Total knee arthroplasty (TKA) is performed at an increasing rate in the United States, with projections to reach anywhere from 1.26 [1] to 3.48 [2] million annual occurrences by the year 2030. With advancements in presurgical risk stratification, patients with a wide spectrum of medical comorbidities may be suitable candidates for knee arthroplasty after appropriate optimization [3].

Presurgical optimization enhances patient care while reducing readmission rates and costs associated with total joint arthroplasty, demonstrating the multidisciplinary approach to arthroplasty surgery in the context of general health and well-being [4]. Despite the multidisciplinary approach to preoperative optimization, disease-specific and joint-centric assessment tools predominate patient-reported outcome measures used in the context of knee arthroplasty [5]. The Knee Injury and Osteoarthritis Outcome Score (KOOS) assesses patients’ perceptions of their knee pain and function and is a valid and responsive instrument [6]. The KOOS for Joint Replacement (KOOS-JR) is a specific short-form designed to address the most relevant issues in patients with end-stage knee osteoarthritis undergoing TKA and is a validated and efficient “knee health” instrument [7]. However, when taking into account the

No author associated with this paper has disclosed any potential or pertinent conflicts of interest, including any possible commercial interests, and sources of funding that may have influence on the submitted work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2020.01.032.

* Reprint requests: Lafi S. Khalil, MD, Department of Orthopedic Surgery, Henry Ford Health System, 2799 W. Grand Blvd, Detroit, MI 48202.

https://doi.org/10.1016/j.arth.2020.01.032
0883-5403/© 2020 Elsevier Inc. All rights reserved.
medical portfolio of many arthroplasty patients, there is a need for an efficient short-form assessment tool that can capture patient outcomes related to the patients’ general well-being in addition to their “knee health.”

Patient-Reported Outcomes Measurement Information System (PROMIS) is a global health (GH) outcome metric that is not specific to a single disease, procedure, or anatomic location. PROMIS has been validated to capture patients' overall health, allowing comparisons across disease states [8]. The standardization of these scores enable clear interpretation of results between studies and across medical specialties [9]. The technical advantage of administering PROMIS is the use of computer adaptive testing algorithms, which enhance the patient experience by taking less time to complete while increasing precision [8]. A practical advantage is the standardized format normalized to the general population with a universal mean ± standard deviation of 50 ± 10, facilitating administration and interpretation by providers from various specialties.

The purpose of this study was therefore to determine whether PROMIS-GH scales were a valid and effective tool in detecting change in patient outcomes following TKA as compared with legacy scores. Specifically, the validity, responsiveness, and minimal clinically important differences (MCIDs) of PROMIS-GH forms were evaluated with respect to KOOS-JR scales from preoperative baseline through 12 months postoperative follow-up. Our primary outcome was whether PROMIS-GH scales are a valid metric that would demonstrate responsiveness in knee arthroplasty patients that correlated with KOOS-JR scores. Secondary outcomes included correlation of MCID between PROMIS and KOOS-JR scales.

Materials and Methods

Institutional review board approval was obtained before retrospective review of clinical data. A consecutive series of all patients undergoing primary unilateral TKA between December 2017 through April 2019 by 3 fellowship trained arthroplasty surgeons were included. Authors JD, TB, and RW were responsible for 29%, 43%, and 27% of the cases, respectively. KOOS-JR and PROMIS-GH short forms were collected in clinic preoperatively and postoperatively per the surgeons’ routine clinical practice. Only patients completing at least one preoperative and one postoperative set of surveys were included for this review. For the purposes of this retrospective chart review, patient surveys were categorized according to postoperative day completed, with 0–45 days postoperatively assigned as 1-month scores, days 46–135 assigned as 3-month scores, days 136–270 assigned as 6-months, and after 270 days assigned as 1-year postoperatively. The surveys were completed on a tablet computer (iPad tablet; Apple Inc, Cupertino, CA). Patients who underwent staged bilateral TKAs were excluded to avoid any confounding effects of the second surgery. Patients unable to communicate in English were not included as they do not routinely complete the survey forms necessary for the study. Patient demographics, such as weight, height, body mass index, and age, were obtained through chart review of electronic medical records.

Completion of the KOOS-JR and PROMIS surveys was done voluntarily as part of routine practice. All survey data were collected and managed through REDCap (Research Electronic Data Capture), a password-encrypted, web-based application approved for data capture and storage by our institutional review board. The PROMIS-GH survey was divided into the 2 subcomponent PROMIS-Physical Health (PH) and PROMIS-Mental Health (MH) scores. The PROMIS-GH surveys are comprised of 10 short questions, and the forms may be viewed at www.healthmeasures.nei-t/score-and-interpret/calculate-scores).

Statistical Analysis

All statistical analyses were carried out by a trained biostatistician, using R (R Foundation for Statistical Computing 2018, Vienna, Austria).

Fig. 1. Correlation between PROMIS Physical Health (PH) and KOOS-JR scores. The postoperative PROMIS-PH and KOOS-JR interval scores were significantly correlated with one another at all time points (P < .001). KOOS-JR, Knee Injury and Osteoarthritis Outcome Score for Joint Replacement; PROMIS, Patient-Reported Outcomes Measurement Information System.
Austria). Alpha was set at 0.05 for all analyses, with a beta of 0.20. Analysis of variance was conducted to identify significant changes in PROMIS-PH, PROMIS-MH, or KOOS-JR scores at different time points. Floor and ceiling effects were calculated as the proportion of patients who have the minimum or maximum possible scores, respectively. Floor and ceiling effects are indicators of a scale’s ability to differentiate among patients at each end of the scale, evaluating a questionnaire’s performance in its range, accuracy, and response bias [10]. External validity was represented by Pearson correlations between PROMIS and KOOS-JR scores, with greater strength of correlations nearing the maximum value of 1.0 for KOOS-JR interval scores and −1.0 for KOOS-JR raw scores. Effect size indices (ESIs) were generated to display the responsiveness of each score at each postoperative time point, with an ESI of <0.2 implying a low effect, between 0.2 and 0.8 implying a moderate effect, and >0.8 implying a large effect. ESI is routinely used to assess the responsiveness of patient-reported outcomes to change, with a higher ESI indicating a better ability to detect change [11,12]. Pearson correlations coefficients (r) were assessed for interdomain relationships between KOOS-JR and PROMIS and were interpreted as follows: 0.80 to 1.00, very strong; 0.60 to 0.79, strong; 0.40 to 0.59 moderate; 0.20 to 0.39, weak; and 0.00 to 0.19, very weak [13]. Significant correlations between KOOS-JR and PROMIS PH would outline the convergent validity between the two surveys.

MCIDs were calculated for KOOS-JR and PROMIS using the distribution-based method, which was found by dividing by half the standard deviation of the preoperative outcome score, as previously demonstrated in a variety of orthopedic procedures [11,14–18]. Furthermore, an anchor-based method for calculating MCID was used for comparison. KOOS-JR was used as our anchor to identify the cutoff value of change in PROMIS scores, at the time of longest follow-up, that maximized the sensitivity and specificity of achievement of the KOOS-JR MCID [11]. Receiver operating characteristic curve method was used to determine which preoperative PROMIS-PH and KOOS-JR scores predict achieving MCID with greatest sensitivity and specificity [14].

**Results**

There were 67 patients excluded who had staged bilateral TKAs, 2 patients who had a conversion of a unicompartmental knee arthroplasty to a TKA, and 2 patients who had revision TKA as the index procedure (with previous primary TKAs performed at an outside institution). After these exclusions, 875 patients met inclusion criteria. The mean ± standard deviation for age and body mass index were 67.5 ± 9.2 years and 32.7 ± 6.2 kg/m², respectively. Surgery was conducted on the patients’ right side in 51% of cases. The floor and ceiling effects of the PROMIS-PH scales were both 0%, as the maximum (57.7) and minimum (19.9) scores in our cohorts did not reach the highest (67.7) or lowest (16.2) possible scores in the questionnaire. Significant correlations between PROMIS-PH and KOOS-JR interval scores were found both preoperatively and also at all postoperative time points (Fig. 1, Table 1).

The postoperative PROMIS-PH and KOOS-JR interval scores were significantly increased from baseline at each postoperative time point ($P < .001$; Table 1 and Fig. 2). The PROMIS-MH score did not change between time points ($P > .05$) and therefore was not shown to be responsive by the ESI (Table 1). The responsiveness of the PROMIS-PH score was shown to be moderate at 1 and 3 months (ESI >0.2) and excellent at 6 and 12 months (ESI >0.8), while the KOOS-JR was shown to have excellent responsiveness at all time points (Table 1).

The MCIDs for PROMIS-PH were similar when calculated with the distribution and anchor-based methods (2.3 and 2.5, respectively), and the numbers of patients achieving MCID for PROMIS-PH and KOOS-JR were similar at each postoperative time point (Table 2). Using the receiver operating characteristic analysis, a preoperative PROMIS-PH score less than or equal to 38 predicted achieving MCID at some point within 1-year follow-up with 59% sensitivity and 70% specificity. A preoperative PROMIS-PH score of less than 32.5 resulted in 93% specificity, with 88 of the 86 patients achieving MCID. Patients who achieved MCID in PROMIS-PH scores at each time point showed excellent responsiveness and correlation with KOOS-JR scores (Fig. 3).

**Discussion**

Our study found that the PROMIS-GH form, particularly the PH form, is a valid and responsive tool in patients undergoing primary primary TKA for osteoarthritis. The low floor and ceiling effects found in our study (0% for both) indicate that PROMIS-PH has the ability to differentiate between patients along either end of the spectrum, improving measurement accuracy and limiting response bias [10]. Patients demonstrated significant improvement throughout the 12-month follow-up at all time points in our study, relative to preoperative evaluation, as demonstrated by both KOOS-JR and PROMIS-PH scores. The responsiveness of KOOS-JR was more robust by 3-months postoperatively, while both KOOS-JR and PROMIS-PH demonstrated excellent responsiveness by 6-months to 12-months postoperatively (ESI >0.8). PROMIS-PH correlated significantly with the KOOS-JR score at all time points. Furthermore, each outcome tool was able to detect similar number of patients achieving MCID at each follow-up period.

In a retrospective review of 2291 patients who underwent primary TKA, Lyman et al concluded that the KOOS-JR was a valid and responsive outcome tool with a favorable floor and ceiling effect profile and high internal and external consistency [7]. Likewise, we demonstrated favorable floor and ceiling effects as well as high external validity. The responsiveness of the PROMIS-PH scores in our study mirrored that of KOOS-JR scores, a tool shown to be valid and responsive in the literature [7]. Further analysis of the ESI of KOOS-JR and PROMIS-PH in our study, which allows comparison

### Table 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Baseline (n = 875)</th>
<th>1-Mo (n = 699)</th>
<th>3-Mo (n = 170)</th>
<th>6-Mo (n = 134)</th>
<th>1-Y (n = 57)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROMIS-PH</td>
<td>38.5 ± 4.5</td>
<td>40.9 ± 4.6*</td>
<td>42.2 ± 4.9*</td>
<td>42.9 ± 5.1*</td>
<td>43.2 ± 4.7*</td>
</tr>
<tr>
<td>ESI PROMIS-PH</td>
<td>0.45</td>
<td>0.68</td>
<td>1.00</td>
<td>0.90</td>
<td>0.83</td>
</tr>
<tr>
<td>PROMIS-MH</td>
<td>46.2 ± 5.2</td>
<td>46.2 ± 5.1</td>
<td>46.0 ± 4.9</td>
<td>46.4 ± 5.0</td>
<td>47.1 ± 4.2</td>
</tr>
<tr>
<td>ESI PROMIS-MH</td>
<td>0.04</td>
<td>0.10</td>
<td>0.08</td>
<td>0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>KOOS-JR</td>
<td>47.8 ± 13.6</td>
<td>60.7 ± 11.5*</td>
<td>64.8 ± 12.5*</td>
<td>69.7 ± 15.3*</td>
<td>66.3 ± 15.3*</td>
</tr>
<tr>
<td>ESI KOOS-JR</td>
<td>0.90</td>
<td>1.09</td>
<td>1.29</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>Correlation between PROMIS-PH &amp; KOOS-JR</td>
<td>0.51*</td>
<td>0.40*</td>
<td>0.38*</td>
<td>0.47*</td>
<td>0.54*</td>
</tr>
</tbody>
</table>

KOOS-JR, Knee Injury and Osteoarthritis Outcome Score for Joint Replacement; MH, Mental Health; PH, Physical Health; PROMIS, Patient-Reported Outcomes Measurement Information System; SD, standard deviation.

* Compared with baseline, $P < .001$. 

Downloaded for Anonymous User (n/a) at Henry Ford Hospital / Henry Ford Health System (CS North America) from ClinicalKey.com by Elsevier on August 30, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.
cross scales as each outcome scale is normalized \([11,12]\), demonstrated moderate to strong responsiveness in both scales. Similar results are reported in arthroscopy populations, with both KOOS and PROMIS scales demonstrating moderate to strong responsiveness as measured by ESI \([11]\). Our study instead compares the PROMIS-PH short form, a patients’ self-perceived assessment of their global physical health, to the KOOS-JR short form, demonstrating that each respective short form is valid and responsive.

Demonstrating the validity and responsiveness of these PROMIS metrics is particularly important given concerns that the items included on the PROMIS scales might not capture changes related to specific diseases or interventions, as the more general health items could underrepresent the effects of a specific intervention (eg, TKA) on a specific disease \([19]\). Conversely, the influence of health issues unrelated to the disease undergoing intervention may potentially mask the response to intervention in more disease-specific and less global tools, or the influence of the intervention on global health outcomes might not be captured by disease-specific interventions. The ability of an outcome tool to detect change throughout the course of intervention for a specific disease enhances the applicability to trend outcomes. Maintaining this element of responsiveness after TKA while simultaneously capturing an assessment of a patient’s overall health, rather than solely “knee health,” provides clear advantages in the patient’s medical care \([11]\). Furthermore, it allows nonorthopedic providers and healthcare administrators to understand the patient’s general health state before and after surgical intervention as PROMIS is reported in a universally standardized scale \([8,19]\). In a prior retrospective review of PROMs in 124 knee arthroplasty patients, PROMIS Physical Function, Pain Interference, and Pain Intensity scores demonstrated a modest correlation between with KOOS-JR \([19]\). These same measures have demonstrated significant correlation in knee arthroscopy patients \([11]\). The present study expands upon these prior investigations by demonstrating the correlation between KOOS-JR and PROMIS-PH short forms, which is constructed to capture patient perceptions of their overall health. Our results note the utility of these general health forms in capturing patients’ response to total knee arthroplasty up to a year after surgery.

Patients undergoing TKA achieved MCID in similar proportions and magnitudes at all follow-up time points when comparing KOOS-JR and PROMIS PH. The greatest percentage of patients reached MCID at 6-months follow-up in both KOOS-JR and PROMIS PH scales (Table 2). This finding suggests that PROMIS-PH scales, despite being a general health metric, are equally as efficacious as a

<table>
<thead>
<tr>
<th>Measure</th>
<th>MCID</th>
<th>1 Mo</th>
<th>3 Mo</th>
<th>6 Mo</th>
<th>12 Mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROMIS Physical Health(^a)</td>
<td>2.3</td>
<td>57%</td>
<td>70%</td>
<td>70%</td>
<td>76%</td>
</tr>
<tr>
<td>KOOS-JR(^b)</td>
<td>6.8</td>
<td>65%</td>
<td>76%</td>
<td>84%</td>
<td>76%</td>
</tr>
<tr>
<td>PROMIS Physical Health(^b)</td>
<td>2.5</td>
<td>43%</td>
<td>58%</td>
<td>62%</td>
<td>59%</td>
</tr>
</tbody>
</table>

\(^a\) Distribution method: one-half of the standard deviation of the preoperative score.

\(^b\) Anchor method: KOOS MCID (6.8) used to calculate PROMIS Physical Health MCID.

**Table 2** Minimal Clinically Important Difference (MCID) Values.

**Fig. 2.** Box plots depicting the median scores and interquartile ranges at each time point for PROMIS-PH and KOOS-JR. The postoperative PROMIS-PH and KOOS-JR interval scores were significantly increased from baseline at each postoperative time point \((P \leq .001)\). KOOS-JR, Knee Injury and Osteoarthritis Outcome Score for Joint Replacement; PROMIS-PH, Patient-Reported Outcomes Measurement Information System Physical Health.

**Fig. 3.** Responsiveness in KOOS-JR in patients meeting PROMIS-PH MCID. The change in KOOS-JR scores for patients who achieved MCID for PROMIS-PH plotted by postoperative time point. Patients who achieved PROMIS-PH MCID demonstrated similar responsiveness on the KOOS-JR scale. Distribution-based MCID was calculated by the following formula: \(0.5 \times SD \) (preoperative). Anchor-based MCID was calculated by using the KOOS-JR MCID as the anchor and identifying the cutoff value of change in PROMIS scores at the time of longest follow-up that maximized the sensitivity and specificity of predicting achievement of the KOOS-JR MCID. KOOS-JR, Knee Injury and Osteoarthritis Outcome Score for Joint Replacement; PROMIS-PH, Patient-Reported Outcome Measurement Information System—Physical Health; MCID, minimal clinically important difference; SD, standard deviation.
disease-specific scale in detecting clinically significant change related to an intervention for that disease. The findings are consistent with the literature in a wide range of pathologies in orthopedics, such as knee arthroscopy and rotator cuff repair [9,11,20]. This finding is particularly important as medical care reimbursement is transitioning to a value-based system where the effectiveness of an intervention may be detected by the perceived improvement of the patient [21]. The ability of an outcome scale to detect MCID allows a physician to quantify and document a patient's improvement, enhancing a value-based model of patient care. The effectiveness of PROMIS-PH as an outcome metric is complemented by its efficient administration and low burden [9].

Limitations

This study is not without limitations. As a retrospective review, this study is limited by the availability of data in the electronic medical record and in the REDCap system, but because this study seeks to report on the validity of the PROMIS metric in routine clinical practice, the inclusion of only the data which is obtained as part of routine patient care enhances the generalizability and reproducibility of these findings to other clinical settings. An important limitation of the current report is that the majority of patients included have not completed the PROMIS and KOOS-JR surveys at 6-months or 1-year postoperatively. The timing of patients' final postoperative outcome measures ranges from 1-month to 12-months postoperatively, with some patients having been lost to follow-up in the REDCap system after the index procedure (Table 1). However, all available data were included, and statistically and clinically significant results were found with the numbers available. Further study is warranted to assess the responsiveness of the PROMIS and KOOS-JR PROMs in the mid-term to long-term postoperative time period.

Conclusions

The PROMIS Physical Health scale demonstrates similar effectiveness as the KOOS-JR in detecting disease-specific change following primary TKA in patients with knee osteoarthritis. While KOOS-JR showed greater responsiveness by 3-months postoperatively, this difference was not apparent by 6-months and 12-months. Both scales demonstrated moderate to strong responsiveness, suggesting both are able to capture the improvement in patient outcomes throughout the early postoperative time period up to 1 year. Further study will determine the ability of PROMIS scales to demonstrate improvement in the intermediate to late postoperative time period. The ability of PROMIS-PH to detect MCID in a similar capacity to KOOS-JR allows this scale to be used by clinicians in an efficient, effective, and value-based care system, which may help guide decision-making in both clinical and administrative contexts.

Acknowledgments

Yusuf Qadeer, BS: Wayne State University—School of Medicine 540 E. Canfield Ave., Detroit, MI 48201, USA. Michael Bazydlo, MS, Henry Ford Hospital Department of Biostatistics 2799 W. Grand Blvd, Detroit, MI 48202, USA.

References