms randomized controlled trial, clinical practice guideline, and systematic review were not required, in order to ensure that all studies were captured. Reference lists of systematic reviews, clinical practice guidelines, and included randomized controlled trials were manually searched for additional studies. A detailed protocol for this review has not been previously published. A database search strategy is shown in the Appendix.

**Study Selection and Inclusion Criteria**
Randomized controlled trials, clinical practice guidelines, or systematic reviews were included if they examined the effect of preoperative interventions involving exercise and/or education on postoperative outcomes in patients undergoing THA or TKA, and were published since 2000 as full-text, English-language journal articles. An exercise intervention was defined as any aerobic, range-of-motion, strength, or proprioceptive activity requiring physical effort and prescribed with a defined number of sets and repetitions. There were no restrictions on study dates, patient characteristics, or the follow-up time point at which postoperative outcomes were measured. Trials comparing
2 interventions (i.e., those that did not have a control group receiving no treatment or the standard of care) were excluded. Two reviewers independently assessed eligibility in 2 stages. First, titles and abstracts were screened, and classified as either “not meeting the inclusion criteria” or “uncertain.” Articles considered “uncertain” were obtained as full manuscripts and reviewed to determine eligibility. Disagreements were discussed and a third independent reviewer was consulted to achieve consensus. Details of the literature search are reported using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines21,22.

**Outcome Measures and Data Extraction**

The primary postoperative outcome was a patient-reported measure of pain and function. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)23 subdomains for pain and function were extracted from each study when applicable; otherwise, alternate measures of pain and function were used. Secondary outcome measures of interest included postoperative measures of quadriceps and hamstring strength, anxiety, and length of stay. Two reviewers independently extracted means and standard deviations, mean differences, or effect sizes for the outcomes of interest. Reviewers also extracted the following information from each article: sample size; patient demographics including age, sex, and body mass index (BMI); intervention details; and the efficacy of the intervention prior to surgery. Program effectiveness was defined as a significant improvement in the outcome of interest during the preoperative period from before to immediately after the program. Studies that did not report follow-up on the intervention prior to surgery were treated conservatively and identified as ineffective. Outcomes were extracted using the longest follow-up time point for each study. Authors were contacted when information was insufficient. If authors could not be contacted, information was extracted from original figures or from review articles that discussed the study.

**Methodological Quality Assessment**

Two reviewers independently evaluated the quality of included studies using the Cochrane risk-of-bias tool24, which has 7 items: sequence generation, allocation concealment, blinding of patients and outcome assessors, incomplete data, selective reporting, and “other.” The "other" category included whether the intervention and control groups were similar at baseline and whether a study followed intention-to-treat (ITT) principles. Each item was rated as a low, unclear, or high risk of bias. Disagreements were discussed and a third independent reviewer was consulted to achieve consensus.

**Statistical Analysis**

Agreement between reviewers was evaluated using the kappa (κ) statistic25,26. Before data pooling, studies were categorized according to joint, follow-up, and program effectiveness. The SMD for an outcome in a particular study, at the longest follow-up time point, was calculated as the postoperative difference between the control and experimental groups divided by the pooled standard deviation, using the reported sample sizes, means, and standard deviations or using the reported mean difference. For each outcome, a pooled estimate for the SMD and an accompanying 95% confidence interval (CI) were calculated using the inverse-variance method and a random-effects model. An outcome favoring the experimental group was represented as a negative value, and the magnitude of the SMD was interpreted using the Cohen d value27. Publication bias was assessed using the Egger regression test27, and if such bias was present, a trim-and-fill method28 was used to reassess the treatment effect after adjustment for selective reporting. The proportion of variability associated with heterogeneity was assessed using the I² and Q statistics29,30. I² was interpreted as low (25%), moderate (50%), or high (75%) heterogeneity31. If significant heterogeneity (p < 0.05) was found, the effect was further evaluated by appropriately removing outliers and by performing sensitivity analyses among subgroups involving the (1) joint (hip or knee), (2) follow-up time (≤6 weeks after surgery, at 3 months after surgery, or ≥6 months after surgery), or (3) program effectiveness (defined as significant improvement following completion of the program but prior to surgery). For the primary outcome measures, previously reported standard deviations from a typical population with osteoarthritis were used to translate the SMDs into a mean difference in WOMAC pain and function change scores32. Each meta-analysis was performed using Comprehensive Meta-Analysis software (version 2; Biostat). A p value of <0.05 or a 95% CI that did not include 0 was considered significant.

**Results**

**Study Selection**

After removing duplicates, 5,392 potentially relevant articles were identified; 160 full-text articles were screened, and 35 randomized controlled trials33-67 met the eligibility criteria and were included in the final analyses (Fig. 1). Interrater agreement was excellent for determining eligibility on the basis of titles and abstracts (κ = 0.77) and full-text articles (κ = 0.80). The outcome measures for all 35 randomized controlled trials are described in Table I. Nine authors were contacted, and 8 responded by providing additional information that was not reported in the original article.

**Study Characteristics**

Data were extracted for 2,959 patients: 1,151 underwent TKA (mean age [and standard deviation], 67 ± 3.2 years; 419 male, 732 female), 1,193 underwent THA (69 ± 6.4 years; 535 male, 658 female), and the joint (hip or knee) was not specified for 615 patients (68 ± 2.8 years; 190 male, 425 female) in 5 randomized controlled trials. Fifteen
studies used exercise alone (10 TKA studies \(48,51,53,55,57,58,60,62,65,66\), 4 THA studies \(35,52,54,56\), 1 study of TKA and THA \(45\)), 9 studies used education alone (6 THA studies \(33,37,41,44,64,67\), 3 studies of TKA and THA \(34,46,63\)), and 11 studies used a combination (3 TKA studies \(39,42,59\), 6 THA studies \(38,40,43,47,49,50\), 2 studies of TKA and THA \(36,61\) ) (Table I). Preoperative programs typically ran between 4 and 8 weeks (minimum, 2 weeks; maximum, 24 weeks) and 1 to 5 times per week (mean, approximately 3 times per week).

**Quality Assessment of Included Studies**

Interrater agreement for the Cochrane risk-of-bias tool was good (\(\kappa = 0.71\)). Disagreements were most common among randomized controlled trials, with variable outcome reporting and incomplete data. Consensus ratings for the methodological quality are given in the Appendix. Across most studies, risk of bias was low for completeness of outcome data and selectivity of outcome reporting, low for whether random sequence generation and allocation concealment were used, and high for whether participants, personnel, and outcome assessors were blinded. Other sources of bias that existed in some studies included dissimilarity between groups at baseline and no intention-to-treat protocols. No articles were excluded on the basis of quality.

**Outcome Measures**

**Pain**

Twenty-five studies reported the effects of prehabilitation on postoperative pain (Table II) \(33,34,37,42,45,48,51,53,55,61,63-66\). Overall, there was a small, significant difference favoring prehabilitation for improving pain (SMD = 0.14, \(p = 0.007\)). The funnel plot was symmetric and negative for the existence of publication bias (intercept \(=-0.17\), \(p = 0.779\)). When analyzed by joint, there was a similar but nonsignificant SMD in patients treated for knee osteoarthritis (SMD = 0.11, \(p = 0.136\)), and a larger SMD in patients with hip osteoarthritis (SMD = 0.15, \(p = 0.017\)); the latter represented a significant improvement in pain in patients who underwent prehabilitation compared with controls.

When analyzed with respect to study follow-up, SMDs for improvements in postoperative hip pain were small and nonsignificant at \(\leq 6\) weeks and \(\leq 6\) months (Table II). At 3 months, there was a moderate, significant difference favoring prehabilitation for improving pain.
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size, I/C</th>
<th>Age, I/C (yr)</th>
<th>Female Sex, I/C (%)</th>
<th>BMI, I/C (kg/m²)</th>
<th>Joint</th>
<th>Preop. Intervention</th>
<th>Follow-up</th>
<th>Loss to Follow-up, I/C (%)</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doering et al. (2000)</td>
<td>46/54</td>
<td>58.7 ± 10.8/60.4 ± 8.7</td>
<td>46/51</td>
<td></td>
<td>Hip</td>
<td>Educational video (1 day)</td>
<td>4 d</td>
<td>0/0</td>
<td>VAS, LOS, STAI</td>
</tr>
<tr>
<td>Weaver et al. (2003)</td>
<td>69/67</td>
<td>72 ± 7.8</td>
<td>86/100</td>
<td></td>
<td>Both</td>
<td>Education for postop. care, 1 home care visit by nurse and PT</td>
<td>6 mo</td>
<td>12/7</td>
<td>WOMAC pain and function</td>
</tr>
<tr>
<td>Gilbey et al. (2003)</td>
<td>7/31</td>
<td>66.73 ± 10.19/63.29 ± 12.01</td>
<td>57/68</td>
<td>27.71 ± 4.78/28.20 ± 3.60</td>
<td>Hip</td>
<td>2 clinic and 2 home sessions, 1 h/wk for 8 wk (aerobic, strength, hydrotherapy)</td>
<td>6 mo</td>
<td>14/19</td>
<td>WOMAC function</td>
</tr>
<tr>
<td>Crowe et al. (2003)</td>
<td>65/68</td>
<td>66.9 ± 11.9/70.7 ± 10.7</td>
<td>78/81</td>
<td>29.3 ± 5.9/29.6 ± 5.9</td>
<td>Both</td>
<td>Custom conditioning program up to 24 wk plus education and a hospital tour</td>
<td>DC</td>
<td>8/0</td>
<td>LOS</td>
</tr>
<tr>
<td>Giraudet-Le Quintrec et al. (2003)</td>
<td>48/52</td>
<td>62.7 ± 8.3/64.3 ± 9.5</td>
<td>56/62</td>
<td></td>
<td>Hip</td>
<td>Small group education sessions, 1 ×/wk for 2-6 wk</td>
<td>1 wk</td>
<td>2/0</td>
<td>VAS, LOS, STAI</td>
</tr>
<tr>
<td>McGregor et al. (2004)</td>
<td>19/20</td>
<td>70.8 ± 9.3/72.8 ± 10.1</td>
<td>—</td>
<td>—</td>
<td>Hip</td>
<td>Education, gait aid instruction, and exercise for 2-4 wk</td>
<td>3 mo</td>
<td>21/0</td>
<td>WOMAC pain and function</td>
</tr>
<tr>
<td>Beaupre et al. (2004)</td>
<td>65/66</td>
<td>67 ± 7/67 ± 6</td>
<td>60/50</td>
<td>32 ± 6/31 ± 5</td>
<td>Knee</td>
<td>Mobility and postop. education plus supervised exercise, 3 ×/wk for 3 wk</td>
<td>12 mo</td>
<td>8/3</td>
<td>WOMAC pain and function, Q5/HS, LOS</td>
</tr>
<tr>
<td>Gocen et al. (2004)</td>
<td>30/30</td>
<td>46.93 ± 11.48/55.50 ± 14.44</td>
<td>45/27</td>
<td>24.94 ± 3.70/27.69 ± 3.70</td>
<td>Hip</td>
<td>Home exercise and education 3 ×/day</td>
<td>24 mo</td>
<td>3/0</td>
<td>VAS, HHS</td>
</tr>
<tr>
<td>Mitchell et al. (2005)</td>
<td>80/80</td>
<td>70.0 ± 7.2/70.6 ± 8.2</td>
<td>63/52</td>
<td>—</td>
<td>Knee</td>
<td>3 home visits/wk (pain relief, range of motion, gait education) for 8 wk</td>
<td>3 mo</td>
<td>29/29</td>
<td>WOMAC pain and function, LOS</td>
</tr>
<tr>
<td>Siggeirsdottir et al. (2005)</td>
<td>27/23</td>
<td>69/66</td>
<td>52/52</td>
<td>—</td>
<td>Hip</td>
<td>Surgical education and postop. exercise instruction</td>
<td>6 mo</td>
<td>0/13</td>
<td>OHS, LOS</td>
</tr>
<tr>
<td>Rooks et al. (2006)</td>
<td>H, 32/31; K, 22/23</td>
<td>H, 65 ± 11.59/7; K, 65 ± 8/69 ± 8</td>
<td>H, 63/52; K, 50/57</td>
<td>H, 28.4 ± 5.3/30.3 ± 9.1; K, 35.7 ± 9.2/33.9 ± 6.5</td>
<td>Both</td>
<td>Custom strength and conditioning, 3 ×/wk for 6 wk</td>
<td>6 mo</td>
<td>H, 22/23; K, 36/35</td>
<td>WOMAC pain and function, Q5</td>
</tr>
<tr>
<td>Johansson et al. (2007)</td>
<td>62/61</td>
<td>59.7/65.2</td>
<td>52/51</td>
<td>—</td>
<td>Hip</td>
<td>Educational concept maps for 120 min, 2 wk prior</td>
<td>DC</td>
<td>11/16</td>
<td>LOS</td>
</tr>
<tr>
<td>Pour et al. (2007)</td>
<td>46/48</td>
<td>60.5 ± 10.4/61.2 ± 7.6</td>
<td>48/50</td>
<td>25.5 ± 2.8/26.5 ± 2.7</td>
<td>Hip</td>
<td>Accelerated protocol of exercises plus education, 2-3 sessions</td>
<td>DC</td>
<td>0/0</td>
<td>HHS, LOS</td>
</tr>
<tr>
<td>Williamson et al. (2007)</td>
<td>60/61</td>
<td>70.0 ± 8.79/69.6 ± 10.0</td>
<td>52/54</td>
<td>32.8 ± 5.68/32.7 ± 6.45</td>
<td>Knee</td>
<td>Small group exercise 1 ×/wk for 6 wk</td>
<td>3 mo</td>
<td>38/31</td>
<td>VAS, OHS, LOS, HAD</td>
</tr>
</tbody>
</table>

continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size, I/C</th>
<th>Age, I/C (yr)</th>
<th>Female Sex, I/C (%)</th>
<th>BMI, I/C (kg/m²)</th>
<th>Joint</th>
<th>Preop. Intervention</th>
<th>Follow-up</th>
<th>Loss to Follow-up, I/C (%)</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larsen et al. (2008)²⁶</td>
<td>46/45</td>
<td>64 ± 10.8/66 ± 9.2</td>
<td>56/45</td>
<td>—</td>
<td>Both</td>
<td>Expanded education program to establish postop. goals</td>
<td>3 mo</td>
<td>0/0</td>
<td>LOS</td>
</tr>
<tr>
<td>Vukomanović et al. (2008)²⁷</td>
<td>23/22</td>
<td>60.05 ± 11.01/56.2 ± 18.45</td>
<td>70/80</td>
<td>—</td>
<td>Hip</td>
<td>Education by physiatrist plus mobility and exercise instruction, 1 education and 2 exercise sessions</td>
<td>DC</td>
<td>22/18</td>
<td>VAS, HHS, LOS</td>
</tr>
<tr>
<td>Ferrara et al. (2008)²⁸</td>
<td>12/11</td>
<td>63.82 ± 9.01/63.08 ± 6.89</td>
<td>64/58</td>
<td>—</td>
<td>Hip</td>
<td>Group and individual exercise, 60 min, 5 days/wk for 4 wk, and postop. education</td>
<td>3 mo</td>
<td>0/17</td>
<td>WOMAC pain and function, QS</td>
</tr>
<tr>
<td>Topp et al. (2009)²¹</td>
<td>26/28</td>
<td>64.1 ± 7.05/63.6 ± 6.68</td>
<td>73/64</td>
<td>32.16 ± 5.87/32.0 ± 6.09</td>
<td>Knee</td>
<td>Resistance, flexibility and step training, 1 supervised and 2 home sessions, 3×/wk</td>
<td>3 mo</td>
<td>—</td>
<td>VAS, QS</td>
</tr>
<tr>
<td>Hoogeboom et al. (2010)²²</td>
<td>10/11</td>
<td>77.3 ± 3/75.0 ± 5</td>
<td>70/64</td>
<td>26.0 ± 2.6/27.4 ± 4.2</td>
<td>Hip</td>
<td>Supervised exercise program and home exercise, 2×/wk for 3-6 wk</td>
<td>DC</td>
<td>0/9</td>
<td>LOS</td>
</tr>
<tr>
<td>Walls et al. (2010)²³</td>
<td>9/5</td>
<td>64.4 ± 8.0/63.2 ± 11.4</td>
<td>67/80</td>
<td>30.7 ± 3.0/32.8 ± 6.3</td>
<td>Knee</td>
<td>Home-based NMES plus exercise, 20 min/day for 5 days/wk for 8 wk</td>
<td>3 mo</td>
<td>18†</td>
<td>WOMAC pain and function, LOS</td>
</tr>
<tr>
<td>Bitterli et al. (2011)²⁴</td>
<td>41/39</td>
<td>65.37 ± 10.77/68.42 ± 9.74</td>
<td>46/31</td>
<td>27.63 ± 3.60/27.07 ± 3.56</td>
<td>Hip</td>
<td>Written and verbal instructions (2 sessions) prior to sensorimotor training, 2×/day for 2-6 wk</td>
<td>12 mo</td>
<td>12/23</td>
<td>WOMAC function, LOS</td>
</tr>
<tr>
<td>Gottoettner et al. (2011)²⁵</td>
<td>18/20</td>
<td>72.8±66.9</td>
<td>89/70</td>
<td>27.4/28.2</td>
<td>Knee</td>
<td>Proprioceptive training program, 1 supervised session/wk, 45 min, daily practice for 6 wk</td>
<td>1.5 mo</td>
<td>17/0</td>
<td>WOMAC pain and function</td>
</tr>
<tr>
<td>Oosting et al. (2012)²⁶</td>
<td>15/15</td>
<td>76.9 ± 6.3/75.0 ± 6.3</td>
<td>93/67</td>
<td>28.6 ± 5.6/27.8 ± 4.2</td>
<td>Hip</td>
<td>Supervised home exercise, 30 min, 2×/wk and 4 unsupervised exercise sessions for 3-6 wk</td>
<td>1.5 mo</td>
<td>7/20</td>
<td>HOOS pain and function, LOS</td>
</tr>
<tr>
<td>McKay et al. (2012)²⁷</td>
<td>10/12</td>
<td>63.5 ± 4.93/60.58 ± 8.05</td>
<td>50.0/66.7</td>
<td>35.03 ± 6.13/33.78 ± 7.05</td>
<td>Knee</td>
<td>Lower extremity exercise, 3×/wk for 6 wk</td>
<td>3 mo</td>
<td>30/17</td>
<td>WOMAC pain and function, QS</td>
</tr>
<tr>
<td>Brown et al. (2012)²⁸</td>
<td>17/15</td>
<td>—</td>
<td>—</td>
<td>38.8 ± 8.8/34.6 ± 7.6</td>
<td>Knee</td>
<td>Supervised strength and conditioning for 45 min 3×/wk for 8 wk</td>
<td>3 mo</td>
<td>35/53</td>
<td>SF-36 pain and function</td>
</tr>
<tr>
<td>Huang et al. (2012)²⁹</td>
<td>126/117</td>
<td>69.8 ± 7.2/70.5 ± 7.4</td>
<td>68.8/73.5</td>
<td>27.1 ± 4.0/27.2 ± 4.5</td>
<td>Knee</td>
<td>Education and home exercise strengthening for 4 wk</td>
<td>DC</td>
<td>0/0</td>
<td>VAS, LOS</td>
</tr>
</tbody>
</table>

*Continued*
### TABLE I (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size, I/C</th>
<th>Age, I/C (yr)</th>
<th>Female Sex, I/C (%)</th>
<th>BMI, I/C (kg/m²)</th>
<th>Joint</th>
<th>Preop. Intervention</th>
<th>Follow-up</th>
<th>Loss to Follow-up, I/C (%)</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tungtrongjit et al. (2012)</td>
<td>30/30</td>
<td>63 ± 7.6/65.9 ± 7.2</td>
<td>86.7/80.0</td>
<td>24.3 ± 2.4/25.3 ± 3.8</td>
<td>Knee</td>
<td>Quad. strengthening home program, 2 × /day for 3 wk</td>
<td>6 mo</td>
<td>12‡</td>
<td>WOMAC pain and function, QS</td>
</tr>
<tr>
<td>Villadsen et al. (2014)</td>
<td>84/81</td>
<td>67.9 ± 8.6/66.9 ± 8.3</td>
<td>56/56</td>
<td>29.6 ± 4.5/31.1 ± 6.1</td>
<td>Both</td>
<td>NEMEX (core, strength and balance), 1 hr, 2 × /wk for 8 wk, plus education</td>
<td>3 mo</td>
<td>5/10</td>
<td>KOOS or HOOS pain and function</td>
</tr>
<tr>
<td>Matassi et al. (2014)</td>
<td>61/61</td>
<td>66 ± 7.2/67 ± 7.7</td>
<td>54/43</td>
<td>29 ± 4.3/28 ± 3.7</td>
<td>Knee</td>
<td>Exercise instruction plus unsupervised exercise, 5 days/wk for 6 wk</td>
<td>12 mo</td>
<td>—</td>
<td>LOS</td>
</tr>
<tr>
<td>Blau et al. (2015)</td>
<td>106/103</td>
<td>66 (range, 60-74)</td>
<td>58%</td>
<td>25 (range, 23-28)‡</td>
<td>Hip</td>
<td>Small group education, 1 session 4 wk prior to surgery</td>
<td>DC</td>
<td>5/9</td>
<td>NRS, LOS</td>
</tr>
<tr>
<td>Cooke et al. (2016)</td>
<td>45/46</td>
<td>67†</td>
<td>63.40†</td>
<td>32.3 ± 6.46#</td>
<td>Both</td>
<td>Education DVD plus home self-efficacy activities, 4 × prior to admission</td>
<td>1.5 mo</td>
<td>20/19</td>
<td>NRS, STAI</td>
</tr>
<tr>
<td>Skoff er et al. (2016)</td>
<td>30/29</td>
<td>70.7 ± 7.3/70.1 ± 6.4</td>
<td>63/59</td>
<td>30.0 (range, 22.6-42.5)/31.8 (range, 24.3-42.2)</td>
<td>Knee</td>
<td>Supervised resistance training with PT for 60 min, 3 × /wk for 4 wk</td>
<td>1.5 mo</td>
<td>4/28</td>
<td>KOOS pain and function, QS/HS</td>
</tr>
<tr>
<td>Calatayud et al. (2016)</td>
<td>25/25</td>
<td>66.8 ± 4.8/66.7 ± 3.1</td>
<td>84+</td>
<td>32 ± 4.2/31 ± 3.8</td>
<td>Knee</td>
<td>8-wk training program 3 days/wk</td>
<td>3 mo</td>
<td>12/12</td>
<td>WOMAC pain and function, QS/HS</td>
</tr>
<tr>
<td>Jepson et al. (2016)</td>
<td>21/23</td>
<td>67 ± 11.2/65 ± 10.7</td>
<td>33/57</td>
<td>—</td>
<td>Hip</td>
<td>In-home education by an occupational therapist</td>
<td>6 mo</td>
<td>14/9</td>
<td>LOS, NEAIX, HADS-A</td>
</tr>
</tbody>
</table>

*‡ = intervention group, C = control group, VAS = visual analog scale for pain, LOS = length of stay, STAI = State-Trait Anxiety Inventory, PT = physical therapy, WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index, DC = discharge, QS = quadriceps strength, HS = hamstring strength, HHS = Harris hip score, NRS = numeric rating scale for pain, AIMS = Arthritis Impact Measurement Scale, OHS = Oxford Hip Score, H = hip, K = knee, HAD = Hospital Anxiety and Depression Score, NMEES = neuromuscular electrical stimulation, HOOS = Hip injury and Osteoarthritis Outcome Score, SF-36 = Short Form-36, NEMEX = neuromuscular exercise, KOOS = Knee injury and Osteoarthritis Outcome Score, NEADL = Nottingham Extended Activities of Daily Living, and HADS-A = Hospital Anxiety and Depression Scale (anxiety subdomain). Values are given as the mean with or without the standard deviation. If patient demographics were not separated by randomized group.

(SMD = 0.34, p = 0.015). SMDs for improvements in postoperative knee pain were small and nonsignificant, but increased with longer follow-up duration (Table II).

In effective programs, a small but significant SMD in patients with hip osteoarthritis favored prehabilitation for improvements in postoperative pain (SMD = 0.19, p = 0.027). In ineffective programs, the SMD was smaller (0.11) and nonsignificant (p = 0.285) (Table II). However, in patients with knee osteoarthritis, prehabilitation was not significantly favored over the controls for postoperative improvements in pain even for effective programs (Table II).

Heterogeneity was significant for effective programs in patients with knee osteoarthritis (I² = 62.85%, p = 0.019); after removing a single outlier, the SMD decreased from 0.11 to 0.00 and heterogeneity was reduced to a nonsignificant level (I² = 37.80%, p = 0.169).

The pooled SMD of 0.14 translated into a change of 0.63 (95% CI, 0.18 to 1.13) out of 20 in the WOMAC pain subscore. For effective programs, the corresponding WOMAC pain change scores were 0.85 (95% CI, 0.09 to 1.63) for THA and 0.50 (95% CI, −0.18 to 2.12) for TKA. For ineffective programs, the corresponding WOMAC pain change scores were 0.50 (95% CI, −0.41 to 1.36) and 0.45 (95% CI, −0.18 to 1.08), respectively.

### Function

Twenty-three studies reported the effects of prehabilitation on postoperative function (Table III). Overall, there was a moderate, significant difference favoring prehabilitation for improving function (SMD = 0.34, p < 0.001). The funnel plot was symmetric but positive for significant publication bias (intercept = 2.49, p = 0.011); using a random-effects trim-and-fill model did not change the SMD. The SMDs were similar for...
THA and TKA, significantly favoring prehabilitation for improving function in both patients with hip osteoarthritis (SMD = 0.32, p < 0.001) and knee osteoarthritis (SMD = 0.32, p = 0.015). Heterogeneity was significant in patients with knee osteoarthritis (I² = 66.98%, p < 0.001). After removing a single outlier, the SMD for improvement in function was adjusted from 0.32 to 0.21 and heterogeneity was no longer significant (I² = 43.73%, p = 0.075).

When analyzed with respect to study follow-up, SMDs for improvements in postoperative hip function were similar and significant, favoring prehabilitation for improvement, for longer follow-ups (SMD = 0.31 and 0.39, p = 0.026 and 0.010) but not for follow-up of ≤6 weeks (SMD = 0.16, p = 0.321) (Table III). Heterogeneity was significant for follow-up of ≥6 months (I² = 58.11%, p = 0.026). After removing a single outlier, the SMD at ≥6 months postoperatively decreased from 0.39 to 0.26, and heterogeneity was no longer significant (I² = 11.27%, p = 0.343). SMDs for improvements in postoperative knee function were largest at 3 months postoperatively but did not reach significance (SMD = 0.39, p = 0.059); heterogeneity was significant (I² = 76.32%, p < 0.001). After removing a single outlier, the SMD at 3 months postoperatively decreased from 0.39 to 0.18, and heterogeneity was no longer significant (I² = 45.50%, p = 0.088).

In patients with hip osteoarthritis, a small-to-moderate SMD significantly favored prehabilitation for improvements in function in effective programs (p = 0.001) but not in ineffective programs (p = 0.168) (Table III); heterogeneity was significant for ineffective programs (I² = 67.78%, p = 0.015). After removing a single outlier, the SMD decreased from 0.28 to 0.06, and heterogeneity was no longer significant (I² = 0%, p = 0.826). In patients with knee osteoarthritis, prehabilitation was not significantly favored compared with controls for improvements in function in effective or ineffective programs (p = 0.053 and 0.191) (Table III). The SMD for effective programs was moderate and approached significance (p = 0.053); heterogeneity was significant (I² = 79.45%, p = 0.001). After removing a single outlier, the SMD decreased from 0.53 to 0.20, and heterogeneity was no longer significant (I² = 0%, p = 0.416).

The pooled SMD of 0.34 translated into a change of 5.57 (95% CI, 2.78 to 8.35) out of 68 in the WOMAC function score. For effective programs, the WOMAC function change scores were 6.06 (95% CI, 2.95 to 9.01) for THA and 8.68 (95% CI, 2.06 to 17.36) for TKA. For ineffective programs, the WOMAC function change scores were 4.58 (95% CI, 1.97 to 7.21) for THA and 2.95 (95% CI, 0.16 to 5.80) for TKA.

**Quadriceps Strength**

Eight studies reported the effects of prehabilitation on postoperative quadriceps strength (Table IV).
Overall, there was a moderate, significant difference favoring prehabilitation for improving quadriceps strength (SMD = 0.29, p = 0.042) (Table IV). The funnel plot was symmetric and negative for publication bias (intercept = 0.51, p = 0.813). Heterogeneity was significant ($I^2 = 48.55\%, p = 0.049$); this was reduced when patients with hip and knee osteoarthritis were evaluated separately. The SMD was larger and significant in patients with knee osteoarthritis (SMD = 0.42, p = 0.002), but not significant in patients with hip osteoarthritis (SMD = −0.28, p = 0.241). There were too few studies reporting quadriceps strength for further THA subgroup analyses.

When analyzed with respect to study follow-up, the SMD for postoperative quadriceps strength was large and significant at 3 months postoperatively (SMD = 0.59, p = 0.006) but not significant at ≥6 months (SMD = 0.23, p = 0.130).

When analyzed on the basis of program effectiveness, prehabilitation was moderately favored for improvements in quadriceps strength in both program subgroups for patients with knee osteoarthritis, although only the p value for ineffective programs reached significance (SMD = 0.48 and p = 0.098 for effective programs; SMD = 0.41 and p = 0.011 for ineffective ones) (Table IV). Although the SMD was slightly larger for effective programs, the p value was not significant; heterogeneity was significant ($I^2 = 71.38\%, p = 0.030$). After removing a single outlier$^{66}$, the SMD decreased from 0.48 to 0.19, and heterogeneity was no longer significant ($I^2 = 0\%, p = 0.561$).

Hamstring Strength
Three studies$^{39,65,66}$ reported hamstring strength; these included only patients with knee osteoarthritis (Table IV). Overall, there was a large but nonsignificant difference favoring prehabilitation for improving hamstring strength (SMD = 1.32, p = 0.132). The funnel plot was symmetric and negative for publication bias (intercept = 1.27, p = 0.165). Heterogeneity was significant ($I^2 = 95.91\%, p < 0.001$). After removing a single outlier$^{66}$, the SMD decreased from 1.32 to 0.14, and heterogeneity was no longer significant ($I^2 = 11.63\%, p = 0.287$). There were too few studies reporting hamstring strength for further subgroup analyses.

Anxiety
Six studies reported the effects of prehabilitation on postoperative anxiety (Table V)$^{33,37,41,48,63,67}$. Overall, there was a small, nonsignificant difference favoring prehabilitation for improving postoperative anxiety (SMD = 0.06, p = 0.723) (Table V). The funnel plot was symmetric and negative for publication bias (intercept = 1.97, 95% CI = −7.54 to 11.48, p = 0.596). The SMDs were similar in magnitude but in opposite directions for patients with knee osteoarthritis (SMD = −0.23, p = 0.492) and hip osteoarthritis (SMD = 0.17, p = 0.128). Heterogeneity was...
large and significant ($I^2 = 59.99\%$, $p = 0.029$), and likely due to the heterogeneity among TKA studies ($I^2 = 76.95\%$, $p = 0.037$). There were too few studies reporting anxiety for planned TKA subgroup analyses. At ≤6 weeks postoperatively, prehabilitation was associated with a small improvement in anxiety in THA patients, but this did not did not reach significance ($SMD = 0.20$, $p = 0.089$). At a follow-up duration of ≥6 months, a
small-to-moderate difference favoring prehabilitation was likewise nonsignificant (SMD = 0.34, p = 0.156).

When analyzed with respect to intervention effectiveness, the SMD for effective programs in THA patients was small and nonsignificant (SMD = 0.20, p = 0.178) (Table V). Only 1 THA study had an ineffective program, which was therefore not analyzed further. Both studies involving TKA reported ineffective programs for patients with knee osteoarthritis (Table V). The SMD for anxiety in these studies did not favor prehabilitation (SMD = -0.23, p = 0.492) and had high heterogeneity (I² = 76.95%, p = 0.037), but planned subgroup analyses could not be completed because the number of studies was too small.

**Length of Stay**

Nineteen studies reported the effects of preoperative programs on postoperative length of stay (Table VI). Overall, there was a small-to-moderate, significant difference favoring prehabilitation (SMD = 0.37, p = 0.001) for decreasing hospital length of stay. The funnel plot was symmetric and negative for publication bias (intercept = 0.58, 95% CI, -2.62 to 3.77, p = 0.707). Heterogeneity was significant (p < 0.001). When THA and TKA were evaluated separately, there was a similar SMD in patients with hip osteoarthritis (SMD = 0.31, p = 0.027) and a higher SMD in patients with knee osteoarthritis (SMD = 0.54, p < 0.001); heterogeneity remained elevated (p < 0.001).

Neither effective nor ineffective programs significantly improved the hospital length of stay for patients with hip osteoarthritis (p = 0.237 and 0.057) (Table VI). However, for patients with knee osteoarthritis, the decrease in length of stay was moderate and significant for both effective (SMD = 0.57, p = 0.012) and ineffective programs (SMD = 0.52, p = 0.036).

**Discussion**

This systematic review and meta-analysis summarizing evidence from 35 randomized controlled trials provides greater clarity regarding the role of prehabilitation for patients undergoing arthroplasty. Our results suggest that preoperative programs provide improvements in postoperative outcomes; however, effect sizes varied depending on the joint, follow-up duration, and program effectiveness.

Previous reviews reported conflicting results and were based on few, low-quality studies. Wang et al. found small improvements in pain and function that were similar for THA and TKA, but were not maintained beyond 3 months. This finding is consistent with our results, although improvements in function were maintained beyond 6 months after THA. Additionally, our results suggest larger improvements in outcomes if the program was effective preoperatively. Unfortunately, findings were less consistent for patients undergoing TKA. Previous evidence suggests that ~20% of patients are unsatisfied with pain and function after TKA. Proposed reasons for greater dissatisfaction following TKA may include preoperative expectations of pain and function after surgery; therefore, in addition to the importance of targeted preoperative exercise programs, future research should focus on targeted educational programs to address patient expectations.

Although we found significant SMDs for pain and function, interpretation of these findings is challenging without knowledge of the clinical importance of the improvement. Converting our effect sizes into WOMAC change scores for pain (scale, 0 to 20) and function (scale, 0 to 68) suggested negligible differences between groups in WOMAC pain, and this difference was consistent across all analyses. There were, however, larger differences between groups in WOMAC function scores across all studies (5.57), with the largest differences occurring in studies with effective preoperative programs (6.06 for THA, 8.68 for TKA). Interestingly, the upper limit of the 95% CI

---

| TABLE VI Standardized Mean Differences (SMDs), Compared with Controls, and Heterogeneity (I²) in Studies Reporting Hospital Length of Stay |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                  | SMD             | 95% CI          | Z               | P Value         | I² (%)          | P Value         |
| All studies (n = 19)            | 0.37            | 0.16, 0.57      | 3.46            | 0.001           | 76.53           | <0.001          |
| THA only (n = 12)               | 0.31            | 0.04, 0.59      | 2.21            | 0.027           | 77.79           | <0.001          |
| TKA only (n = 9)                | 0.54            | 0.24, 0.84      | 3.52            | <0.001          | 78.92           | <0.001          |
| Preop. effectiveness of the intervention |
| Effective                      |
| THA (n = 7)                    | 0.12            | -0.08, 0.31     | 1.18            | 0.237           | 87.78           | <0.001          |
| TKA (n = 5)                    | 0.57            | 0.13, 1.02      | 2.52            | 0.012           | 76.88           | 0.002           |
| Ineffective                    |
| THA (n = 5)                    | 0.58            | -0.18, 1.18     | 1.9             | 0.057           | 26.83           | 0.224           |
| TKA (n = 4)                    | 0.52            | 0.03, 0.99      | 2.10            | 0.036           | 85.32           | <0.001          |
for patients who received an effective program and underwent TKA may be considered clinically important. However, because of the relatively large CIs, considerable uncertainty remains.

Clinical interpretation of these findings should also account for the potential effects of performance bias. Providing care beyond “standard or usual care” may positively influence patient perceptions and outcome measures of interest; however, blinding patients to exercise by including a sham intervention would have been challenging, and no studies in this review included a third randomized arm. Objective measures such as strength may be less influenced by biases, but despite numerous randomized controlled trials incorporating a strength-training component in the prehabilitation, few studies (primarily on TKA) included measures of postoperative strength. Among TKA patients, improvements were observed in quadriceps strength at, but not beyond, 3 months. Notably, our results show similar quadriceps strength differences between groups in studies with effective and ineffective programs.

Strategies for decreasing length of stay are gaining much attention from hospital administrators and policy-makers. The findings of this review suggest favorable decreases in length of stay for patients receiving prehabilitation, with larger decreases in patients undergoing TKA compared with THA. These effects appear to be independent of program effectiveness, suggesting that additional patient-specific factors may play a role. None of the included studies directly measured patient costs; however, patients with greater disease severity, poorer function, and greater pain typically have greater health-care resource use. Similarly, Lavernia et al. investigated the effect of disease severity on the cost-effectiveness of THA and found that worse preoperative WOMAC scores were associated with a less cost-effective intervention. Therefore, prehabilitation for improving pain and function may influence overall health-care costs associated with THA and TKA.

Patient anxiety may also affect outcomes and quality of life following arthroplasty; however, the effectiveness of prehabilitation with respect to postoperative anxiety remains unclear. Isolating the true effects of prehabilitation on behavioral or psychological metrics is difficult. Specifically, the effects of prehabilitation may be masked by previously established patient expectations about exercise effectiveness, hindering any potential benefits that exercise may have for the patient. Clinical trials inconsistently report the mental state or beliefs of study participants, and this requires further investigation. Aydin et al. reported positive effects of prehabilitation on preoperative anxiety but no effect on postoperative anxiety. Similar findings by McDonald et al. suggested that stratification is needed for patients with depression, those with anxiety, or other subgroups not likely to receive benefit. The ability to perform this type of analysis hinges on better reporting of the patient characteristics most applicable to the outcome of interest. Consideration should also be given to additional patient-specific factors that may confound potential effects of prehabilitation during the postoperative period, such as family support, resource availability, and discharge planning.

Limitations and Future Research

To our knowledge, this study is the first to investigate the impact of preoperative prehabilitation program effectiveness on postoperative outcomes, highlighting an important area for future research. The duration, frequency, and type of exercise and the level of supervision are known contributors to the effectiveness of prehabilitation; yet, are often not reported. Limitations of the individual studies include large differences between the interventions tested, poor-quality interventions, and minimal reporting on compliance and other factors potentially impacting the findings (e.g., comorbidities and pain management). A standardized approach for reporting clinical trial interventions is warranted. Future research should also consider better reporting of patient compliance with prehabilitation. Particularly in patients with osteoarthritis, pain and other associated comorbidities may be a primary barrier to participation in prehabilitation activities, consequently limiting the potential improvements in postoperative outcomes, and this warrants further investigation to better understand how various patient subgroups respond to prehabilitation. For the overall review, inclusion of all studies regardless of their risk-of-bias level may be a limitation. The effects of high and low-quality studies can influence the overall effect (i.e., removing outliers decreased SMDs); therefore, SMDs are reported both with and without the outlier. Future research should focus on the role of prehabilitation in the overall rehabilitation process following arthroplasty, with specific investigations regarding functional exercise (exercise that is more easily translated to activities of daily living): optimal dosage, intensity, and type of exercise; exercise programs using a patient-centered approach; supervised versus unsupervised programs; and whether prehabilitation results in faster return to work or recreational activities.

Conclusions

Results from this systematic review and meta-analysis suggest that prehabilitation programs provide small-to-moderate improvements that vary by joint. In patients undergoing THA, significant improvements were observed for pain, function, and length of stay. In patients undergoing TKA, significant improvements were observed for function, quadriceps strength, and length of stay. Future clinical trials should collect a standardized regimen of outcomes, including standardized postoperative time points for the data collection follow-ups. Further investigation into factors contributing to effective exercise and targeted educational programs to improve postoperative patient expectations may result in clinically important improvements.
The Value of Preoperative Exercise and Education for Patients Undergoing THA and TKA

Appendix

The MEDLINE search strategy and a table summarizing the risk of bias in each included study are available with the online version of this article as a data supplement at jbj.org (http://links.lww.com/JBJSREV/A286).

NOTE:
The authors thank Glenn Wark and Darren Bittner for their assistance with data collection and analysis and with initial preparation of the manuscript.

Kathy Ikert, PT2,
initial preparation of the manuscript.
The authors thank Glenn Wark and


