

# UCSF

## UC San Francisco Previously Published Works

### Title

Are Patient-Reported Outcome Measures for Anterior Cruciate Ligament Injuries Validated for Spanish Language and Culture?

### Permalink

<https://escholarship.org/uc/item/0rg8q1vc>

### Journal

Orthopaedic Journal of Sports Medicine, 12(8)

### ISSN

2325-9671

### Authors

Siu, Jeremy  
Garcia-Lopez, Edgar  
Pandya, Nirav  
[et al.](#)

### Publication Date

2024-08-01

### DOI

10.1177/23259671241256413

Peer reviewed

# Are Patient-Reported Outcome Measures for Anterior Cruciate Ligament Injuries Validated for Spanish Language and Culture?

Jeremy W. Siu,<sup>\*†</sup> BS, Edgar Garcia-Lopez,<sup>‡</sup> MD, MS, Nirav K. Pandya,<sup>‡</sup> MD, Brian Feeley,<sup>‡</sup> MD, and Lauren M. Shapiro,<sup>‡</sup> MD, MS

*Investigation performed at University of California, San Francisco, California, USA*

**Background:** Patient-reported outcome measures (PROMs) have been adopted as a way to measure patient self-rated physical function and health status for patients with anterior cruciate ligament (ACL) injuries. Although multiple PROMs exist and have been translated into various languages, the cross-cultural adaptation and validity of these PROMs for Spanish-speaking patients is unknown.

**Purpose:** To evaluate the adaptation quality and psychometric properties of Spanish-language adaptations of PROMs for patients with ACL injuries.

**Study Design:** Scoping review; Level of evidence, 3.

**Methods:** Under PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, we reviewed published studies related to adaptation quality and psychometric properties of Spanish PROMs in patients with ACL injuries. The methodological quality of the included studies was assessed using the Guidelines for the Process of Cross-Cultural Adaptation of Self-Reported Measures, the Quality Criteria for Psychometric Properties of Health Status Questionnaires, and the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. The level of evidence for each PROM was determined based on the number of studies, methodological quality, consistency of results, and sample size.

**Results:** The initial search strategy identified 5687 articles. After removal of duplicates, 1882 titles were screened, and 114 articles were assessed for eligibility. Six articles were selected for final review, comprising 4 PROMs: the Lysholm knee score, the Anterior Cruciate Ligament–Return to Sport After Injury (ACL-RSI), the Lower Extremity Functional Scale, and the Lower Limb Functional Index. Three studies followed all 6 processes for cross-cultural adaptation. None of the studies demonstrated all 14 domains required for cross-cultural validity (eg, description of translator expertise). The ACL-RSI achieved the highest level of evidence, with 3 of 9 domains demonstrating moderate evidence.

**Conclusion:** This review identified 4 instruments that have been translated for Spanish-speaking patients with ACL injuries, none of which demonstrated appropriate adaptation or robust psychometric properties. The study highlights the need for improvement in PROMs for Spanish-speaking patients and the potential for mismeasurement and inappropriate application of PROM results in patients with ACL injuries.

**Keywords:** ACL; culture; language; patient-reported outcome measures; Spanish; validity

Anterior cruciate ligament (ACL) injuries are among the most common orthopaedic injuries in the United States and around the world, with a global incidence estimated to be 29 to 39 per 100,000 persons.<sup>16-18,26,27,44</sup> Patient-

reported outcome measures (PROMs) are utilized alongside objective measures such as knee stability and range of motion to assess function, treatment outcomes, and quality of life after ACL injury. Previous studies have reported the Lysholm knee score and the International Documentation Committee (IKDC) Subjective Knee Form (SKF) as the most commonly used PROMs.<sup>29</sup>

The majority of PROMs have been developed in English-speaking populations, and validation studies often lack

The Orthopaedic Journal of Sports Medicine, 12(8), 23259671241256413  
DOI: 10.1177/23259671241256413  
© The Author(s) 2024

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

data related to the ethnic or cultural backgrounds of the study population.<sup>14,19,28,30,48,55</sup> As such, the questions and responses on these PROMs may not apply to, or be valid for, patients of different ethnic or cultural backgrounds.<sup>19,21</sup> For example, Paz et al<sup>38</sup> demonstrated that English- and Spanish-speaking patients responded to 44% of items differently on linguistically translated PROMs, despite having similar levels of underlying physical function. This is particularly important for Spanish-speaking patients. Spanish is the second most widely spoken language globally; more than 500 million people speak Spanish as their first or second language, and 54 million people speak Spanish in the United States alone.<sup>47</sup> Failing to utilize culturally appropriate PROMs can result in missing or inaccurate data due to misunderstood or culturally irrelevant items.<sup>56</sup> This is reflective of a larger trend in which diverse patient populations are not included in orthopaedic research.<sup>5</sup> Therefore, it has been recommended by governing bodies, including the US Food and Drug Administration (FDA), to use PROMs that have been appropriately translated, cross-culturally adapted, and validated.<sup>53</sup>

As PROM use for ACL injuries continues to grow, it is important to ensure these tools are utilized appropriately, particularly as outcomes after ACL surgery are more complex than graft retear rates. As such, the purpose of this study was to evaluate the adaptation quality of PROMs for ACL injuries in the Spanish language and to assess the psychometric properties of PROMs for ACL injuries in the Spanish language. We hypothesized that PROMs for ACL injuries in Spanish-speaking patients will exhibit inadequate adaptation quality and psychometric properties.

## METHODS

### Literature Search

This review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.<sup>36</sup> We searched the PubMed, Cinahl via EBSCO, Medline via Ovid, Embase, and Web of Science databases on March 3, 2023, with the objective of identifying all published studies on linguistic and cultural adaptations and validation of instruments pertaining to ACL injuries in Spanish. We utilized the names of all of the PROMs that are described for ACL injuries, as guided by previous work.<sup>31</sup> We considered studies that evaluated lower extremity conditions and other traumatic conditions that included patients with ACL injuries. The searches were run using common terms related to PROMs; Spanish language and Spanish-speaking countries; cross-cultural

adaptation or validity; and lower extremity, knee, or ACL. The full search strategy is shown in the Supplemental Material (available separately).

### Eligibility Criteria

There was no time restriction for the included studies. Our inclusion criteria were (1) studies related to linguistic or cultural adaptation of PROMs assessing ACL injuries in the Spanish language; (2) studies reporting the process of linguistic or cross-cultural adaptation to Spanish; (3) studies reporting testing of linguistic or cross-cultural adaptations to Spanish; (4) studies with a full-text original article; (5) studies published in peer-reviewed journals; (6) studies written in either English or Spanish; (7) studies that included adults aged  $\geq 18$  years who had sustained ACL injuries; and (8) studies with evidence levels of 1 to 4. Excluded were comments, letters, editorial guidelines, conference reports, and reviews.

The retrieved articles were uploaded to Rayyan, an open-source online platform used for systematic reviews, and duplicates were subsequently eliminated.<sup>35</sup> Afterward, 2 independent reviewers (J.S., E.G.L.) screened the titles and abstracts, and the full-text articles were further assessed based on predetermined inclusion and exclusion criteria. Additional relevant studies that were missed during the primary search were identified by examining the references. The reviewers convened to discuss the included articles, and any disagreements were resolved through consensus. If a consensus was unachievable, a third reviewer (L.M.S.) was available to facilitate an agreement.

The initial search yielded a total of 5687 studies. After removal of duplicates, 1882 studies remained for further analysis. Our team conducted a thorough review of the references, but no additional studies were identified. The literature search, screening, and review process is depicted in Figure 1.

### Data Extraction and Quality Assessment

The following information was extracted from each study: lead author, year of publication, inclusion criteria, sample size, age (mean and standard deviation), sex distribution, condition studied, country of study, and PROMs. The methodological quality of the included studies was evaluated by 2 reviewers (J.S., E.G.L.) using 3 checklists:

1. The Guidelines for the Process of Cross-Cultural Adaptation of Self-Reported Measures, which consists of 6

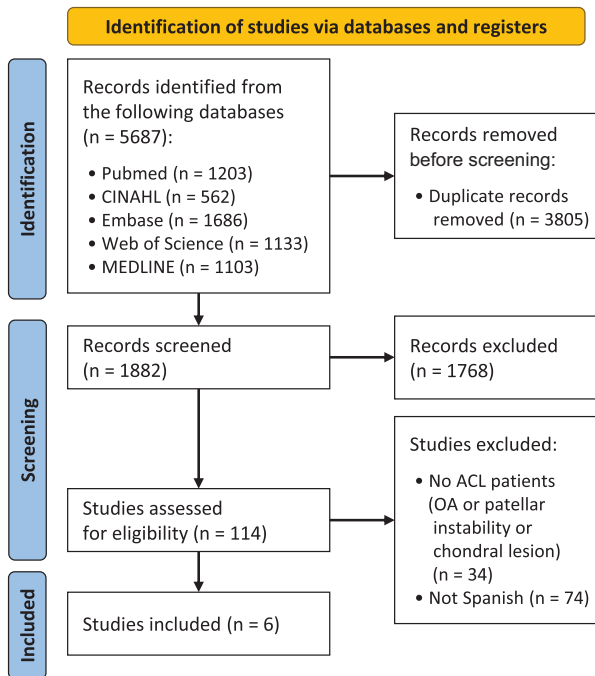
\*Address correspondence to Jeremy W. Siu, BS, School of Medicine, University of California, San Francisco, 500 Parnassus Avenue, MU 340, San Francisco, CA 94133, USA (email: Jeremy.siu@ucsf.edu).

†School of Medicine, University of California San Francisco, San Francisco, California, USA.

‡Department of Orthopedic Surgery, University of California San Francisco, San Francisco, California, USA.

Final revision submitted August 30, 2023; accepted November 16, 2023.

One or more of the authors has declared the following potential conflict of interest or source of funding: N.K.P. has received education payments from Evolution Surgical. B.F. has received education payments from Evolution Surgical. L.S. has received education payments from Southtech Orthopedics and Evolution Surgical and hospitality payments from Axogen and TriMed. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.



**Figure 1.** Flowchart summarizing the literature search, screening, and review. ACL, anterior cruciate ligament; OA, osteoarthritis.

items that outline the standardized adaptation process for translating and culturally adapting an instrument.<sup>4,15,54</sup>

2. The Quality Criteria for Psychometric Properties of Health Status Questionnaires, which evaluates the content validity, criterion validity, construct validity, agreement, reliability, responsiveness, floor or ceiling effects, and interpretability of instruments.<sup>49</sup>
3. The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) Checklist for Cross-Cultural Validity, which evaluates the methodological quality of studies and the cross-cultural validity of a PROM.<sup>33,34</sup> Cross-cultural validity measures the degree to which the performance of items on a translated or adapted instrument adequately reflects the performance of the original version.<sup>33,34</sup> We utilized the “worst score counts” principle to evaluate the COSMIN checklist (ie, the lowest rating of any standard in the box is recorded).<sup>24,46</sup>

Agreement between the 2 reviewers was evaluated using the quadratic weighted  $\kappa$  statistic ( $\kappa_w$ ). If  $\kappa_w$  was  $>0.75$ , the results were adopted. Disagreements were resolved through consensus discussion or with the involvement of a third reviewer (L.M.S.).

### Level of Evidence Assessment

We determined the level of evidence for each PROM by combining the results of the studies for each domain. In accordance with previous investigation, we considered the number of studies, the methodological quality of the

studies, and the consistency of the results.<sup>7,24,32,40,49-51</sup> Sample size was evaluated in accordance to previous investigations and as dictated by COSMIN, in which sample sizes of  $>100$  were considered excellent, 50 to 100 good, 30 to 50 fair, and  $<30$  poor.<sup>24,32,33,40,51</sup> We rated the measurement properties for each PROM as strong, moderate, limited, or conflicting evidence based on these factors.<sup>23,31,39,50-52</sup>

## RESULTS

Ultimately, 6 studies were selected for final review,<sup>2,8-10,39,43</sup> of which 4 were published in English<sup>2,8,9,43</sup> and 2 in Spanish.<sup>10,39</sup> The evaluated studies included 4 PROMs: the Lysholm knee score, the Anterior Cruciate Ligament–Return to Sport After Injury (ACL-RSI), the Lower Extremity Functional Scale (LEFS), and the Lower Limb Functional Index (LLFI). Three of the studies focused specifically on ACL injuries,<sup>2,39,43</sup> and the other 3 studies examined a range of conditions including ACL injuries.<sup>8-10</sup> The study populations comprised persons from 3 different countries; the key characteristics of the included studies are listed in Table 1.

### Quality Assessment of the Adaptation Process

The interrater reliability for the quality assessment of the adaptation process yielded a  $\kappa_w$  value of 0.94, indicating substantial agreement. As shown in Table 2, all 6 studies included the initial translation, synthesis of the translation, and back-translation steps.<sup>2,8-10,39,43</sup> Four studies incorporated an expert committee as well as testing of the prefinal version.<sup>8,10,39,43</sup> Three studies included an appraisal of the adaptation process<sup>10,39,43</sup>; notably, these same 3 studies completed all 6 steps of the adaptation process.

### Measurement Property Methodology

The  $\kappa_w$  value for interrater reliability regarding the assessment of measurement properties was 0.95, indicating substantial agreement (Table 2). All of the included studies were rated positively for internal consistency, assessed via Cronbach alpha, and for reliability, assessed via intra-class correlation coefficient (ICC). Two studies reported both construct and criterion validity,<sup>2,9</sup> whereas no other study had information available on criterion validity. Four studies received an indeterminate rating for content validity,<sup>8,10,39,43</sup> as they lacked a clear description of the measurement aim, target population, or concepts being measured. Meanwhile, 2 studies were rated negatively with regard to content validity,<sup>2,9</sup> as they failed to describe the target population. None of the studies evaluated all of the measurement properties.

### COSMIN Cross-Cultural Validity

The  $\kappa_w$  value for interrater reliability regarding the cross-cultural validity assessment was 0.89, indicating substantial agreement. None of the studies demonstrated all 14

TABLE 1  
Characteristics of Included Studies<sup>a</sup>

Characteristic	Arroyo Morales (2019) <sup>2</sup>	Peña (2021) <sup>39</sup>	Sala-Barat (2019) <sup>43</sup>	Cruz-Díaz (2006) <sup>8</sup>	Dell'Era (2016) <sup>10</sup>	Cuesta-Vargas (2014) <sup>9</sup>
Inclusion criteria	Adult patients with ACL injury	Adult patients with ACL injury	Football players undergoing primary ACLR with or without associated meniscal injury	Various acute/subacute/chronic MSK diagnoses (9% were lower extremity ligamentous injuries)	Various MSK injuries (22% were lower extremity ligamentous injuries)	Various lower limb injuries (24% were ligamentous injuries of the knee)
Sample size	95	93	114	132	127	136
Age, y (mean ± SD)	21.8 ± 5.4	M: 36; F: 45 <sup>b</sup>	21.8 ± 5.22	27.11 ± 6.22	43 ± 17.1	48 ± 19 <sup>c</sup>
Sex, n (%)	M: 64 (67); F: 31 (33%)	M: 40 (43); F: 53 (57%)	M: 87 (85); F: 17 (15%)	M: 73 (55); F: 59 (45)	M: 59 (47); F: 68 (54%)	M: 62 (46); F: 74 (54%)
Country	Spain	Colombia	Spain	Spain	Argentina	Spain
PROM evaluated	Lysholm	Lysholm	ACL-RSI	LEFS	LEFS	LLFI

<sup>a</sup>ACL, anterior cruciate ligament; ACLR, anterior cruciate ligament reconstruction; ACL-RSI, Anterior Cruciate Ligament–Return to Sport After Injury; F, female; LEFS, lower extremity functional scale; LLFI, lower limb functional index; M, male; MSK, musculoskeletal.

<sup>b</sup>No SD reported.

<sup>c</sup>Age 50 ± 19 years for patients with knee injuries.

TABLE 2  
Quality Assessment of the Adaptation Process and Quality Criteria for Measurement Properties<sup>a</sup>

Assessment	Arroyo Morales (2019) <sup>2</sup>	Peña (2021) <sup>39</sup>	Sala-Barat (2019) <sup>43</sup>	Cruz-Díaz (2006) <sup>8</sup>	Dell'Era (2016) <sup>10</sup>	Cuesta-Vargas (2014) <sup>9</sup>
Quality assessment of the adaptation process						
Initial translation	+	+	+	+	+	+
Synthesis of the translations	+	+	+	+	+	+
Back-translation	+	+	+	+	+	+
Expert committee	o	+	+	+	+	o
Testing the prefinal version	+	+	+	-	+	-
Appraisal of the adaptation process	o	+	+	-	+	-
Quality criteria for measurement properties						
Criterion validity	+	o	o	o	o	+
Construct validity	+	o	+	+	?	+
Content validity	-	?	?	?	?	-
Internal consistency	+	+	+	+	+	+
Agreement	o	+	o	+	o	+
Reliability	+	+	+	+	+	+
Floor/ceiling effect	+	o	+	+	o	o
Responsiveness	-	o	o	-	+	-
Interpretability	o	o	o	+	?	o

<sup>a</sup>+, performed; ?, indeterminate rating; -, not performed, o, unavailable or not clear.

domains required for cross-cultural validity (Table 3). In addition, only 2 studies reported the percentage of missing items (item 1).<sup>2,9</sup> Notably, all studies reported an adequate sample size (item 3), and in all studies, the items were translated forward and backward (item 7).

#### Level of Evidence Assessment

A  $\kappa_w$  value of 0.826 was achieved for interrater reliability regarding the level of evidence assessment, indicating substantial agreement between the reviewers. The ACL-RSI

achieved the highest level of evidence, with 3 of the 9 domains demonstrating moderate evidence (Table 4). No domain of any PROM demonstrated strong evidence. Internal consistency and reliability showed the most evidence across all PROMs, while interpretability was available for only 1 study.<sup>8</sup>

#### DISCUSSION

In this review, we identified 4 instruments, none of which demonstrated appropriate adaptation or robust

TABLE 3  
COSMIN Checklist for Cross Cultural Validity<sup>a</sup>

Checklist Item <sup>b</sup>	Arroyo Morales (2019) <sup>2</sup>	Peña (2021) <sup>39</sup>	Sala-Barat (2019) <sup>43</sup>	Cruz-Díaz (2006) <sup>8</sup>	Dell’Era (2016) <sup>10</sup>	Cuesta-Vargas (2014) <sup>9</sup>
1	+	-	-	-	-	+
2	+	-	-	-	-	+
3	+	+	+	+	+	+
4	+	+	+	+	+	+
5	-	+	+	+	0	+
6	0	+	+	+	+	+
7	+	+	+	+	+	+
8	0	+	+	+	+	0
9	0	+	+	+	+	0
10	+	+	+	-	+	-
11	+	-	+	-	+	-
12	+	0	+	+	+	0
13	+	-	-	-	-	-
14	+	-	+	+	-	+

<sup>a</sup>+, yes; -, no; 0, unavailable or not clear. COSMIN, Consensus-based Standards for the selection of health Measurement Instruments; HR-PRO, health-related patient-reported outcomes.

<sup>b</sup>Items: 1, Was the percentage of missing items given? 2, Was there a description of how missing items were handled? 3, Was the sample size included in the analysis adequate? 4, Were both the original language in which the HR-PRO instrument was developed, and the language in which the HR-PRO instrument was translated described? 5, Was the expertise of the people involved in the translation process adequately described? 6, Did the translators work independently from each other? 7, Were items translated forward and backward? 8, Was there an adequate description of how differences between the original and translated version were resolved? 9, Was the translation reviewed by a committee? 10, Was the HR-PRO instrument pretested to check interpretation, cultural relevance of the translation, and ease of comprehension? 11, Was the sample used in the pretest adequately described? 12, Were the samples similar for all characteristics except language and/or cultural background? 13, Were there any important flaws in the design or methods of the study? 14, For Classical Test Theory: was confirmatory factor analysis performed? For Item Response Theory: Was differential item function between language groups assessed?

TABLE 4  
Ratings of Measurement Properties for Level of Evidence<sup>ab</sup>

Domain	Lysholm	ACL-RSI	LEFS	LLFI
Criterion validity	+	NA	NA	+
Construct validity	+	+	+	+
Content validity	+	+	+	NA
Internal consistency	++	++	++	+
Agreement	+	NA	+	+
Reliability	++	++	++	+
Responsiveness	+	++	+	NA
Floor and ceiling	+	+	+	NA
Interpretability	NA	NA	+	NA

<sup>a</sup>ACL-RSI, Anterior Cruciate Ligament–Return to Sport After Injury; LEFS, lower extremity functional scale; LLFI, lower limb functional index; NA, not available (not performed or described).

<sup>b</sup>Grading: +++ or ---, multiple studies of good quality OR 1 study of excellent quality: strong evidence of positive/negative result; ++ or --, multiple studies of fair quality OR 1 study of good quality: moderate evidence of positive/negative result; + or -, 1 study of fair quality: limited evidence of positive/negative result; +/-, conflicting findings; ?, only studies of poor quality: unknown, due to poor methodological quality.

psychometric properties. Our findings highlight the need for improvement in outcome measurement for Spanish-speaking patients with ACL injuries and the potential for mismeasurement and inappropriate application of PROM results in this population. Given the growing use of

PROMs in guiding care and assessing treatment of patients with ACL injuries, appropriate utilization across diverse patient groups is crucial. Linguistic and cultural adaptation of PROMs promotes accurate measurement, applicability, and patient engagement. This also leads to

the inclusion of a more diverse patient population in orthopaedic research, particularly with the ability to capture language-based outcomes.

With increasing cross-national research, it is crucial to adapt PROMs to specific populations to ensure appropriate measurement. This is illustrated by the growing body of evidence demonstrating that a patient's language and culture can significantly influence their PROM scores.<sup>1,6,38,53</sup> Notably, 3 studies (of 3 PROMs) followed all 6 steps of the comprehensive adaptation process outlined by Beaton et al.<sup>3,10,39,43</sup> The 3 studies that did not include all steps of the adaptation process completed the initial steps but failed to test the prefinal version and appraise the adaptation process (Table 2).<sup>2,8,9</sup> These data demonstrate superior cultural adaptation as compared with other conditions in orthopaedics; for example, a similar evaluation of PROMs for patients with distal radius fractures demonstrated that no studies followed and reported all 6 steps of cross-cultural adaptation.<sup>25</sup> While the cross-cultural adaptation of PROMs evaluating ACL injuries in the Spanish-speaking population may be superior to that of other conditions, high-quality PROM adaptation can be enhanced by following standardized guidelines and documenting the process.<sup>3</sup>

Psychometric evaluation, typically conducted after the adaptation of a PROM, is important to understand and ensure the validity of a PROM in measuring the intended constructs. Construct validity, internal consistency, reliability, and floor/ceiling effect were reliably measured by all of the included studies.<sup>2,8-10,39,43</sup> Notably, interpretability was only reported in 1 study.<sup>8</sup> Given the variance of literacy levels of patients, it is important to have a PROM that is easily readable after being translated.<sup>50</sup> Finally, the content validity of every study was either negatively rated or indeterminant. Content validity, the extent to which a PROM measures the concept of interest (eg, physical function) in a target population (eg, Spanish-speaking patients with ACL injuries) is a critical step in the evaluation and validation of a PROM and is considered to be the most important measurement property of a PROM, such that its importance is stressed by the US FDA and the European Medicines Agency.<sup>7,12,51,52</sup> Notably, the lack of content validity negatively affects other measurement properties (eg, responsiveness, internal consistency), and may indicate a PROM is ineffective at measuring a specific concept in a specific population.<sup>7,12,51,52</sup> Ensuring that components of the Quality Criteria for Psychometric Properties of Health Status Questionnaires guidelines are addressed and documented appropriately helps ensure the proper use of PROM tools and understand opportunities for improvement.<sup>13,49</sup>

Finally, the COSMIN checklist guides the assessment of cross-cultural validity, gauging how well translated or culturally adapted instruments capture the original version's item performance.<sup>49</sup> Our research revealed that no PROMs specifically related to ACL injuries had undergone rigorous cross-cultural validation for use in Spanish-speaking populations. Sala-Barat et al<sup>43</sup> was the most rigorous, reporting 11 out of 14 items. Similarly, it was not obvious whether multiple studies had included data pertaining to some of

the items (Table 3). It is possible that the authors of these studies were not specifically following the COSMIN guidelines, which highlights the importance of utilizing a systematic approach to the cross-cultural adaptation of PROMs. Future research should detail how missing data were handled, address committee member disagreements, and describe the translation committee. These aspects are crucial for establishing PROM validity and reliability in diverse populations.

The lack of appropriately adapted and validated instruments has implications for both research and clinical practice. From a research perspective, multiple studies have demonstrated the lack of diversity of patients included in research studies, particularly randomized controlled trials. For example, a systematic review evaluating orthopaedic randomized controlled trials demonstrated that the reporting of race and ethnicity of study participants occurred in 7% and 3% of publications, respectively.<sup>37</sup> This phenomenon has been demonstrated in the spine and hand literature as well.<sup>11,22</sup> The lack of appropriately translated and adapted PROMs may contribute to the lack of diversity in trials and has been cited as a barrier to inclusion of diverse populations in research.<sup>20,45</sup> Because clinical practice guidelines (ie, tools based upon best available evidence to guide treatment) are informed by these research studies, it is critical these studies include diverse patient populations such that the clinical practice guidelines are broadly applicable. Similarly, there is growing evidence that demonstrates differences in outcomes based on race, ethnicity, and language.<sup>23,41,42</sup> As PROM use grows in guiding clinical care and evaluating outcomes, the appropriate adaptation and use of these tools is critical to prevent health care disparities.

## Limitations

This review has limitations. First, the scope of the review was confined solely to peer-reviewed articles, thus potentially introducing publication bias. This bias may have also affected the validity of our results, given that studies presenting positive outcomes are more likely to receive publication. Further, we recognize that many other PROMs exist that may have been translated and adapted for Spanish-speaking patients with ACL injuries. For example, many Patient Reported Outcomes Measurement Information System (PROMIS) tools have been translated and may be used in research and in clinical practice; however, our search strategy did not identify any relevant development, adaptation, or validation studies specific for this population. Second, the inclusion criteria only considered studies with complete original texts, which may have excluded relevant research solely available in abstract form or conference reports. Another limitation arose from the absence of a gold standard by which to compare with the newly created or adapted questionnaire (eg, in evaluating criterion validity). Identifying such a measure for ACL injuries would permit a more accurate evaluation of each questionnaire's reliability. Furthermore, we did not seek clarification or retrieval of unpublished data from the authors of the included papers.

## CONCLUSION

Our study demonstrated that, despite ACL injury prevalence and patient diversity, there was insufficient evidence supporting the adaptation quality and psychometric properties of PROMs for Spanish-speaking patients. We identified 4 instruments, none of which demonstrated appropriate adaptation or robust psychometric properties. The findings highlight the need for improvement in PROMs for Spanish-speaking patients and the potential for mismeasurement and inappropriate application of PROM results in this population. These improvements in the quality and psychometric properties may be possible with adherence to standard guidelines and reporting processes.

## REFERENCES

- Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of health-related quality of life. *Qual Life Res.* 1993;2(6):369-395.
- Arroyo-Morales M, Martin-Alguacil J, Lozano-Lozano M, et al. The Lysholm score: cross cultural validation and evaluation of psychometric properties of the Spanish version. *PLoS One.* 2019;14(8):e0221376.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000;25(24):3186-3191.
- Beaton DE, Schemitsch E. Measures of health-related quality of life and physical function. *Clin Orthop Relat Res.* 2003;413:90-105.
- Beck JJ, Pandya NK, Carter CW, Mulcahey MK. Current concept review: inclusion and analysis of diverse study populations in orthopaedic research. *J Am Acad Orthop Surg.* 2021;29(10):e479-e487.
- Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality. *J Clin Epidemiol.* 1998;51(11):913-923.
- Center for Drug Evaluation and Research. Patient-focused drug development: selecting, developing, or modifying fit-for-purpose clinical outcome assessments. US Food and Drug Administration. June 2022. Accessed April 22, 2023. <https://www.fda.gov/media/159500/download>
- Cruz-Díaz D, Lomas-Vega R, Osuna-Pérez MC, Hita-Contreras F, Fernández AD, Martínez-Amat A. The Spanish lower extremity functional scale: a reliable, valid and responsive questionnaire to assess musculoskeletal disorders in the lower extremity. *Disabil Rehabil.* 2014;36(23):2005-2011.
- Cuesta-Vargas AI, Gabel CP, Bennett P. Cross cultural adaptation and validation of a Spanish version of the lower limb functional index. *Health Qual Life Outcomes.* 2014;12:75.
- Dell'Era S, Dimaro M, Gamboa A, Spath MB, Salzberg S, Hernández D. Cross-cultural adaptation and Argentine validation of the Lower Extremity Functional Scale Questionnaire [Article in Spanish]. *Medicina (B Aires).* 2016;76(5):279-285.
- Delma S, Langford K, Baylor JL, et al. Race and ethnicity reporting in randomized controlled trials published in upper-extremity journals. *J Hand Surg.* 2023;48(4):340-347.
- European Medicines Agency. Regulatory guidance for the use of health-related quality of life (HRQL) measures in evaluation medicinal products - scientific guideline. Accessed April 22, 2023. <https://www.ema.europa.eu/en/regulatory-guidance-use-health-related-quality-life-hrql-measures-evaluation-medicinal-products-scientific-guideline>
- Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health.* 2007;10(Suppl\_2):S94-S105.
- Gabel CP, Melloh M, Burkett B, Michener LA. Lower limb functional index: development and clinimetric properties. *Phys Ther.* 2012;92(1):98-110.
- Gandek B, Ware JE Jr. Methods for validating and norming translations of health status questionnaires: the IQOLA Project approach. International quality of life assessment. *J Clin Epidemiol.* 1998;51(11):953-959.
- Gianotti SM, Marshall SW, Hume PA, Bunt L. Incidence of anterior cruciate ligament injury and other knee ligament injuries: a national population-based study. *J Sci Med Sport.* 2009;12(6):622-627.
- Granan LP, Bahr R, Steindal K, Furnes O, Engebretsen L. Development of a national cruciate ligament surgery registry: the Norwegian National Knee Ligament Registry. *Am J Sports Med.* 2008;36(2):308-315.
- Granan LP, Forssblad M, Lind M, Engebretsen L. The Scandinavian ACL registries 2004-2007: baseline epidemiology. *Acta Orthop.* 2009;80(5):563-567.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46(12):1417-1432.
- Hussain-Gambles M, Atkin K, Leese B. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. *Health Soc Care Community.* 2004;12(5):382-388.
- Hyland CJ, Guo R, Dhawan R, et al. Implementing patient-reported outcomes in routine clinical care for diverse and underrepresented patients in the United States. *J Patient-Rep Outcomes.* 2022;6(1):20.
- Issa TZ, Lambrechts MJ, Canseco JA, et al. Reporting demographics in randomized control trials in spine surgery - we must do better. *Spine J.* 2023;23(5):642-650.
- Kerman H, Smith SR, Smith KC, et al. Disparities in total knee replacement: population losses in quality-adjusted life years due to differential offer, acceptance, and complication rates in African Americans. *Arthritis Care Res (Hoboken).* 2018;70(9):1326-1334.
- Kleinlugtenbelt YV, Nienhuis RW, Bhandari M, Goslings JC, Poolman RW, Scholtes VAB. Are validated outcome measures used in distal radial fractures truly valid?: A critical assessment using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. *Bone Joint Res.* 2016;5(4):153-161.
- Lemos J, Xiao M, Castro Appiani LM, Katz P, Kamal RN, Shapiro LM. Are patient-reported outcome measures for distal radius fractures validated for Spanish and culture? A systematic review. *J Hand Surg.* 2023;48(7):673-682.
- Lind M, Menhert F, Pedersen AB. The first results from the Danish ACL reconstruction registry: epidemiologic and 2 year follow-up results from 5,818 knee ligament reconstructions. *Knee Surg Sports Traumatol Arthrosc.* 2009;17(2):117-124.
- Lyman S, Koulouvaris P, Sherman S, Do H, Mandl LA, Marx RG. Epidemiology of anterior cruciate ligament reconstruction: trends, readmissions, and subsequent knee surgery. *J Bone Joint Surg Am.* 2009;91(10):2321-2328.
- Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med.* 1982;10(3):150-154.
- Makhni EC, Padaki AS, Petridis PD, et al. High variability in outcome reporting patterns in high-impact ACL literature. *J Bone Joint Surg Am.* 2015;97(18):1529-1542.
- Marmura H, Tremblay PF, Getgood AMJ, Bryant DM. Development and preliminary validation of the KOOS-ACL: a short form version of the KOOS for young active patients with ACL tears. *Am J Sports Med.* 2023;51(6):1447-1456.
- Meta F, Lizzio VA, Jildeh TR, Makhni EC. Which patient reported outcomes to collect after anterior cruciate ligament reconstruction. *Ann Joint.* 2017;2:21.
- Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1171-1179.
- Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement



- properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 19(4):539-549.
34. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 63(7):737-745.
  35. Uuzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev.* 2016;5(1):210.
  36. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.
  37. Paul RW, Lee D, Brutico J, Tjoumakaris FP, Ciccotti MG, Freedman KB. Reporting and analyzing race and ethnicity in orthopaedic clinical trials: a systematic review. *J Am Acad Orthop Surg Glob Res Rev.* 2021;5(5):e21.00027.
  38. Paz SH, Spritzer KL, Morales LS, Hays RD. Evaluation of the patient-reported outcomes information system (PROMIS®) Spanish-language physical functioning items. *Qual Life Res.* 2013;22(7):1819-1830.
  39. Peña ORD, Gómez Gelvez A, Torres LP, García LFG. Cross-cultural adaptation to Spanish and validation of Lysholm knee score [article in Spanish]. *Rev Colomb Ortop Traumatol.* 2021;35(3):223-228.
  40. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147-1157.
  41. Reyes AM, Katz JN. Racial/ethnic and socioeconomic disparities in osteoarthritis management. *Rheum Dis Clin North Am.* 2021;47(1):21-40.
  42. Rosenberg SI, Chu Y, Ouweleen AJ, Hall TB, Patel NM. Is preferred language other than English associated with delayed surgery after ACL injury in children and adolescents? *Clin Orthop Relat Res.* 2023;481(2):292-298.
  43. Sala-Barat E, Álvarez-Díaz P, Alentorn-Geli E, Webster KE, Cugat R, Tomás-Sabado J. Translation, cross-cultural adaptation, validation, and measurement properties of the Spanish version of the anterior cruciate ligament-return to sport after injury (ACL-RSI-Sp) scale. *Knee Surg Sports Traumatol Arthrosc.* 2020;28(3):833-839.
  44. Sanders TL, Maradit Kremers H, Bryan AJ, et al. Incidence of anterior cruciate ligament tears and reconstruction: a 21-year population-based study. *Am J Sports Med.* 2016;44(6):1502-1507.
  45. Sheikh A. Why are ethnic minorities under-represented in US research studies? *PLoS Med.* 2005;3(2):e49.
  46. Sigirtmac IC, Oksuz C. Systematic review of the quality of the cross-cultural adaptations of disabilities of the arm, shoulder and hand (DASH): systematic review of the cross-cultural adaptations of DASH. *Med Lav.* 2021;112(4):279-291.
  47. Spanish in the World. Ministerio de Asuntos Exteriores, Unión Europea y Cooperación. Accessed April 22, 2023. <https://www.exteriores.gob.es/en/PoliticaExterior/Paginas/EIspanolEnElMundo.aspx>
  48. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res.* 1985;198:43-49.
  49. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
  50. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCV. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21(4):651-657.
  51. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res.* 2018;27(5):1159-1170.
  52. US Department of Health and Human Services FDA Center for Drug Evaluation and Research, US Department of Health and Human Services FDA Center for Biologics Evaluation and Research, US Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes.* 2006;4(1):79.
  53. Wagner AK, Gandek B, Aaronson NK, et al. Cross-cultural comparisons of the content of SF-36 translations across 10 countries. *J Clin Epidemiol.* 1998;51(11):925-932.
  54. Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability. *J Clin Epidemiol.* 1998;51(11):945-952.
  55. Webster KE, Feller JA. Evaluation of the responsiveness of the anterior cruciate ligament return to sport after injury (ACL-RSI) scale. *Orthop J Sports Med.* 2021;9(8):23259671211031240.
  56. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health.* 2005;8(2):94-104.