Psychometric Assessment of the Patient’s Knee Implant Performance (PKIP) Questionnaire for the Assessment of Primary Total Knee Arthroplasty

Theresa Coles,1 Kim Dwyer,2 Margaret Mordini,3 Valerie Williams,1 Mark Clatworthy,4 Piers Yates,5 William Hamilton6
1 RTI Health Solutions, Research Triangle Park, NC, United States; 2 DePuy Synthes Joint Reconstruction, Warsaw, IN, United States; 3 Ascot Hospital, Auckland, New Zealand; 4 Fremeantle Hospital, Fremeantle, Australia; 5 Anderson Orthopedic Research Institute, Alexandria, VA, United States

BACKGROUND

PKIP Description
- Newly developed patient-reported outcome (PRO) measure
- Assess knee implant functional performance
- For use with patients before and after primary total knee arthroplasty (TKA)

Factors assessed that lead to patient dissatisfaction and describes current needs in knee functional performance
- Adequate stability, mobility, satisfaction, and confidence with knee implant
- 34 items, each answering a 5- to 14-point critical response scale, with higher values indicating better knee stability (except for items that are reverse-scored)

PKIP Development
- The PKIP has undergone two phases of development:
  - Phase 1: Conceptualization and item generation, focus groups, and in-depth interviews described in Coles et al. (2018)
  - Phase 2: Psychometric evaluation

OBJECTIVE
- To evaluate the psychometric properties of the PKIP (phase 2 of the PKIP development)

METHODS

Study Background
- Prospective, multicenter, nonrandomized, noncomparative, longitudinal study to gather clinical and PRO data regarding knee prostheses currently on the market

Participants
- Adults (aged 22 to 80) with noninflammatory degenerative arthroplasty (TKA)
- A total of 322 patients were followed at five study sites: Ascot Hospital, Auckland, New Zealand; Fremantle Hospital, Fremantle, Australia; Anderson Orthopedic Research Institute, Alexandria, Virginia, United States; Ascot Hospital, Auckland, New Zealand; Anderson Orthopedic Research Institute, Alexandria, VA

RESULTS

Participant Characteristics
- A total of 316 patients completed the presurgery visit, 299 (95.1%) the 1-year visit, 281 (92.6%) the 2-year visit, and 241 (76.4%) the 3-year visit
- Mean age (SD) was 67.5 years (13.0); range 21-92 years
- Female participants outnumbered males (68% vs. 32%)
- The majority of participants were white (71%)
- The average body mass index was 31.2 (SD = 4.6)
- Sample characteristics were similar across knee implant configurations

Response Distributions
- Descriptive statistics did not reflect flavour or ordering effect on responder response. Results were similar across knee configurations

Measurement Reliability
- Three main concepts emerged from the inter-item correlations and exploratory factor analysis based on median to strong inter-item correlations and factor loadings ≥ 0.37: confidence/stability, mostly subjective, and satisfaction
- Although the construct analyses and factor analyses suggested that the confidence and stability items were highly related, qualitative interviews revealed that patients did not always interpret these concepts. Therefore, the confidence and stability items were separated into two subscales
- Confirmatory factor analysis supported the structure depicted in Figure 2 (square root ≥ 0.50; CF: P < 0.001; comparative fit index ≥ 0.80; Tucker-Lewis Index ≥ 0.80; root mean square error of approximation ≤ 0.04)

Construct Validity
- Cronbach’s alphas were generally satisfactory across PKIP subscale and knee configuration, almost attaining the minimum suggested threshold of 0.70
- Although this study was not designed to assess test-retest reliability, an acceptable level of correlation was based on a small subgroup of patients (n = 40) who had the same PKIP Overall score at the minimum 1-year visit and minimum 1-year (postsurgery to 10 months) and minimum 1-year (postsurgery to 10 months) 2 years apart. It was not as easy to be stable.
- The test-retest intraclass correlation coefficient of the PKIP Overall score was 0.77 (95% CI 0.67 to 0.84)

Scoring
- PKIP subscale scoring algorithm was developed that involved reverse scoring two items, simple unweighted total scores
- PKIP subscale scores (Confidence, Stability, Modify Activities, Satisfaction) range from 0 to 10, and have the Overall score ranges from 0 to 100. Higher scores indicate better knee functioning

Table 1: Study Strata and Expected Sample Sizes

<table>
<thead>
<tr>
<th>Study Strata</th>
<th>Total Expected Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presurgery</td>
<td>300</td>
</tr>
<tr>
<td>Visit 1</td>
<td>261</td>
</tr>
<tr>
<td>Visit 2</td>
<td>216</td>
</tr>
<tr>
<td>Visit 3</td>
<td>193</td>
</tr>
<tr>
<td>Visit 4</td>
<td>165</td>
</tr>
<tr>
<td>Visit 5</td>
<td>149</td>
</tr>
</tbody>
</table>

Table 2: Summary of PKIP Validity-Correlation Hypotheses and Results for the PKIP Overall Score

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Presurgery to Postsurgery (10-22 Months)</th>
<th>Postsurgery to 1 Year</th>
<th>Minimum 1 Year</th>
<th>10-22 Months- Minimum 1 Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKS</td>
<td>0.83 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
</tr>
<tr>
<td>OKS</td>
<td>0.88 (p &lt; 0.001)</td>
<td>0.69 (p &lt; 0.001)</td>
<td>0.69 (p &lt; 0.001)</td>
<td>0.69 (p &lt; 0.001)</td>
</tr>
<tr>
<td>EQ-5D-3L Index</td>
<td>0.79 (p = 0.001)</td>
<td>0.65 (p = 0.001)</td>
<td>0.65 (p = 0.001)</td>
<td>0.65 (p = 0.001)</td>
</tr>
<tr>
<td>VAS</td>
<td>0.68 (p &lt; 0.001)</td>
<td>0.63 (p &lt; 0.001)</td>
<td>0.63 (p &lt; 0.001)</td>
<td>0.63 (p &lt; 0.001)</td>
</tr>
</tbody>
</table>

Table 3: Biomechanical PKIP-AES Scores at Minimum 1 Year Post-Surgery

<table>
<thead>
<tr>
<th>PKIP Subscale</th>
<th>Minimum 1 Year</th>
<th>10-22 Months</th>
<th>Minimum 1 Year to 10-22 Months</th>
<th>PKIP Overall</th>
<th>Minimum 1 Year to 10-22 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
</tr>
<tr>
<td>Stability</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
</tr>
<tr>
<td>Modify Activities</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
<td>0.70 (p &lt; 0.001)</td>
</tr>
</tbody>
</table>

Table 4: PKIP Effect Size Estimates

<table>
<thead>
<tr>
<th>PKIP Subscale</th>
<th>Minimum 1 Year</th>
<th>10-22 Months</th>
<th>Minimum 1 Year to 10-22 Months</th>
<th>PKIP Overall</th>
<th>Minimum 1 Year to 10-22 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
</tr>
<tr>
<td>Stability</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
</tr>
<tr>
<td>Modify Activities</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
<td>1.00 (p &lt; 0.001)</td>
</tr>
</tbody>
</table>

DISCRIMINATING ABILITY

- Analyses of variance (ANOVA) were used to explore the discriminability of scores by examining mean differences in scores across known subgroups by comparing the following: Patients who were classified as having the best 25% versus the bottom 25% on the OZI, and the top 50% and bottom 50%
- Patients with AKS total scores ≥ 80 (better knee functioning) versus patients with AKS scores < 80 (worse knee functioning)
- It was hypothesized that patients who were rated in the top 25% or the bottom 25% of the OZI would have statistically better functioning than those who fell between those groups

Ability to Detect Change
- PKIP scores after surgery from presurgery to postsurgery to 10-22 months

Future studies should monitor the performance of the PKIP in new populations, in other languages, and in treatment comparisons.

REFERENCES

Please see handout for complete reference list.

CONTACT INFORMATION

Theresa Coles, PhD
Senior Health Outcomes Analyst
RTI Health Solutions
801 Heritage Corporate Center
Research Triangle Park, NC 27709
Phone: +1.919.316.3843; E-mail: tcoles@rti.org

Presentation submitted for Health Outcomes Congress November 13-15, 2018
Amsterdam, The Netherlands