Factors Correlated With Physical Function 1 Year After Total Knee Arthroplasty in Patients With Knee Osteoarthritis
A Systematic Review and Meta-analysis

Unni Olsen, RN, MSC; Maren Falch Lindberg, RN, PhD; Christopher Rose, CStat, PhD; Eva Denison, PT, PhD; Caryl Gay, Psych, PhD; Arild Aamodt, MD, PhD; Jens Ivar Brox, MD, PhD; Øystein Skare, PI, PhD; Ove Furnes, MD, PhD; Kathryn Lee, RN, PhD; Anners Lerdal, RN, PhD

Abstract

IMPORTANCE More than 1 in 5 patients do not experience improved physical function after total knee arthroplasty (TKA). Identification of factors associated with physical function may be warranted to improve outcomes in these patients.

OBJECTIVE To identify preoperative and intraoperative factors associated with physical function at 12 months after TKA in a systematic review and meta-analysis.

DATA SOURCES Data from January 2000 to October 2021 were searched in Medline, Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane Library, and Physiotherapy Evidence Database (PEDro). No language restrictions were applied.

STUDY SELECTION Prospective observational studies or randomized clinical trials on factors associated with physical function after TKA in adult patients with osteoarthritis were selected. A prespecified peer-reviewed protocol was followed.

DATA EXTRACTION AND SYNTHESIS Following the Preferred Reporting Items for Systematic Reviews and Meta-analyses guideline, 2 reviewers independently screened titles and abstracts and judged risk of bias using Quality in Prognosis Studies (QUIPS). Multivariate random-effects meta-analyses were performed to estimate mean correlations between factors and physical function with 95% CIs. Sensitivity analyses were conducted for each QUIPS domain. Certainty of evidence was evaluated using Grading of Recommendations, Assessment, Development and Evaluations (GRADE). This study was registered with the International Prospective Register of Systematic Reviews (PROSPERO).

MAIN OUTCOMES AND MEASURES The primary outcome was physical function 12 months after TKA. Secondary outcomes were physical function 3 and 6 months after TKA. All estimates are mean correlations between factors and postoperative function. Positive correlations correspond to better function.

RESULTS Among 12,052 articles, 20 studies (including 11,317 patients and 37 factors) were analyzed. Mean correlation with higher BMI was estimated to be −0.15 (95% CI, −0.24 to −0.05; P = .33; moderate-certainty evidence), while mean correlation with better physical function was estimated to be 0.14 (95% CI, 0.02 to 0.26; P = .03; low-certainty evidence) and mean correlation with more severe osteoarthritis was estimated to be 0.10 (95% CI, 0.01 to 0.19; P = .17; high-certainty evidence). In sensitivity analyses, mean correlation with better physical function was estimated to be 0.20 (95% CI, 0.04 to 0.36; P = .02), and so perhaps a larger coefficient than in the main analysis, while mean

Key Points

Question What preoperative and intraoperative factors are correlated with physical function after total knee arthroplasty (TKA)?

Findings In this systematic review and meta-analysis of 20 studies that included 11,317 patients with osteoarthritis, higher preoperative body mass index (BMI) was correlated with worse physical function, while better preoperative physical function and more severe osteoarthritis were correlated with better physical function 1 year after TKA.

Meaning These findings suggest that presurgical BMI, physical function, and osteoarthritis severity may be important factors to include and test in models predicting TKA outcomes.

Supplemental content

Author affiliations and article information are listed at the end of this article.
correlations were estimated to be similar for other factors (BMI: −0.17; 95% CI, −0.28 to −0.06; $P < .001$; osteoarthritis severity: 0.10; 95% CI, −0.01 to 0.20; $P = .05$).

CONCLUSIONS AND RELEVANCE This study found that higher presurgical BMI was correlated with worse physical function (with moderate certainty) and that better physical function (low certainty) and osteoarthritis severity (high certainty) were correlated with better physical function after TKA. These findings suggest that these factors should be included when testing predictive models of TKA outcomes.


Introduction

Total knee arthroplasty (TKA) has become the third most common inpatient surgery in the United States, with 750,000 yearly procedures projected to double in the next decade.¹² TKA is regarded as a cost-efficient and effective treatment for restoring physical function in patients with end-stage osteoarthritis.³ However, more than 1 in 5 patients do not regain physical function after TKA.⁴ Nonimprovement of physical function is a risk factor associated with more expensive revision surgery and an immense burden at individual, health care system, and socioeconomic levels.⁵,⁶

Factors identified in predictive models using high-quality evidence could improve patient outcomes, particularly for those who are unlikely to benefit from surgery or who have unrealistic expectations. Evidence on factors associated with physical function has been reviewed previously, but findings were contradictory, limited in scope, based on pooled data across short-term and longer-term outcomes, or did not address certainty of evidence.⁷,¹³ Thus, there is need for a new synthesis of evidence on short-term TKA outcomes that uses current systematic review methods and captures recently published studies. The aim of this systematic review and meta-analysis was to synthesize evidence on preoperative and intraoperative factors associated with physical function 12 months after TKA (primary outcome) and 3 and 6 months after TKA (secondary outcomes).

Methods

In this systematic review and meta-analysis, we followed a prespecified peer-reviewed protocol¹⁴ and a preprint¹⁵ registered in International Prospective Register of Systematic Reviews (PROSPERO; CRD42018079069), designed and conducted according to Cochrane Handbook guidelines.¹⁶ Results are reported according to the recently revised Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline.

Search Strategy and Data Sources

The search strategy was collaboratively developed by researchers (U.O. and M.F.L.) and research librarians, with feedback from the research team.¹⁴ Published studies from January 1, 2000, to October 8, 2021, were systematically searched, with no language restrictions, in Medline (Ovid), Embase (Ovid), Cumulative Index to Nursing and Allied Health Literature (CINAHL; EBSCO), Cochrane Library, and Physiotherapy Evidence Database. References were managed using Endnote X8 software version 20.2.1 (Clarivate Analytics). Subject headings and keywords for each database are described in eTable 5 in the Supplement, and full search strategies for each database are defined in the protocol.¹⁴
Eligibility Criteria
To be maximally inclusive, studies had to include estimates of association between preoperative or intraoperative factors and physical function at 3, 6, or 12 months after TKA. We considered studies eligible if participants were adults diagnosed with osteoarthritis scheduled for primary TKA. Prospective longitudinal observational studies and randomized clinical trials that provided sufficient estimates of association were eligible. We excluded retrospective and case-control studies, as well as conference abstracts. We also excluded studies with mixed patient populations (eg, rheumatoid arthritis, total hip arthroplasty, or unicompartmental arthroplasty) if separate outcome data were not reported for osteoarthritis and TKA.

Outcomes
The primary outcome was physical function at 12 months after TKA. Secondary outcomes were physical function 3 and 6 months after TKA.

Study Selection and Data Extraction
Data from included studies were extracted to a standardized extraction form, with details in the published protocol. Data included study design, sample size, country, age, sex, body mass index (BMI [calculated as weight in kilograms divided by height in meters squared]), outcome measures used, data collection time points, statistical analyses, and estimates of association. One reviewer performed data extraction (U.O.), while another reviewer checked data accuracy against source material (M.F.L.). Two reviewers (U.O. and M.F.L.) evaluated titles and abstracts for applicability, then read and checked full-text publications against eligibility criteria. Another author (E.D.) was involved in resolving disagreements.

Methodological Quality
Risk of bias was assessed using the Quality in Prognosis Studies (QUIPS) tool, following the strategy described in the protocol, in which 2 reviewers (U.O. and M.F.L.) independently assessed risk of bias and had consensus discussions before arriving at consensus. In cases of disagreement, E.D. was involved in the final decision. QUIPS has 6 risk domains: study participation, attrition, prognostic factor measurement, statistical analysis and reporting, confounding, and outcome measurement.

Certainty of Evidence
Two researchers (U.O. and M.F.L.) rated certainty of evidence by consensus discussion using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) framework. In some cases, a third researcher (E.D.) was involved in discussions. Certainty of evidence was graded as high, moderate, low, or very low. We used GRADEpro GDT (McMaster University) to summarize evidence.

Statistical Analysis
Findings for all included studies were synthesized by outcomes at 3, 6, or 12 months after TKA as described in the protocol. We were unable to complete planned multivariate random-effects meta-analysis because extracted data were too sparse (with a large number of factors reported by relatively few studies). Accordingly, we used a frequentist version of the bayesian multivariate model. Additional protocol deviations are explained in eMethods in the Supplement.

To quantify associations between potential factors and the outcome, we extracted odds ratios (ORs), risk ratios (RRs), linear model coefficients (including differences), or correlations using discrete or continuous scales. We meta-analyzed hyperbolic arc tangent-transformed correlation coefficients, which under reasonable assumptions can be imputed for these measures of association and are invariant under linear transformation. This approach allowed inclusion of studies using various measurement tools and analyses in the meta-analysis.
We anticipated that studies would use different instruments and statistical methods that could lead to between-study heterogeneity. Therefore, multivariate random-effects meta-analysis was conducted to estimate mean correlations (ie, not common correlations) between factors and postoperative physical function.

Heterogeneity was quantified using $I^2$ statistics. We used $P$ scores that measured the certainty that the mean correlation for a factor was larger than those for all other factors.\(^21\) We also performed exploratory univariate meta-analyses and multivariate meta-analyses (after removing factors supported by few studies to reduce the problem of sparsity of estimation). Estimates from 3 models were compared for consistency. Finally, sensitivity analyses on physical function at 12 months after TKA were conducted for each QUIPS domain by excluding studies judged as high risk of bias and rerunning multivariate meta-analysis.

Statistical analyses were performed using Stata statistical software version 16 (StataCorp). We report mean correlations with 95% CIs. We did not prespecify any hypothesis testing but report 2-sided $P$ values for completeness.

### Results

The Figure 1 study flow diagram outlines study selection and reasons for exclusion.\(^22-41\) From 12 052 articles screened for title and abstracts, 391 articles were selected for full-text examination, with 20 studies\(^22-41\) (total sample = 11 317 patients) for qualitative analysis at 3, 6, and 12 months and 17 studies\(^22-33,35-38,41\) for quantitative analysis at 6 and 12 months. Individual study characteristics are detailed in the Table.\(^22-41\) All were prospective longitudinal observational designs; no randomized trial met inclusion criteria. We identified 37 factors across 20 studies. There were 8 studies\(^26-30,34,37,38\) conducted in Europe, 6 studies\(^24,31-33,39,40\) in Asia, 4 studies\(^25,35,36,41\) in North America, and 1 study\(^23\) in Australia, and 1 study\(^23\) was multicontinental (ie, Australia, Europe, and North America). Sample sizes ranged from 49 patients\(^36\) to 5309 patients.\(^32\) Mean age varied from 63 years\(^35\) to 75 years,\(^32\) and representation of women ranged from 49.3%\(^36\) to 90.0%.\(^32\) The most common physical function measure was the Western Ontario and McMaster Universities Arthritis Index (WOMAC). We excluded 6 studies from analysis.\(^42-47\) owing to unsuccessful attempts to obtain missing data. Sedentary behavior,\(^40\) lack of energy,\(^38\) drowsiness,\(^38\) and sleeping difficulties.\(^38\)

### Figure 1. Flowchart of Included Studies

- **19221** Records identified through database searching
- **7169** Duplicates removed
- **12052** Records screened
- **11661** Records excluded after evaluation of title and abstract
- **391** Full-text articles assessed for eligibility
- **371** Full-text articles excluded
  - 90 No regression performed
  - 80 Conference abstract, duplicate, or commentary
  - 54 Inadequate study design or aim
  - 21 Prognostic factor or outcome not evaluated
  - 94 TKA and OA pooled results or total score
  - 20 Insufficient follow-up time
  - 12 Insufficient data or age <18 y

**OA** indicates osteoarthritis; **TKA**, total knee arthroplasty.
| Source                | Country          | Design | Patients analyzed, No. | Data collection | Follow-up, mo | Baseline age, y | Patients, No./Total No. (%) | Analysis                        | Factors measured                        | Outcome measured                        |
|-----------------------|------------------|--------|------------------------|-----------------|---------------|-----------------|-------------------------------|----------------------------------------|-----------------------------------------|
| Berghmans et al,2019  | Netherlands      | PC     | 146                    | NA              | 3             | Mean, 66.4      | 79/150 (53)                  | Stepwise multiple linear regression    | Mental health (SF-36), physical function (WOMAC), knee stiffness (WOMAC) | WOMAC                                    |
| Lindner et al,2018    | Germany          | PC     | 61                     | NA              | 3             | Mean, 67        | 37/61 (61)                  | Stepwise multiple linear regression    | Pain (WOMAC)                           | WOMAC                                    |
| Luo et al,2019        | China            | PC     | 471                    | 2017-2018       | 3             | Mean, 64.3      | 357/471 (76)                 | Pearson correlation                  | Sleep dysfunction (PSQI), day time sleepiness (ESS), sleep quality (self-developed scale (0-10)) | KSS                                      |
| Bugada et al,2017     | Italy            | PC     | 563                    | 2012-2015       | 6             | Median, 72      | 421/606 (69)                | Logistic regression                  | Comorbiditity (ASA Physical Status Classification System) | NRS                                      |
| Engel et al,2007      | US               | PC     | 54                     | NA              | 6             | Mean, 68        | 26/74 (49%)                 | Multiple hierarchical regression      | AHI                                     | WOMAC                                    |
| Escobar et al,2007    | Spain            | PC     | 640                    | 1999-2000       | 6             | Mean, 72        | 473/471 (74%)               | General linear model                 | Age (y), sex (men/women), social support (yes/no), comorbiditiy (CCI), physical function (WOMAC), low back pain (yes/no), mental health (SF-36) | WOMAC                                    |
| Hylkema et al,2019    | US               | PC     | 131                    | 2012-2014       | 6             | Mean, 61        | 114/183 (62)                | Univariate linear regression         | Pain catastrophizing (PCS)            | WPAI-SHP                                |
| Oka et al,2019        | Japan            | PC     | 82                     | 2017-2019       | 6             | Mean, 72.1      | 67/82 (82)                  | Multiple linear regression           | Sedentary behavior (MET)              | New KSS                                 |
| Pua et al,2019        | Singapore        | PC     | 4026                   | 2013-2017       | 6             | Mean, 68        | 3003/4026 (75)             | Proportional odds ordinal regression | Age (y), sex (men/women), BMI, education (primary, secondary, tertiary), ethnicity (Chinese, Indian, Malay, other), social support (yes/no), comorbiditiy (yes/no) contralateral knee pain (KSS), pain (OKQ), knee extension and flexion (goniometer), physical function (categoties), depression (SF-36) | OKQ                                      |
| Sugawara et al,2017   | Japan            | PC     | 59                     | 2011-2012       | 6             | Mean, 75        | 53/59 (90)                  | Stepwise multiple regression         | TSL, S                                 | JKOM                                     |
| Taniuchi et al,2016   | Japan            | PC     | 81                     | 2013-2014       | 6             | Mean, 72        | 73/81 (90)                  | Multiple linear regression           | TUG                                     | TUG                                      |
| Yang et al,2019       | US               | PC     | 107                    | 2010-2011       | 6             | Mean, 65        | 55/107 (51)                 | Multivariate logistic regression     | Mental health (SF-36), pain catastrophizing (PCS), comorbiditiy (CCI), use device (yes/no) | WOMAC                                    |
| Berghmans et al,2019  | Netherlands      | PC     | 144                    | NA              | 12            | Mean, 66.4      | 79/150 (53)                 | Stepwise multiple linear regression  | Physical function (WOMAC), knee function (KSS) | WOMAC                                    |

(continued)
<table>
<thead>
<tr>
<th>Source</th>
<th>Country</th>
<th>Design</th>
<th>Patients analyzed, No.</th>
<th>Data collection</th>
<th>Follow-up, mo</th>
<th>Baseline age, y</th>
<th>Patients, No./Total No. (%)</th>
<th>Analysis</th>
<th>Factors measured</th>
<th>Outcome measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dowsey et al.</td>
<td>Australia</td>
<td>PC</td>
<td>473</td>
<td>2006-2007</td>
<td>12</td>
<td>Mean, 71</td>
<td>331/478 (69) 142/478 (31)</td>
<td>Multivariate linear regression</td>
<td>Age (y), sex (men/women), BMI, comorbidity (CCI), pain (IKSS), physical function (IKSS), mental health (SF-12), K-L grade, cruciate retaining, patella resurface, multicompartement OA</td>
<td>IKSS</td>
</tr>
<tr>
<td>Lindberg et al.</td>
<td>Norway</td>
<td>PC</td>
<td>182</td>
<td>2012-2014</td>
<td>12</td>
<td>Mean, 67</td>
<td>124/182 (68) 58/182 (32)</td>
<td>Multivariate logistic regression</td>
<td>Age (y), sex (men/women), pain (BPI), lack of energy, drowsiness, sleeping difficulties, bloating, worrying, sexuality problems (MSAS-SF)</td>
<td>BPI</td>
</tr>
<tr>
<td>Lingard et al.</td>
<td>UK, US, Canada,</td>
<td>PC</td>
<td>676</td>
<td>1997-1999</td>
<td>12</td>
<td>Distress: median, 70</td>
<td>574/676 (85) 102/676 (15)</td>
<td>Logistic regression</td>
<td>Psychological distress (SF-36) WOMAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nankaku et al.</td>
<td>Japan</td>
<td>PC</td>
<td>115</td>
<td>2013-2015</td>
<td>12</td>
<td>Mean, 71</td>
<td>99/115 (86) 16/115 (14)</td>
<td>Stepwise multiple regression</td>
<td>Age (y), physical function (KSS), TUG</td>
<td>KSS</td>
</tr>
<tr>
<td>Sullivan et al.</td>
<td>Canada</td>
<td>PC</td>
<td>120</td>
<td>NA</td>
<td>12</td>
<td>Mean, 67</td>
<td>73/120 (61) 47/120 (39)</td>
<td>Multiple regression</td>
<td>Age (y), sex (men/women), BMI, comorbidity (CCI), physical function and pain (WOMAC), pain catastrophizing (PCS), depression (PHQ-9), kinesiophobia (TSK), surgery duration (min)</td>
<td>WOMAC</td>
</tr>
<tr>
<td>Tilbury et al.</td>
<td>Netherlands</td>
<td>PC</td>
<td>146</td>
<td>2011-2012</td>
<td>12</td>
<td>Mean, 67</td>
<td>101/146 (69) 87/146 (31)</td>
<td>Multivariate linear regression</td>
<td>BMI, mental health (SF-36), physical function (KOOS), outcome expectancies (HSS hip replacement and knee replacement expectations surveys)</td>
<td>KOOS</td>
</tr>
<tr>
<td>Wylde et al.</td>
<td>UK</td>
<td>PC</td>
<td>220</td>
<td>NA</td>
<td>12</td>
<td>Median, 70</td>
<td>136/220 (62) 84/220 (38)</td>
<td>Ordinary least square regression</td>
<td>Age (y), sex (men/women), comorbidity (SCQ), physical function (WOMAC), depression and anxiety (HADS), pain self-efficacy (PSEQ)</td>
<td>WOMAC</td>
</tr>
</tbody>
</table>

Abbreviations: AHI, Arthritis Helplessness Index; ASA, American Society of Anesthesiologists; BMI, body mass index; BPI, Brief Pain Inventory; CCI, Charlson Comorbidity Index; ESS, Epworth Sleepiness Scale; HADS, Hospital Anxiety and Depression Scale; HSS, Hospital for Special Surgery; IKSS, International Knee Society Score; KOOS, Knee Outcome and Depression Score; KSS, Knee Society Clinical Rating System; MET, Metabolic Equivalent of Tasks; MSAS-SF, Memorial Symptom Assessment Scale Short Form; NA, not applicable; NRS, numerical rating scale; OKQ, Oxford Knee Questionnaire; PC, prospective cohort; PCS, Pain Catastrophizing Scale; PHQ-9, Patient Health Questionnaire; PSEQ, Pain Self-Efficacy Questionnaire; PSQI, Pittsburgh Sleep Quality Index; SCQ, Self-Administered Comorbidity Questionnaire; SF-12, 12-Item Short-Form Health Survey; SF-36, 36-Item Short Form Health Survey; TK, Tampa Scale of Kinesiophobia; TULS, time single legged stand with eyes open; TUG, Timed Up and Go; WOAI-SHP, Work Productivity and Activity Impairment Questionnaire; Specific Health Problem; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

* Study with 2 follow-up times.
bloating, worrying, and problems with sexuality were reported once and were not included in the meta-analysis.

Estimates of correlations of factors with function are reported separately for 6-month and 12-month outcomes (Figure 2 and Figure 3). Results from 2 or more studies that could be statistically combined in multivariate meta-analysis are reported subsequently. Explorations of sensitivity analysis are in eFigure 1 and eTable 1 in the Supplement, while explorations of potential inconsistencies and results from exploratory univariate meta-analyses are in eFigures 2 and 3 in the Supplement. Labels for included factors are defined in eTable 3 and reason for exclusion of the individual studies are described in eTable 6 in the Supplement. Positive correlations correspond to better function postoperatively.

There were 9 studies with 2637 patients that reported estimates for 25 potential factors for our primary outcome, physical function at 12 months after TKA. Preoperative function (6 studies), mental health (including anxiety, depression, and psychological distress [5 studies]), and age (5 studies) were the most frequently reported factors. Several studies were judged as at high risk of bias on 1 or more domains (Figure 4).

Mean correlation with higher BMI was estimated to be −0.15 (95% CI −0.24 to −0.05; \( P = .33 \); P score = 70.0%; 3 studies; moderate-certainty evidence and moderate heterogeneity among reported estimates of association \( I^2 = 46\% \)). Mean correlation with better physical function was estimated to be 0.14 (95% CI, 0.02 to 0.26; \( P = .03 \); P score = 65.6%; 6 studies; low-certainty evidence and substantial heterogeneity among estimates of association \( I^2 = 90\% \)), while mean correlation with better mental health was estimated to be 0.12 (95% CI, −0.01 to 0.25; \( P = .10 \); P score = 60.0%; 5 studies; moderate-certainty evidence and substantial heterogeneity among reported estimates of association \( I^2 = 67\% \)) and mean correlation with more severe osteoarthritis was estimated to be 0.10 (95% CI, 0.01 to 0.19; \( P = .17 \); P score = 53.8%; 2 studies; high-certainty evidence and heterogeneity between reported estimates \( I^2 = 26\% \)).

**Figure 2. Forest Plot of Factors Associated With Physical Function at 12 mo**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Studies, No.</th>
<th>Correlation (95% CI)</th>
<th>Correlated with worse function</th>
<th>Correlated with better function</th>
<th>( P ) value</th>
<th>P score, %</th>
<th>( I^2 ), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>1</td>
<td>0.25 (0.05 to 0.44)</td>
<td></td>
<td></td>
<td>.02</td>
<td>87.4</td>
<td>NA</td>
</tr>
<tr>
<td>Pain self-efficacy</td>
<td>1</td>
<td>0.15 (-0.01 to 0.31)</td>
<td></td>
<td></td>
<td>.048</td>
<td>68.1</td>
<td>NA</td>
</tr>
<tr>
<td>Knee status</td>
<td>1</td>
<td>0.14 (-0.08 to 0.35)</td>
<td></td>
<td></td>
<td>.28</td>
<td>63.3</td>
<td>NA</td>
</tr>
<tr>
<td>Preoperative function</td>
<td>6</td>
<td>0.14 (0.02 to 0.26)</td>
<td></td>
<td></td>
<td>.03</td>
<td>65.6</td>
<td>89.6</td>
</tr>
<tr>
<td>Mental health</td>
<td>5</td>
<td>0.12 (-0.01 to 0.25)</td>
<td></td>
<td></td>
<td>.10</td>
<td>60.0</td>
<td>67.4</td>
</tr>
<tr>
<td>K-L grade</td>
<td>2</td>
<td>0.10 (0.01 to 0.19)</td>
<td></td>
<td></td>
<td>.17</td>
<td>53.8</td>
<td>25.9</td>
</tr>
<tr>
<td>Cruciate retaining</td>
<td>2</td>
<td>0.09 (-0.03 to 0.21)</td>
<td></td>
<td></td>
<td>.21</td>
<td>48.6</td>
<td>NA</td>
</tr>
<tr>
<td>Male sex</td>
<td>4</td>
<td>0.05 (-0.04 to 0.15)</td>
<td></td>
<td></td>
<td>.009</td>
<td>36.1</td>
<td>44.3</td>
</tr>
<tr>
<td>Outcome expected</td>
<td>1</td>
<td>0.04 (-0.21 to 0.29)</td>
<td></td>
<td></td>
<td>.93</td>
<td>37.0</td>
<td>NA</td>
</tr>
<tr>
<td>Preoperative pain</td>
<td>4</td>
<td>0.04 (-0.04 to 0.12)</td>
<td></td>
<td></td>
<td>.008</td>
<td>30.6</td>
<td>74.2</td>
</tr>
<tr>
<td>Patella resurfaced</td>
<td>2</td>
<td>0.04 (-0.09 to 0.16)</td>
<td></td>
<td></td>
<td>.005</td>
<td>31.1</td>
<td>NA</td>
</tr>
<tr>
<td>Surgery duration</td>
<td>2</td>
<td>0.0 (-0.23 to 0.23)</td>
<td></td>
<td></td>
<td>.004</td>
<td>26.7</td>
<td>NA</td>
</tr>
<tr>
<td>Kinesophobia</td>
<td>2</td>
<td>0 (-0.23 to 0.23)</td>
<td></td>
<td></td>
<td>.004</td>
<td>26.7</td>
<td>NA</td>
</tr>
<tr>
<td>Multicompartment OA</td>
<td>1</td>
<td>-0.01 (-0.14 to 0.11)</td>
<td></td>
<td></td>
<td>.004</td>
<td>24.2</td>
<td>NA</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>3</td>
<td>-0.02 (-0.12 to 0.07)</td>
<td></td>
<td></td>
<td>.84</td>
<td>25.8</td>
<td>26.1</td>
</tr>
<tr>
<td>Age</td>
<td>5</td>
<td>-0.05 (-0.13 to 0.04)</td>
<td></td>
<td></td>
<td>.009</td>
<td>33.3</td>
<td>73.0</td>
</tr>
<tr>
<td>BMI</td>
<td>2</td>
<td>0 (-0.24 to 0.05)</td>
<td></td>
<td></td>
<td>.33</td>
<td>70.0</td>
<td>46.1</td>
</tr>
<tr>
<td>Symptomatic joints</td>
<td>1</td>
<td>-0.20 (-0.35 to -0.03)</td>
<td></td>
<td></td>
<td>.31</td>
<td>79.1</td>
<td>NA</td>
</tr>
<tr>
<td>Catastrophizing</td>
<td>1</td>
<td>-0.24 (-0.47 to 0.01)</td>
<td></td>
<td></td>
<td>.07</td>
<td>82.8</td>
<td>NA</td>
</tr>
</tbody>
</table>

BMI indicates body mass index; K-L, Kellgren-Lawrence; NA, not applicable; OA, osteoarthritis. Direction of correlation: increased values of factors correlate with better postoperative function for all factors except dichotomous values (ie, cruciate retaining, male sex, patella resurfaced, and multicompartment OA), for which presence of factor correlates with better postoperative function.
High-certainty evidence and heterogeneity for osteoarthritis may not be important. We were unable to conclude that clinically meaningful correlations did not exist for the other 15 factors owing to limited evidence (ie, wide CIs).

In the prespecified sensitivity analysis (eTable 1 in the Supplement), mean correlation with better physical function was estimated to be 0.20 (95% CI, 0.04 to 0.36; \( P = .02 \) vs coefficient = 0.14; 95% CI, 0.02 to 0.26 when including all studies). Mean correlation with BMI was

![Figure 3. Forest Plot of Factors Associated With Physical Function at 6 mo](image)

**Figure 3. Forest Plot of Factors Associated With Physical Function at 6 mo**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Studies, No.</th>
<th>Correlation (95% CI)</th>
<th>Correlated with worse function</th>
<th>Correlated with better function</th>
<th>P value</th>
<th>P score, %</th>
<th>I², %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility³³</td>
<td>1</td>
<td>0.87 (0.77 to 0.93)</td>
<td></td>
<td></td>
<td>&lt;.001</td>
<td>100</td>
<td>NA</td>
</tr>
<tr>
<td>Preoperative function³⁰-³³</td>
<td>3</td>
<td>0.37 (0.27 to 0.46)</td>
<td></td>
<td></td>
<td>&lt;.001</td>
<td>90.4</td>
<td>92.9</td>
</tr>
<tr>
<td>Pain self-efficacy³⁵</td>
<td>1</td>
<td>0.27 (–0.03 to 0.53)</td>
<td></td>
<td></td>
<td>.09</td>
<td>73.1</td>
<td>NA</td>
</tr>
<tr>
<td>Education³¹</td>
<td>1</td>
<td>0.16 (–0.15 to 0.44)</td>
<td></td>
<td></td>
<td>.45</td>
<td>54.1</td>
<td>NA</td>
</tr>
<tr>
<td>Mental health³⁰,³¹,³⁴</td>
<td>3</td>
<td>0.15 (–0.08 to 0.36)</td>
<td></td>
<td></td>
<td>.28</td>
<td>53.5</td>
<td>80.9</td>
</tr>
<tr>
<td>Social support³⁰,³¹</td>
<td>2</td>
<td>0.09 (–0.01 to 0.19)</td>
<td></td>
<td></td>
<td>.09</td>
<td>40.9</td>
<td>0</td>
</tr>
<tr>
<td>Preoperative pain³¹</td>
<td>1</td>
<td>0.06 (–0.17 to 0.28)</td>
<td></td>
<td></td>
<td>.83</td>
<td>32.8</td>
<td>NA</td>
</tr>
<tr>
<td>Arthritis helplessness³⁰</td>
<td>1</td>
<td>0 (-0.35 to 0.35)</td>
<td></td>
<td></td>
<td>&gt;.99</td>
<td>26.3</td>
<td>NA</td>
</tr>
<tr>
<td>Knee extension³¹</td>
<td>1</td>
<td>0.011 (–0.11 to 0.09)</td>
<td></td>
<td></td>
<td>.96</td>
<td>16.4</td>
<td>NA</td>
</tr>
<tr>
<td>Male sex³⁰,³¹</td>
<td>2</td>
<td>0.03 (–0.14 to 0.07)</td>
<td></td>
<td></td>
<td>.74</td>
<td>22.6</td>
<td>83.2</td>
</tr>
<tr>
<td>Knee flexion³¹</td>
<td>1</td>
<td>0.04 (–0.14 to 0.06)</td>
<td></td>
<td></td>
<td>.61</td>
<td>24.8</td>
<td>NA</td>
</tr>
<tr>
<td>BMI²¹</td>
<td>1</td>
<td>0.05 (–0.14 to 0.05)</td>
<td></td>
<td></td>
<td>.53</td>
<td>26.0</td>
<td>NA</td>
</tr>
<tr>
<td>Comorbidity²⁹,³¹</td>
<td>3</td>
<td>0.09 (–0.19 to 0.02)</td>
<td></td>
<td></td>
<td>.13</td>
<td>39.7</td>
<td>5.5</td>
</tr>
<tr>
<td>Age³⁰,³¹</td>
<td>2</td>
<td>0.09 (-0.19 to 0.00)</td>
<td></td>
<td></td>
<td>.06</td>
<td>41.6</td>
<td>95.3</td>
</tr>
<tr>
<td>Low back pain³⁰</td>
<td>1</td>
<td>0.10 (–0.23 to 0.03)</td>
<td></td>
<td></td>
<td>.15</td>
<td>43.4</td>
<td>NA</td>
</tr>
<tr>
<td>Contralateral knee pain³¹</td>
<td>1</td>
<td>0.12 (–0.22 to –0.01)</td>
<td></td>
<td></td>
<td>.03</td>
<td>48.7</td>
<td>NA</td>
</tr>
<tr>
<td>Catastrophizing³⁵,³⁴</td>
<td>2</td>
<td>0.19 (–0.35 to –0.01)</td>
<td></td>
<td></td>
<td>.03</td>
<td>63.4</td>
<td>84.9</td>
</tr>
<tr>
<td>Indian ethnicity³¹</td>
<td>1</td>
<td>0.20 (–0.35 to –0.05)</td>
<td></td>
<td></td>
<td>.005</td>
<td>68.1</td>
<td>NA</td>
</tr>
<tr>
<td>Walking aid use³¹,³⁴</td>
<td>2</td>
<td>0.31 (–0.45 to –0.17)</td>
<td></td>
<td></td>
<td>&lt;.001</td>
<td>84.3</td>
<td>63.5</td>
</tr>
</tbody>
</table>

BMI indicates body mass index; NA, not applicable. Direction of correlation: increased values of factors correlate with better postoperative function for all factors except dichotomous values (ie, male sex, Indian ethnicity, and walking aid use), for which presence of factor correlates with better postoperative function.

![Figure 4. Risk of Bias](image)

**Figure 4. Risk of Bias**
estimated to be –0.17; 95% CI, –0.28 to –0.06; \( P < .001 \) vs coefficient = –0.15; 95% CI, –0.24 to –0.05 when including all studies), while mean correlation with mental health was estimated to be 0.13 (95% CI, –0.04 to 0.29; \( P = .02 \) vs coefficient = 0.12; 95% CI, –0.01 to 0.25 when including all studies), and mean correlation with osteoarthritis severity was estimated to be 0.10 (95% CI, –0.01 to 0.20; \( P = .05 \) vs coefficient = 0.10; 95% CI, 0.01 to 0.19 when including all studies).

For the secondary outcome, physical function 6 months after TKA, 9 studies with 5743 participants reported estimates on 20 potential factors.\(^{29-33,35,36,40,41}\) Estimated correlation coefficients from multivariate meta-analysis are in Figure 3.\(^{29-33,35,36,41}\) Mean correlation with more catastrophizing was estimated to be –0.19 (95% CI, –0.35 to –0.01; \( P = .03 \); \( P \) score = 63.4%; 2 studies\(^{35,41}\); very low–certainty evidence and substantial heterogeneity between reported estimates of association \([I^2 = 85%]\)). While mean correlation with walking use was estimated to be –0.31 (95% CI, –0.45 to –0.17; \( P < .001 \); \( P \) score = 84.1%; 2 studies\(^{31,41}\); high–certainty evidence and substantial heterogeneity between reported estimates of association \([I^2 = 63%]\)). Mean correlation with better physical function was estimated to be 0.37 (95% CI, 0.27 to 0.46; \( P < .001 \); \( P \) score = 90.4; 3 studies\(^{30-32}\); moderate–certainty evidence and substantial heterogeneity among reported estimates of association \([I^2 = 93%]\)), while mean correlation with better mental health was estimated to be 0.15 (95% CI, –0.08 to 0.36; \( P = .28 \); \( P \) score = 53.5; 3 studies\(^{30,31,41}\); high–certainty evidence and substantial heterogeneity among reported estimates of association \([I^2 = 81%]\)). We were unable to conclude that clinically meaningful correlations did not exist for the other 15 factors owing to limited evidence (ie, wide CIs). For the 3-month outcome, we were unable to perform multivariate meta-analysis, as shown in eTable 2 in the Supplement.

QUIPS domains most frequently assessed as at low risk of bias were prognostic factor measurement (16 studies\(^{23-28,30-33,37,41}\)) and outcome measurement (17 studies\(^{22-28,30-38,40}\)). For high risk of bias, QUIPS domains most often assessed were attrition (7 studies\(^{23-26,30,35,37}\)) and statistical analysis (7 studies\(^{24,25,29,32,34,36,39}\)), as shown in Figure 4.

Our GRADE certainty of evidence judgements are included in previously listed data and in eTable 4 in the Supplement. The most common reasons for downgrading certainty of evidence were risk of bias and imprecision.

**Discussion**

To our knowledge, this study is the first prespecified systematic review and meta-analysis using wide eligibility criteria and evaluating certainty of evidence to identify preoperative and intraoperative factors correlated with physical function at 12 months after TKA. Evidence from 7 observational studies\(^{22,24-28,37}\) suggested that higher BMI was correlated with poorer physical function 12 months after TKA and that better preoperative physical function and more severe osteoarthritis were correlated with better physical function 12 months after TKA. Importantly, our findings did not suggest that individual patients with a poor risk factor profile will not experience functional improvement if they undergo TKA. Our findings merely suggest that identified factors were correlated with poorer or better physical function in an absolute sense and may therefore be useful for guiding expectations about TKA outcomes.

We found moderate–certainty evidence for a correlation between higher preoperative BMI and worse function at 12 months, with equal correlation in the sensitivity analysis, in which studies judged to be at high risk of bias were removed. This finding is similar to that of another meta-analysis,\(^{13}\) in which participants without obesity reported lower rates of disability than participants with obesity. The evidence was not graded, however, and the study included retrospective studies with follow-up at 6 months to 10 years. Although we found a correlation between obesity and poorer physical function after TKA, patients with obesity still experience improved function from baseline\(^{48}\) and should thus be considered for surgery. However, the surgeon needs to consider the functional benefit against the risk for complications (eg, septic revisions are more prevalent in patients with...
severe obesity and super obesity\textsuperscript{49) for each patient and discuss these issues with the patient to encourage realistic expectations before proceeding with TKA.\textsuperscript{49}

We found a correlation between better preoperative and better postoperative function at 12 months (low-certainty evidence) and 6 months (moderate-certainty evidence). The correlation remained, with increased coefficients, in the sensitivity analysis. It is not surprising that patients who were healthier before surgery may also have been healthier after surgery. However, our results conflict with those of a systematic review\textsuperscript{8} concluding that lower preoperative function was associated with better function 12 months after TKA. To resolve these conflicting findings, evidence is needed from well-conducted studies using standardized methods to measure factors and outcomes. We also estimated a correlation between more severe osteoarthritis (Kellgren-Lawrence grade) and better physical function at 12 months (high-certainty evidence) in multivariate meta-analysis and sensitivity analysis. These findings are consistent with those of a systematic review\textsuperscript{8} that included retrospective studies with follow-up extending to 1 year. Uncertainty remains regarding evidence for osteoarthritis severity as a factor associated with the outcome at 12 months.\textsuperscript{50,51}

Major strengths of our study include following the recently revised Cochrane Handbook\textsuperscript{16} and guidelines for peer-reviewed protocols,\textsuperscript{14} including longitudinal prospective studies reporting associations at predefined times after TKA, and using multivariate meta-analysis when the number of factors was large compared with the number of studies.\textsuperscript{15} Several previous systematic reviews were unable to perform meta-analysis owing to heterogeneity associated with measurement issues, and others used vote counting, a method discouraged in current guidelines.\textsuperscript{16} We used recommended tools to assess risk of bias (QUIPS) and certainty of evidence (GRADE). Additionally, we prioritized transparency with the systematic use of prespecified methods documented in the protocol,\textsuperscript{14} preprint,\textsuperscript{15} and this article’s supplemental materials.

Limitations

This study has several limitations. To obtain trustworthy estimates without prejudging which factors may have been associated with the outcome, we included a wide range of factors but only from prospective studies reporting associations at specific postoperative times. This necessarily included estimates from studies measuring factors using a range of methods, and so we accounted for heterogeneity in our random-effects meta-analyses. Less heterogeneity was observed across studies using a common measure, particularly 9 studies that used WOMAC to measure physical function. Narrower inclusion criteria could increase the potential for excluding important evidence.\textsuperscript{16} Some studies had large sample sizes and therefore provided precise estimates (ie, narrow CIs). \(I^2\) may be misleading when study estimates are very precise because it is statistically easier to distinguish (ie, detect heterogeneity) between study estimates. In this situation, it is important to consider the degree to which study estimates vary from one another and whether this is clinically important, rather than relying solely on \(I^2\). In particular, \(I^2\) from prognostic studies may be misleading so \(I^2\) statistics should be interpreted cautiously.\textsuperscript{18} Because studies with high risk of bias can lead to biased main results and heterogeneity, we performed prespecified sensitivity analyses and excluded studies assessed as high risk for each QUIPS domain.\textsuperscript{14} We planned to perform analyses of nonreporting bias, small study effects, and subgroup analyses,\textsuperscript{14} but the number of included studies did not meet our prespecified threshold.

We also downgraded certainty of evidence if we judged studies to be at risk of bias. Several studies\textsuperscript{11,52-54} had insufficient reporting of important QUIPS domains (such as attrition and statistical analysis), thus lowering the certainty that study estimates were unbiased. We suggest that researchers use tools like QUIPS at the study design stage to encourage low risk of bias in their findings regarding prognostic factors. This review identified some key areas for future research. Uncertainty remains regarding which patients may benefit most from TKA. Because patient preoperative status (ie, BMI, physical function, and osteoarthritis severity) may be correlated with overall outcomes, evidence from high-quality studies is fundamental for developing a prediction model to better identify patients at increased risk of poor outcomes after TKA. Prehabilitation

\textsuperscript{49)}

\textsuperscript{50,51)}

\textsuperscript{52-54)}

\textsuperscript{14,15)}
interventions to improve modifiable factors (eg, better mental health) are not well-established.\textsuperscript{55,56} We could not synthesize data for a number of factors given that they were studied only once. For these and other factors and outcomes, such as associations between physical function during the first year after TKA and biomechanical aspects of surgery (eg, implant) or pain management, evidence is lacking, highlighting the need for research from these perspectives with appropriate design and power. Additionally, our study provided evidence at the population level not at the level of individual patients. Our results are important for investigating factors to include in predictive models but should be used with caution at the individual level.

**Conclusions**

This study found that there is evidence (with moderate certainty) that higher BMI was correlated with worse physical function and that better physical function (low-certainty evidence) and more severe osteoarthritis (high-certainty evidence) were correlated with better physical function 12 months after TKA. Our findings suggest that these factors should be included in development of predictive models aimed at identifying patients at increased risk of poor function after TKA.
Role of the Funder/Sponsor: The funders had no role in the design, conduct of the study, collection, management, analysis, and interpretation of the data, preparation, review, or approval of the manuscript nor the decision to submit the manuscript for publication.

Additional Contributions: The authors thank the members from the user board, Richard Madsen, MA, and Jan Otto Veiseth, BS (Lovisenberg Diaconal Hospital), for their contributions. They were compensated for the time used for meetings. We also would like to thank the medical librarians, Gunn Kleven, BS, and Hilde Flaaten, BS (University of Oslo), for their work in the development of the search strategy and the systematic search for articles. The medical librarians were not compensated for their work.

REFERENCES


**SUPPLEMENT.**

*Methods.* Multivariate Meta-analysis

*eFigure 1.* Sensitivity Analysis

*eFigure 2.* Exploring Potential Inconsistency at 6 and 12 mo

*eFigure 3.* Univariate Meta-analysis

*eTable 1.* Sensitivity Analysis

*eTable 2.* Reported Associations at 3 mo After TKA

*eTable 3.* Definition and Labels of Factors
eTable 4. Grading of Recommendation Assessment, Development and Evaluation

eTable 5. Search Strategy

eTable 6. Reason for Exclusion of Individual Studies