Original articles

Development and preliminary validation of the measurement properties of the S-Series Functional Scales for the shoulder and knee

Ton That Minh Dat[1*], Nguyen Thi Hue[2], Truong Thi Thuy Giang[2], Vo Chi Dung[2]

[1]Department of Rehabilitation, Hue University of Medicine and Pharmacy, Hue, Vietnam

[2]Post Graduate student, Department of Rehabilitation, Hue University of Medicine and Pharmacy, Hue, Vietnam

***Corresponding author:** Ton That Minh Dat; Email: ttmdat@huemed-univ.edu.vn

Received: 30/01/2026; **Accepted:** 19/05/2026; **Published:** 25/08/2026

DOI: 10.34071/jmp.2026.4.940

Abstract

Background: Shoulder and knee musculoskeletal disorders are major contributors to functional limitation. Although established patient-reported outcome measures (PROMs) such as SPADI and KOOS are widely used, their length and lack of structural standardization across joints limit feasibility for routine monitoring and digital health applications.

Objective: To describe the development of the S-Series Functional Scales and to evaluate the preliminary measurement properties of the Shoulder Function Scale (SFS) and Knee Function Scale (KFS).

Methods: The S-Series was developed according to a 4S philosophy (Simple, Short, Structured, Standardized) and grounded in the ICF framework. A cross-sectional validation study was conducted in Vietnam involving 208 participants (104 shoulder, 104 knee), including asymptomatic and clinical groups. Internal consistency (Cronbach's α), test–retest reliability (ICC), convergent validity (with SPADI, KOOS-12, and EQ-5D-5L), discriminant validity, and feasibility were analyzed.

Results: Both scales demonstrated excellent reliability (KFS: α = 0.96, ICC = 0.95; SFS: α = 0.97, ICC = 0.95). Strong convergent validity was observed with reference instruments (ρ = 0.82–0.83) and the EQ-5D-5L index (ρ ≥ 0.80). Receiver operating characteristic analysis showed excellent discriminative ability, with preliminary cut-off values of ≤ 21 for KFS and ≤ 19 for SFS. More than 84% of participants rated the assessment as helpful.

Conclusion: The S-Series Functional Scales (SFS and KFS) provide a concise, standardized, and ICF-aligned PROM system with promising preliminary psychometric properties, supporting their use in clinical practice and digital health research.

Keywords: *S-Series; PROMs; shoulder; knee; validation; digital health.*

1. INTRODUCTION

1.1. Burden of Disease and the Need for Assessment

Musculoskeletal (MSK) disorders represent a substantial global health burden and are a leading cause of years lived with disability worldwide [1]. Accurate assessment of functional status is therefore essential for clinical decision-making, outcome evaluation, and health system planning.

In Vietnam, knee disorders represent a major component of musculoskeletal morbidity. Knee osteoarthritis is highly prevalent, with radiographic evidence reported in approximately 34.2% of the general population [2]. In addition, anterior cruciate ligament (ACL) injuries are common, particularly among adolescents and young adults, with an estimated incidence of 68.6 per 100,000 person-years [3]. Together, degenerative and post-traumatic knee conditions highlight the need for reliable and efficient tools to monitor knee function across clinical contexts.

Shoulder disorders similarly constitute a significant source of pain and functional limitation. A recent systematic review reported a global median point prevalence of shoulder pain of approximately 16% in the general population [4,5]. With the increasing use of orthopedic interventions such as arthroscopy and joint replacement, accurate postoperative and longitudinal functional assessment is essential. In this context, patient- reported outcome measures (PROMs) are widely used to capture functional status from the patient's perspective.

1.2. Limitations of Current PROMs

Despite the availability of well-established joint-specific PROMs-such as the Shoulder Pain and Disability Index (SPADI) [6] and the Knee injury and Osteoarthritis Outcome Score (KOOS) [7]-several limitations constrain their use in routine clinical practice and digital health settings:

- Length and Response Burden: Many instruments include 10 - 42 items (e.g., the full KOOS), limiting feasibility for high-frequency or long-term

monitoring.

• Lack of Structural Consistency: Differences in item structure and domain composition across joint-specific scales hinder data harmonization and comparative analyses within digital health systems.

• Limited Suitability for Self-Monitoring: Most existing PROMs were designed for episodic clinical assessment rather than continuous, patient-driven self-monitoring during rehabilitation.

These limitations highlight the need for a concise, standardized PROM system capable of supporting both clinical care and digital health applications.

1.3. Proposed Solution: The S-Series Functional Scales

To address these challenges, we developed the S-Series Functional Scales, a system of ultra-short PROMs guided by a "4S" philosophy: Simple, Short, Structured, and Standardized. The S-Series is designed to facilitate efficient functional assessment while maintaining conceptual clarity and clinical relevance.

The system is grounded in the World Health Organization's International Classification of Functioning, Disability and Health (ICF) framework [8]. Key features include:

• Standardized Structure: A uniform six-item core applied across joints, covering functional perception (activity, participation) and mechanical symptoms (pain, perceived range of motion, perceived strength/control).

• Modular Design: Joint-specific modules built on a shared framework, enabling cross-joint comparability.

• Dual Versions: The S-Series framework includes a self-report patient version (S-version) and a clinician-assessed version (C-version, clinROM) integrating structured physical examination elements such as range of motion and muscle performance assessment.

• Patient Engagement: Optional self-check elements intended to enhance patient involvement and understanding of functional limitations.

1.4. Study Objectives

The present study aims to conduct an initial evaluation of the measurement properties and

feasibility of the Shoulder Function Scale (SFS-S) and Knee Function Scale (KFS-S).

The specific objectives are to:

1. Assess reliability, including internal consistency (Cronbach's alpha) and test–retest reliability (intraclass correlation coefficient, ICC).

2. Evaluate validity, including:

• Known-groups (discriminant) validity by comparing asymptomatic and symptomatic participants.

• Convergent validity through correlations with established reference instruments (SPADI, KOOS-12) and health-related quality of life (EQ-5D-5L).

3. Explore discriminative ability and preliminary cut-off values for screening purposes using receiver operating characteristic (ROC) analysis.

4. Assess feasibility, based on user feedback and acceptance rates.

Note: Although the S-Series framework includes additional joint modules (e.g., a lumbar spine module), this study reports validation data exclusively for the shoulder and knee modules.

2. MATERIALS AND METHODS

2.1. Scale Development

The S-Series Functional Scales were developed as a unified system of brief patient-reported outcome measures (PROMs) based on the World Health Organization's International Classification of Functioning, Disability and Health (ICF) [8]. The development followed a 4S design philosophy: Simple, Short, Structured, and Standardized.

Initial item development involved iterative conceptual mapping based on the ICF framework, review of domains commonly represented in established musculoskeletal PROMs (including SPADI and KOOS), and clinical rehabilitation experience. Draft items were informally reviewed by rehabilitation clinicians and further refined during a pilot feasibility phase involving Vietnamese users prior to the present validation study. However, a formal COSMIN-based content validity process, including Delphi consensus procedures and structured cognitive interviewing, was not performed in this preliminary phase.

Each S-Series scale consists of six items covering two conceptual domains:

1. Perceived Functioning, including activity, participation, and satisfaction

2. Mechanical Symptoms (Pain, Perceived Range of Motion, and Perceived Strength/Control)

All items are scored on a 5-point Likert scale (0 - 4), yielding a total score ranging from 0 to 24, with higher scores indicating better function. In the initial development phase, three self-reported instruments were designed: the Shoulder Function Scale (SFS-S), Knee Function Scale (KFS-S), and Lumbar Function Scale (LFS-S). The present study reports validation results for the SFS-S and KFS-S only.

2.2. Study Design, Setting, and Participants

A cross-sectional validation study with a test–retest component was conducted at Hue University

of Medicine and Pharmacy and Hue University Hospital, Vietnam. Data collection took place between October 2025 and January 2026.

Participants were recruited using a convenience sampling strategy and comprised two groups:

- Asymptomatic group: Medical students and postgraduate trainees at Hue University of Medicine and Pharmacy who reported no current shoulder or knee pain, injury, or functional limitation.
- Clinical group: Patients with shoulder or knee pain or functional limitation receiving outpatient rehabilitation care at the Department of Rehabilitation, Hue University Hospital. The knee cohort included a significant proportion of post-operative patients (e.g., post-ACL reconstruction, meniscal repair) requiring functional monitoring.

Eligible participants were aged 18 years or older, able to read Vietnamese, and willing to participate. Exclusion criteria included red-flag conditions (e.g., suspected malignancy, infection), systemic inflammatory disease, neurological disorders, cognitive impairment, or other conditions that could invalidate self-reported functional assessment.

The target sample size was determined pragmatically based on recommendations for preliminary psychometric validation studies, which commonly suggest at least 5–10 participants per item for reliability and construct validity assessment.

2.3. Data Collection Procedure

Data were collected using a digital self-administration approach. Participants accessed the survey by scanning a QR code linking to a Google Forms questionnaire. The survey was anonymous and did not collect personally identifiable information.

- Baseline (T0): Participants completed demographic questions, followed by the relevant S-Series scale (SFS or KFS) and comparator instruments. Each S-Series scale included optional simple functional self-checks with standardized safety instructions.
- Test-retest (T1): A subsample of participants was invited to complete the same S-Series scale a second time after an interval of approximately 5–7 days.

2.4. Measures

- Knee Cohort: Knee Function Scale (KFS-S), the KOOS-12 (reference standard), and the EQ-5D-5L. Use of the KOOS-12 was authorized under an academic user agreement with Mapi Research Trust (Work Order No. 2515700).
- Shoulder Cohort: Shoulder Function Scale (SFS-S), the SPADI (reference standard), and the EQ-5D-5L.
- General Health: Licensed Vietnamese versions of the EQ-5D-5L were used (Registration ID: 75703). Health state profiles were converted into index values using the specific Vietnamese value set [11].
- Feasibility Measures: Feasibility was assessed based on user acceptability. Participants answered specific questions regarding the clarity of instructions (rated on a 4-point scale from “very difficult to understand” to “very easy to understand”), whether they completed the optional self-test (yes/no), and the perceived helpfulness of the assessment (yes/no).

2.5. Statistical Analysis

All analyses were performed using R Statistical Software (R Foundation for Statistical

Computing, Vienna, Austria).

- Descriptive statistics: Continuous variables were summarized using means and standard deviations or medians and interquartile ranges, as appropriate. Floor and ceiling effects were calculated as the proportion of participants achieving the minimum or maximum possible scores.
- Reliability: Internal consistency was evaluated using Cronbach’s alpha. Test-retest reliability was assessed using the intraclass correlation coefficient (ICC; two-way random-effects model, absolute agreement; ICC[A,1]).
- Validity:

○ Known-groups validity: Differences between asymptomatic and symptomatic participants were examined using the independent t-test or Wilcoxon rank-sum test, depending on data distribution.

○ Convergent validity: Spearman’s rank correlation coefficients (ρ) were calculated between S-Series total scores and reference instruments (SPADI, KOOS-12) and EQ-5D-5L.

- Discriminative ability: Receiver operating characteristic (ROC) curve analysis was conducted to evaluate the ability of the scales to distinguish between asymptomatic and symptomatic individuals, with area under the curve (AUC) values reported.

A two-sided p-value < 0.05 was considered statistically significant.

2.6. Ethics Approval

The study was approved by the Institutional Review Board of Hue University of

Medicine and Pharmacy, Vietnam (Approval No. H2025/627). All participants received detailed study information and provided written informed consent prior to participation. The study was conducted in accordance with the principles of biomedical research ethics. Participants’ confidentiality was strictly maintained, and participation was voluntary, with the right to withdraw at any time without consequence.

3. RESULTS

3.1. Demographic Characteristics

The final analysis included 208 participants, with 104 individuals in the Knee module

(KFS) and 104 individuals in the Shoulder module (SFS).

In both cohorts, participants were classified into asymptomatic control and symptomatic clinical groups. The demographic characteristics of the knee and shoulder cohorts are summarized in Table 1.

The knee clinical cohort primarily included postoperative rehabilitation cases, particularly following ACL reconstruction and meniscal procedures, with a smaller number of degenerative musculoskeletal conditions. The shoulder cohort mainly consisted of individuals with rotator cuff-related disorders and adhesive capsulitis encountered in outpatient rehabilitation practice.

Table 1. Characteristics of the Study Population

Characteristic	Knee Cohort (n = 104)	Shoulder Cohort (n = 104)
Control Group (Asymptomatic)	n = 79 (76.0%)	n = 83 (79.8%)
- Median Age (IQR)	22 (3)	22 (3)
Clinical Group (Symptomatic)	n = 25 (24.0%)	n = 21 (20.2%)
- Median Age (IQR)	41 (18)	56 (13)
Gender (Female %)	69 (66.3%)	71 (68.3%)
Age range (years)	20-70	20-70

3.2. Reliability

Both the Knee Function Score (KFS) and Shoulder Function Score (SFS) demonstrated excellent internal consistency and test–retest reliability (Table 2).

Table 2. Reliability Statistics for S-Series Scales

Scale	Cronbach's α	Test-Retest ICC (95% CI)	Interpretation [12]
KFS (Knee)	0.96	0.95 (0.90 - 0.98)	Excellent
SFS (Shoulder)	0.97	0.95 (0.89 - 0.98)	Excellent

Cronbach's alpha values above 0.90 indicate excellent internal consistency for both scales. Corrected item–total correlations were high for both instruments, ranging from 0.87 to 0.93 for the KFS and 0.85 to 0.96 for the SFS. Importantly, the deletion of any individual item did not improve the overall alpha coefficient for either scale, confirming the robust homogeneity of the items.

Test–retest reliability was assessed in subsamples of 29 participants for the KFS and 42 participants for the SFS. Median scores remained stable between baseline and retest assessments for both scales, with ICC values of 0.95 confirming high score stability over time.

3.3. Validity

Known-Groups Validity:

Clear score separation was observed between asymptomatic and symptomatic participants for both joints.

For the knee cohort, median KFS scores were significantly higher in asymptomatic participants (median 24, IQR 0) than in symptomatic participants (median 11, IQR 4), confirming strong known-groups validity ($p < 0.001$; effect size $r = 0.87$; Figure 1).

Similarly, in the shoulder cohort, asymptomatic participants demonstrated markedly higher SFS total scores (median 24, IQR 0) compared with symptomatic participants (median 9, IQR 6), with a large effect size ($r = 0.83$; $p < 0.001$).

Ceiling effects were evident in asymptomatic participants, particularly for the SFS (approximately 63%), whereas floor effects were minimal for both scales (< 2%).

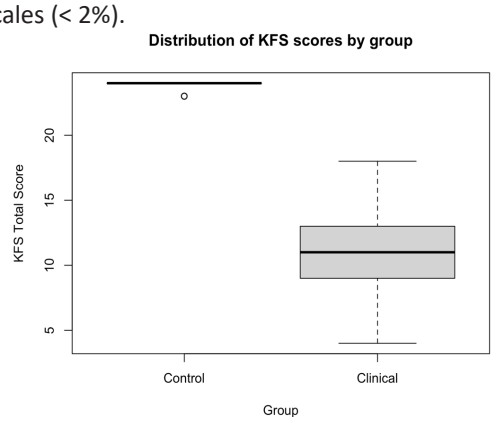

Figure 1. Boxplot of KFS scores demonstrating clear separation between asymptomatic and symptomatic groups.

Convergent Validity:

The S-Series scales showed strong correlations with established legacy instruments (Table 3).

Table 3. Convergent Validity (Spearman's ρ)

S-Series Scale	Reference Scale	Correlation Coefficient	P-value
KFS (Knee)	KOOS-12 Total	0.82	< 0.001
	EQ-5D Index (VN)	0.82	< 0.001
	EQ-VAS	0.49	< 0.001
SFS (Shoulder)	SPADI Total	-0.83	< 0.001
	SPADI Pain	-0.80	< 0.001
	EQ-5D Index (VN)	0.80	< 0.001
	EQ-VAS	0.58	< 0.001

For the KFS, convergent validity was supported by strong correlations with KOOS-12 and the EQ-5D-5L index, and a moderate correlation with EQ-VAS.

For the SFS, strong inverse correlations were observed with SPADI scores and the EQ-5D-5L index, alongside a moderate association with EQ-VAS. The relationship between shoulder function and shoulder-specific disability is illustrated in Figure 2.

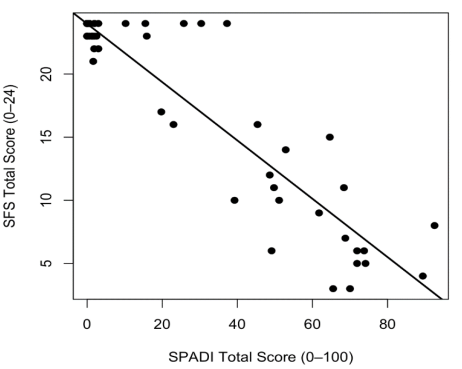

Figure 2. Scatterplot illustrating the strong inverse association between SFS total scores and SPADI scores

3.4. ROC Analysis and Cut-off Points

ROC curve analysis demonstrated excellent discriminative ability for both scales.

For the KFS (Knee), the area under the curve (AUC) was 1.00 (95% CI: 1.00 - 1.00). The optimal cut-off score for distinguishing symptomatic from asymptomatic individuals was ≤ 21 points, yielding 100% sensitivity and 100% specificity.

For the SFS (Shoulder), the AUC was 0.996 (95% CI: 0.987 - 1.000), indicating near-perfect discrimination. Based on the Youden index, the optimal cut-off score was ≤ 19 points, with 96% sensitivity and 100% specificity.

The extremely high AUC values likely reflect the clear separation between the asymptomatic student cohort and the clinical patient population in this validation sample.

3.5. Feasibility

The S-Series scales proved highly feasible for routine use (Table 4).

Table 4. Feasibility Indicators

Indicator	KFS (Knee)	SFS (Shoulder)
Instruction Clarity (Median)	4 / 4	4 / 4
Self-test Completion	83.5%	70.4%
Rated as "Helpful"	84.5%	98.1%

Participants rated the instruction clarity highly. For the KFS, 83.5% of participants completed the self-test independently, and 84.5% reported that the assessment was helpful. For the SFS, 70.4% completed the self-test independently, and 98.1% perceived the assessment as helpful. No major difficulties in understanding or completing the questionnaires were reported.

4. DISCUSSION

This study describes the development and preliminary validation of the S-Series Functional Scales, a unified patient-reported outcome measure (PROM) system for multi-joint musculoskeletal assessment. Using the Shoulder Function Scale (SFS) and Knee Function Scale (KFS) as representative modules, the results provide consistent evidence supporting the validity, reliability, and feasibility of the S-Series in a Vietnamese population.

4.1. Reliability and Conceptual Integrity:

Across both scales, internal consistency was excellent, with Cronbach's alpha values exceeding 0.96. These findings indicate strong coherence among the six core domains-activity, participation, satisfaction, pain, perceived range of motion, and perceived strength/control-supporting the conceptual integrity of the ICF-based structure. Although very high alpha values may raise concerns regarding potential item redundancy [15], this phenomenon is commonly observed in short, unidimensional functional scales, particularly in samples with restricted score variability. Importantly, item–total correlations were uniformly high, and removal of any item did not meaningfully improve

internal consistency, supporting the contribution of each item to the overall construct.

Test–retest reliability was also excellent for both scales, with ICC values exceeding 0.95. These findings support the temporal stability of the S-Series instruments under stable clinical conditions despite their ultra-short structure.

4.2. Validity and Discriminatory Power:

Construct validity was strongly supported for both the Shoulder Function Scale (SFS) and the Knee Function Scale (KFS). Known-groups validity analyses demonstrated large to very large effect sizes ($r = 0.83–0.87$), indicating a clear ability of both instruments to discriminate between asymptomatic individuals and patients with shoulder or knee disorders.

Discriminatory performance was further supported by ROC analyses, which showed excellent classification accuracy for both scales. The SFS demonstrated a near-perfect area under the curve (AUC = 0.996), while the KFS achieved perfect separation between groups (AUC = 1.00). These results suggest that the S-Series scales demonstrate excellent discriminatory ability for clinically meaningful functional impairment in this validation sample. However, the extremely high AUC values should be interpreted with caution, as they likely partially reflect the contrast between a predominantly healthy volunteer group and a clearly symptomatic clinical cohort, resulting in substantial separation between groups. Consequently, the observed ROC performance may overestimate discrimination in more clinically nuanced scenarios, such as differentiating mild versus moderate impairment within patient populations. Future studies should further evaluate discriminatory performance across diagnostically homogeneous cohorts and varying levels of symptom severity.

Convergent validity was confirmed through strong correlations with established condition-specific instruments, including the SPADI for shoulder disorders and the KOOS-12 for knee conditions, as well as with the EQ-5D-5L index. As hypothesized, correlations with EQ-VAS were more moderate, consistent with the broader and less condition-specific nature of this global self-rated health measure.

Exploratory cut-off points derived from ROC analyses suggest potential thresholds for screening purposes, although these values should be considered preliminary and require confirmation in larger prospective studies and independent validation samples.

4.3. Ceiling Effects and Clinical Context:

Ceiling effects were observed for both SFS and KFS, particularly among asymptomatic participants (63-65%). This distributional pattern was expected given the high proportion of healthy individuals in the sample and the functional focus of the scales. Rather than indicating a measurement flaw, these ceiling effects suggest that the S-Series is optimized for identifying functional impairment and monitoring recovery in clinical populations, where scores are distributed across a wider range.

However, high ceiling effects may reduce sensitivity for detecting subtle functional improvements among individuals who already demonstrate near-normal performance. This limitation may be particularly relevant in athletic populations or during later rehabilitation stages. Similar ceiling phenomena have been reported in early validation studies of other musculoskeletal PROMs and tend to diminish in samples with greater

symptom severity or longitudinal follow-up.

4.4. Feasibility and Potential Role in Digital Health

Feasibility findings were consistently favorable across both scales. Participants reported high instruction clarity, high perceived usefulness, and high rates of independent completion. The inclusion of a simple guided self-test component for perceived strength/control was generally well accepted and may enhance patient engagement and understanding of functional limitations. These characteristics align with the original design goals of the S-Series, particularly its intended use in digital health, tele-rehabilitation, and high-frequency self-monitoring contexts.

The S-Series differs from existing shortened PROMs in several important respects. Rather than creating isolated joint-specific short forms, the S-Series was designed as a unified cross-joint framework with standardized structure, scoring logic, and conceptual domains. This architecture may facilitate implementation in digital rehabilitation systems, longitudinal self-monitoring, cross-joint comparison, and simplified integration across multiple musculoskeletal conditions.

Importantly, the S-Series was not intended to replace comprehensive region-specific PROMs such as SPADI or KOOS. Instead, the ultra-short six-item structure was designed to support rapid repeated assessment, scalable digital implementation, and routine functional monitoring within rehabilitation and telehealth workflows.

Limitations

Several limitations should be acknowledged. The use of convenience sampling and the predominance of young asymptomatic participants, primarily medical students, may limit the generalizability of the findings and have contributed to the large between-group differences observed in the present study. In contrast, the clinical cohort consisted of older patients receiving rehabilitation care. The test–retest analyses were based on modest subsamples, and responsiveness to clinical change was not assessed.

Future multicenter longitudinal studies should evaluate responsiveness, minimal clinically important change, measurement error parameters (e.g., minimal detectable change), and performance across diagnostically homogeneous subgroups, broader demographic populations, and different cultural contexts.

5. CONCLUSION

The S-Series Functional Scales (SFS and KFS), as representative joint-specific modules examined in this study, represent a concise, standardized, and ICF-aligned PROM system with strong preliminary psychometric properties. The consistent findings across the shoulder and knee modules support the potential broader applicability of the S-Series as a unified framework for musculoskeletal functional assessment in both clinical practice and digital health research, pending further validation of additional joint modules.

Conflict of Interest: The authors declare no conflicts of interest.

REFERENCES

1. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390(10100):1211-1259.

2. Ho-Pham LT, Lai TQ, Mai TV, Doan TV, Nguyen TV. Prevalence of radiographic osteoarthritis of the knee and its relationship to self-reported pain. PLoS One. 2014;9(4):e94563.

3. Sanders TL, Maradit Kremers H, Bryan AJ, Larson DR, Dahm DL, Levy BA, 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.

4. Lucas J, van Doorn P, Hegedus E, Lewis J, van der Windt D. A systematic review of the global prevalence and incidence of shoulder pain. BMC Musculoskelet Disord. 2022;23:1073.

5. Luime JJ, Koes BW, Hendriksen IJ, Burdorf A, Verhagen AP, Miedema HS, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. Scand J Rheumatol. 2004;33(2):73-81.

6. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. Arthritis Care Res. 1991;4(4):143-149.

7. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS). Health Qual Life Outcomes. 2003;1:64.

8. World Health Organization. International classification of functioning, disability and health: ICF. Geneva: World Health Organization; 2001.

9. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. KOOS-12: A shorter version of KOOS. J Orthop Sports Phys Ther. 2016;46(12):1059-1068.

10. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-1736.

11. Mai VQ, Sun S, Minh HV, Luo N, Giang KB, Lindholm L, et al. An EQ-5D-5L value set for Vietnam. Qual Life Res.

2020;29(8):1923-1933.

12. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155-163.

13. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.

14. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol 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. 2010;63(7):737-745.

15. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53-55.

APPENDIX

Appendix Table 1. Conceptual ICF Alignment of the S-Series Core Items

Item	Construct	ICF Domain	Example ICF Components
Item 1	Daily activity performance	Activities	Mobility, self-care, daily activities
Item 2	Social/work participation	Participation	Work, community, and social participation
Item 3	Satisfaction with function	Personal perception of functioning	Global perceived functional status
Item 4	Pain symptoms	Body Functions	Sensation of pain
Item 5	Perceived range of motion	Body Functions	Mobility of joint functions
Item 6	Perceived strength/ control	Body Functions	Muscle power and movement control