Original Article

Validity and reliability of EQ-5D-5L among patients with axial spondyloarthritis in Singapore

Jun Jie Benjamin Seng[1] , Yu Heng Kwan[2] , Warren Fong[1,3,4] , Jie Kie Phang[3] , Nai Lee Lui[3] , Julian Thumboo[2,3,4] , Ying Ying Leung[1,3,4]

Abstract

Objective: This study aims to evaluate the reliability and validity of EuroQOL-5 Dimensions-5 Levels (EQ-5D-5L) among patients with axial spondyloarthritis (SpA) in Singapore.

Methods: A cross-sectional study was conducted involving patients with axial SpA in an Asian tertiary hospital from 2017 to 2018. This study followed the COnsensus-based Standards for selection of health Measurement Instruments framework. Construct validity was evaluated by testing 22 a priori hypotheses with other patient-reported outcomes measures. Cronbach's alpha was used to estimate the internal consistency of the EQ-5D-5L, while its test-retest reliability was assessed using weighted kappa and the intraclass correlation coefficient (ICC). The measurement error was assessed by analyzing minimal detectable change (MDC).

Results: The median age of included patients (n=118) was 35 years (interquartile range: 28, 49). Ninety-six (81.4%) patients were male, while 112 (94.9%) patients were of Chinese ethnicity. The EQ-5D-5L demonstrated good internal consistency with a Cronbach's alpha of 0.79. The test-retest reliability of the EQ-5D-5L was good with a weighted kappa of ≥0.61 for mobility, self-care, usual activities, and anxiety/depression; the ICC was 0.92 and 0.99 for the EQ-5D-5L index and visual analog scale (VAS) scores, respectively. The weighted kappa for the EQ-5D-5L pain/discomfort was moderate [0.53, 95% confidence interval: 0.41-0.60]. The MDC for EQ-5D-5L index and VAS scores was 0.06 and 4.5, respectively. Convergent validity was supported as all hypotheses were confirmed in the results.

Conclusion: This study supports EQ-5D-5L as a valid and reliable instrument for assessing health-related quality of life among patients with axial SpA in Singapore.

Keywords: Ankylosing, health care, outcome assessment, psychometrics, rheumatology, Singapore, spondylitis, quality of life

ORCID iDs of the authors:
J.J.B.S. 0000-0002-3039-3816;
Y.H.K. 0000-0001-7802-9696;
W.F. 0000-0003-1891-1892;
J.K.P. 0000-0003-1357-8568;
N.L.L. 0000-0003-4275-1737;
J.T. 0000-0001-6712-5535;
Y.Y.L. 0000-0001-8492-6342.

[1] Department of Medicine, Duke-NUS Medical School, 8 College Road, Singapore
[2] Program in Health Services and Systems Research, Duke-NUS Medical School, 8 College Road, Singapore
[3] Department of Rheumatology and Immunology, Singapore General Hospital, Outram Road, Singapore
[4] Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, 12 Science Drive 2, Singapore

Address for Correspondence:
Yu Heng Kwan; Program in Health Services and Systems Research, Duke-NUS Medical School, 8 College Road, Singapore

E-mail: yuheng@u.duke.nus.edu

Introduction

Spondyloarthritis (SpA) encompasses a group of chronic, debilitating inflammatory diseases that result in severe physical limitation and poor quality of life for patients (1). Globally, the prevalence of SpA has been estimated between 0.07% and 1.4% (2, 3). Common clinical features across the spectrum of SpA include axial joint inflammation, asymmetric oligoarthritis, dactylitis, and enthesitis (4). Owing to the lack of disease-modifying treatment available for SpA, its management primarily focuses on improving physical function and pain control to allow preservation of patients' health-related quality of life (HrQoL).

Patient-reported outcome measures (PROMs) are widely used in the management of patients with SpA. Owing to the poor correlation of disease severity with clinical parameters such as C-reactive protein and erythrocyte sedimentation rate (5, 6), tools such as Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and Bath Ankylosing Spondylitis Function Index (BASFI) are used clinically to achieve a more comprehensive understanding of patients with SpA. Additionally, PROMs that evaluate patients' HrQoL are also gaining popularity and importance as they permit physicians to assess the health status of each patient. HrQoL is increasingly recognized as a multidimensional construct of patients' physical function, psychological state, and social relationship (7). Instruments used for the evaluation of patients' HrQoL can be categorized as "generic" or "disease-specific," which have their inherent limitations and strengths. Generic tools allow the quantification and comparison of HrQoL between patients from a general population but are less sensitive to aspects related to par-

Seng et al. Validity of EQ-5D-5L in axial SpA patients

Eur J Rheumatol 2020; 7(2): 71-8

ticular diseases. In contrast, disease-specific instruments are more sensitive but limit comparisons between patients with different comorbidities.

EuroQOL-5 Dimensions-5 Levels (EQ-5D-5L) is a popular instrument used for the assessment of generic HrQoL, cost-utility of healthcare interventions, and computation of quality-adjusted life years for patients. It comprises two components—a health descriptive component and a visual analog scale (VAS). A recent study performed in Hong Kong had shown that EQ-5D-5L demonstrated acceptable psychometric properties for the evaluation of Chinese patients with SpA (8). Although it has been utilized to evaluate the health state of patients with SpA in Europe and Asia (9), no study has evaluated its psychometric properties among patients with axial SpA in Singapore (10, 11). As there may be cross-cultural differences in HrQoL among patients with SpA in different countries, the goal of this study was to evaluate the reliability and validity of the EQ-5D-5L in patients with axial SpA living in a multiethnic Asian country. The COnsensus-based Standards for the selection of health Measurement Instrument (COSMIN) guidelines were followed and adhered to during the assessment (12).

Methods

Study design
We performed a cross-sectional study in a tertiary hospital in Singapore. It involved all patients with axial SpA seen at the specialty rheumatology clinic from 2017 to 2018. Diagnosis of axial SpA was made as per Assessment of Spondyloarthritis International Society Classification (13, 14). All patients included in the study provided informed consent. The study protocol was vetted and approved by the institution review board.

Main Points

- The EuroQOL-5 Dimensions-5 Levels (EQ-5D-5L) is an instrument commonly used to assess health related quality of life (HrQOL) among axial spondyloarthritis (SpA) patients.

- Validation of EQ-5D-5L has not been performed for its use among axial SpA patients in Singapore.

- EQ-5D-5L was shown to be a valid and reliable instrument for assessment of HrQOL among axial SpA patients.

Information pertaining to patient sociodemographic and clinical characteristics was collected. Additionally, PROMs that included the EQ-5D-5L, 36-item Short Form Survey (SF-36), BASDAI, BASFI, Health Assessment Questionnaire-Disability Index (HAQ-DI), Patient Global Assessment (PGA) of disease activity, pain scores, and Work Productivity and Activity Impairment Questionnaire: Spondyloarthritis (WPAI:SpA) were self-administered by patients. We excluded illiterate patients and patients who did not complete the EQ-5D-5L questionnaire.

The EQ-5D-5L data from all stable patients with axial SpA were collected for test-retest reliability. A patient with axial SpA was classified as stable, if there were no changes to the medication therapy and no disease flares in the past 3 months of follow-up (15). Thereafter, each patient self-administered the EQ-5D-5L questionnaire at home at 1 week and 2 weeks after the baseline assessment. The time interval selected was as per the recommendations by Deyo et al. (16), which allows an adequate length to minimize recall effects while providing an indicator of nonspecific score changes that occur naturally in a PROM instrument. The completed EQ-5D-5L self-classifier questionnaires were subsequently mailed to the clinic using the envelope provided.

Patient-reported outcome measures

EuroQOL-5 dimensions-5 levels
The EQ-5D-5L is a PROM that measures HrQoL and comprises a health descriptive component and a VAS (17). The descriptive component evaluates five items related to health which encompass mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. There are five possible responses for each item: no problems, slight problems, moderate problems, severe problems, and extreme problems. Scores from the five items can be used to derive a single utility score. This composite score ranges from -1.00 to 1.00, representing the worst possible health state to the perfect health state (18).

The VAS is a 20-cm vertical scale that is scored from 0 to 100 points. Similarly, a score of 0 and 100 indicate the "worst imaginable health state" and "best imaginable health state," respectively.

36-item short form survey
SF-36 is a generic HrQoL instrument that assesses eight domains of perceived health and has been validated for use in Singapore (19). The domains are, namely, physical functioning, role limitation due to physical problem, bodily pain, general health, vitality, social functioning, role limitation due to emotional problem, and mental health. Scores obtained from the eight subscales are subsequently consolidated into two normalized summary measures—physical component summary (PCS) and mental component summary (MCS) scores (20). The PCS and MCS scores range from 0 to 100, with lower scores representing poorer HrQoL.

Health assessment questionnaire-disability index
The HAQ-DI is a validated, self-administered measure of disability that comprises eight domains (21). It evaluates the level of difficulty patients faced when executing activities such as arising, eating, hygiene, and grooming. Each item is scored from 0 to 3 and a lower score indicates lesser disability (22).

Other PROMs
The BASDAI is a self-reported tool that is used for quantifying and assessing disease activity among patients with SpA (23). It comprises six items that look at patients' fatigue, spinal pain, arthralgia, enthesitis, duration and severity of morning stiffness. The responses range from 0 to 10, with a lower score indicating lower disease activity. The BASFI is self-reported, 8-item instrument (range 0-10) that is used to assess the functional status of patients with SpA (23). It evaluates how symptoms arising from SpA affect patients' daily activities such as their ability to put on socks without help and picking up pens from the floor by bending forward. Similarly, higher scores are an indicator of poorer function. Additionally, the pain and PGA of disease activity (24) scores of each patient were collected. The two items were scored from 0 to 100 and higher scores reflect poorer pain control and well-being. Finally, information pertaining to patients' WPAI:SpA was also collected (25). The WPAI:SpA (response range: 0-100%) evaluates four areas related to the overall impact of SpA on work productivity and daily activities i.e., presenteeism, absenteeism, work productivity loss, and ac-

Eur J Rheumatol 2020; 7(2): 71-8

Seng et al. Validity of EQ-5D-5L in axial SpA patients

Table 1. Sociodemographic, clinical characteristics, and PROM scores of patients with axial SpA[a] (n=118).

Characteristics	Median (Interquartile range) or n (%)
Age	35 (28, 49)
Gender	
Male	96 (81.4)
Race	
Chinese	112 (94.9)
Malay	1 (0.8)
Indian	1 (0.8)
Others[b]	4 (3.4)
Highest education qualification	
No formal education	1 (0.8)
Primary	2 (1.7)
Secondary	17 (14.4)
Tertiary	98 (83.1)
Marital status	
Single	59 (50.0)
Married	54 (45.8)
Divorced	5 (4.2)
Living arrangement	
Staying alone	10 (8.5)
Staying with family/friends	108 (91.5)
Occupation	
Employed	85 (72.0)
Unemployed	7 (5.9)
Retired	9 (7.6)
Others[c]	17 (14.4)
Disease duration, years	6.6 (1.5, 12.2)
PROM score (Range)	
SF-36	
Physical functioning (0-100)	85 (70, 95)
Role-physical (0-100)	78.1 (62.5, 100)
Bodily pain (0-100)	62 (51, 74)
General health (0-100)	57 (45, 72)
Vitality (0-100)	62.5 (50, 75)
Social functioning (0-100)	87.5 (75, 100)
Role-emotional (0-100)	83.3 (75, 100)
Mental health (0-00)	80 (65, 90)
Physical Component Summary (Median: 50, SD 10)	46.2 (38.6, 53.0)
Mental Component Summary (Median: 50, SD 10)	47.5 (39.8, 53.9)
HAQ-DI (0-3)	1.1 (1, 1.4)
BASDAI (0-10)	2.1 (1.5, 4.2)
BASFI (0-10)	1.1 (0.3, 2.8)
PGA (0-10)	3 (1, 4)
Pain (0-10)	2 (1, 4)
WPAI:SpA	
Presenteeism (0-100)	20 (0, 30)
Absenteeism (0-100)	0 (0, 0)
Work productivity loss (0-100)	20 (10, 47.7)
Activity loss (0-100)	20 (0, 30)

[a]As diagnosed with the 2011 Assessment of Spondyloarthritis International Society (ASAS) criteria.
[b]Others refer to Bamar, Filipino, and Ceylonese ethnicity.
[c]Others refer to full-time students, patients serving full-time military training or homemakers.
BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; GCE: general certificate of education; PGA: Patient Global Assessment; HAQ-DI: Health Assessment Questionnaire-Disability Index; PROM: patient reported outcome measures; PSLE: primary school leaving examination; SD: standard deviation; WPAI:SpA: Work Productivity and Activity Impairment Questionnaire: Spondyloarthritis.

tivity impairment (25). Higher percentages indicate greater impairment on productivity. The use of WPAI:SpA has been validated for use among patients with SpA in Singapore (Results unpublished).

Statistical analyses

Data analyses in this study were conducted using the Stata software, version 14.0 (Stata Corporation; College Station, Texas, USA). Test for normality of data was performed using Shapiro-Wilk's test. Descriptive statistics were expressed as mean±standard deviation (SD), median [interquartile range (IQR)], or number (%), where appropriate.

In this study, the preference weights from Singapore were utilized to derive the EQ-5D-5L index score (26). In accordance to the COSMIN checklist (Supplementary File 1), the percentage of missing items, description of how missing items were processed, the distribution of EQ-5D-5L index and VAS scores (including floor and ceiling effect) were reported for interpretability. Floor and ceiling refer to the proportion of observations at the lowest and highest possible value, respectively. A floor or ceiling percentage that is greater than 15% is considered significant (27).

Pertaining to sample size computation, there are currently no guidelines on sample size computation for validation of HrQoL tools. However, it is generally recommended by the COSMIN guidelines to have at least 50-100 respondents (28, 29).

Validity refers the extent to which the tool is able to quantify what it seeks to measure (30, 31). A valid tool should permit the differentiation of HrQoL among patients with unique disease characteristics. The construct validity between EQ-5D-5L subscales and other PROM scores was evaluated using Spearman rank or Pearson's correlation coefficients, where appropriate. High (r=0.5-0.8) and moderate correlation coefficients (r=0.3 to <0.5) indicate that the scores from two PROMs are correlated, while low correlation coefficients (r≤0.3) indicate that the PROMs are quantifying different constructs (32). The validities of convergent construct were evaluated with 22 a priori hypotheses based on the literature search and experience from clinical practice. Construct validities of the EQ-5D-5L was supported if at least 75% of the results are in accordance with hypotheses (33). A p-value of <0.001 was considered statistically significant after applying Bonferroni's correction.

The hypotheses were as follows:
- The EQ-5D-5L index (18, 34) and VAS scores (35) are positively and moderately correlated with SF-36 PCS and SF-36 MCS.
- The EQ-5D-5L index and VAS scores are negatively correlated and highly correlated with PGA (36) scores.
- The EQ-5D-5L index and VAS scores are negatively correlated and moderately correlated with summary scores for BASDAI (37), BASFI (37), HAQ-DI (38), pain (39), WPAI absenteeism, presenteeism, work productivity loss, and activity impairment (40).

Reliability refers to the overall consistency of an instrument. Observed values in a reliable instrument are true within acceptable errors of measurements (30). This was evaluated primarily through internal consistency on the assumption that every item was similar and assessed a single construct. Although the EQ-5D-5L comprises several dimensions with single items, the utilization of Cronbach's alpha for assessing its reliability has been studied in multiple studies (41-44). Internal consistency was hence estimated using Cronbach's alpha and was only supported if it was more than 0.70 (45). The

test-retest reliability was investigated using weighted kappa for the five items of the EQ-5D-5L and intraclass correlation coefficient (ICC) (two-way mixed effects model) with a 95% confidence interval (CI) for the EQ-5D-5L VAS and index scores (46). A weighted kappa of ≤0.2, 0.21-0.4, 0.41-0.6, 0.61-0.8, and ≥0.81 indicated poor, fair, moderate, good, and very good agreement of the responses, respectively, between the repeated evaluations (47). Excellent reliability is demonstrated by an ICC that is greater than or equal to 0.70 (33). Measurement error refers to the systematic and random error of PROM score that cannot be attributed to true changes in the construct to be quantified (48). It was assessed by evaluating the minimal detectable change (MDC_{95}) that exceeds measurement error and noise at a 95% CI. The formula used was as follows:

MDC_{95}=1.96×"2×Standard error of Measurement (SEM)=2.77×SEM. The SEM of the EQ-5D-5L index and VAS scores were computed using the formula: SEM=SD of the EQ-5D-5L score in the patient sample×"(1 − reliability of the score) (49). Changes that are smaller than the MDC_{95} are more likely to be attributed to measurement errors, while changes greater than the MDC should be considered as real changes.

Results

Study sample
A total of 118 patients with axial SpA were recruited for this study from 2017 to 2018. No patients were excluded from the study. The sociodemographic and clinical characteristics of the patients and their PROM scores are reported in Table 1. The median age of patients was 35 years (IQR: 28, 49). Majority of the patients were male (n=96, 81.4%), Chinese (n=112, 94.9%), and received at least tertiary education (n=98, 83.1%).

Interpretability of EQ-5D-5L
There was no patient with missing data in the EQ-5D-5L items and the index score was computed for all patients. The distribution of the EQ-5D-5L index scores for all patients at baseline is reported in Figure 1. The median EQ-5D-5L index score was 0.83 (IQR: 0.75, 0.85). There was no floor (n=0) or ceiling effect (n=0) for EQ-5D-5L index score. The distribution of EQ-5D-5L VAS scores for all included patients is shown in Figure 2. The median EQ-5D-5L VAS score was 75 (IQR: 70, 90). No floor effect for EQ-5D-5L VAS score was observed, while its ceiling effect was acceptable at 1.7%.

Construct validity
Table 2 shows that construct validity of the EQ-5D-5L instrument. With regard to structural validity, there were significant associations between the EQ-5D-5L index and VAS scores with BASDAI, BASFI, HAQ-DI, pain, PGA, SF-36 PCS and SF-36 MCS, WPAI:SpA absenteeism, presenteeism, work productivity loss, and activity impairment (p<0.001). The Spearman correlation coefficients ranged from moderate (HAQ-DI score with EQ-5D-5L VAS score: -0.39, p<0.001) to high (WPAI:SpA—activity impairment with EQ-5D-5L index score: -0.65, p<0.001). All 22 hypotheses were confirmed in the results.
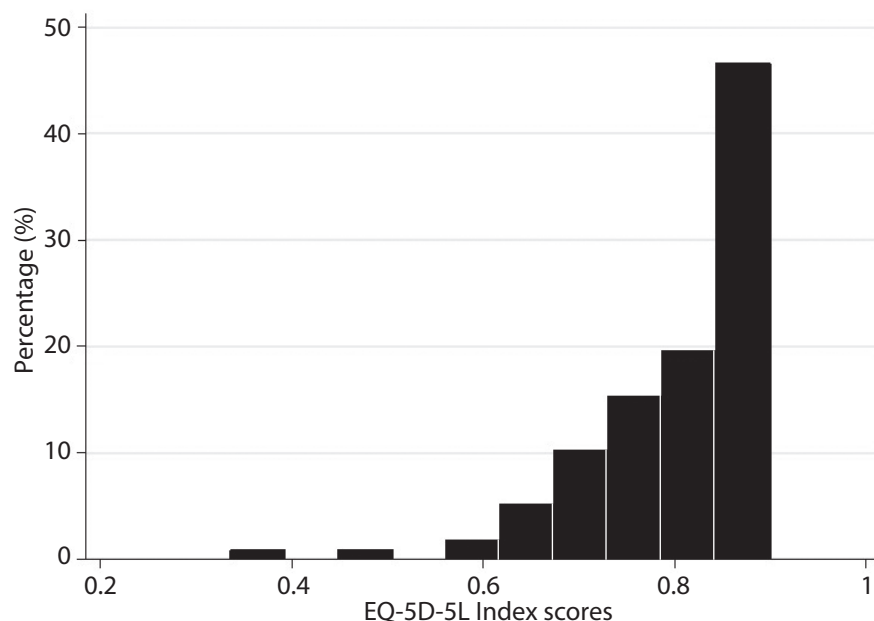

Figure 1. Histogram showing the distribution of EQ-5D-5L index scores of patients with axial spondyloarthritis at baseline (n=118).
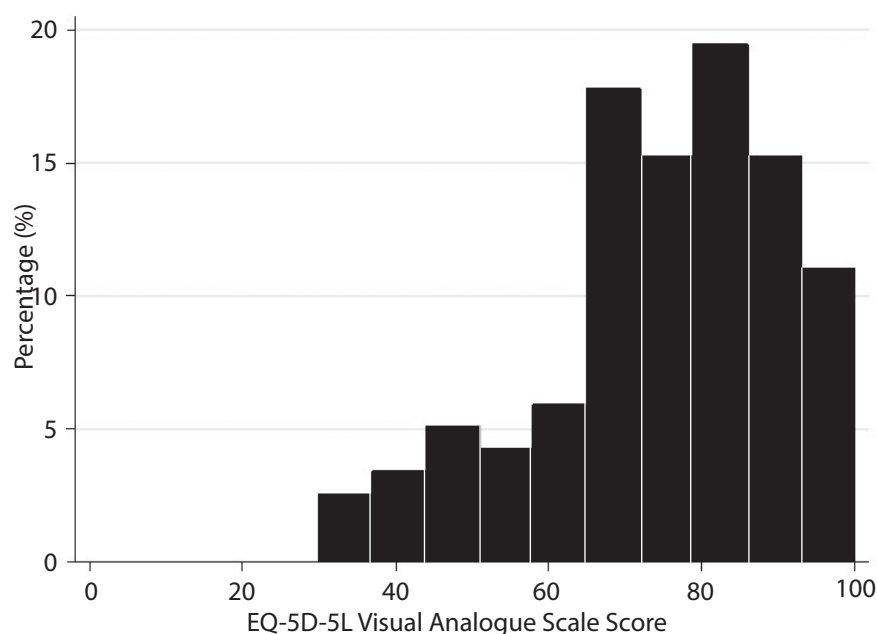

Figure 2. Histogram showing the distribution of EQ-5D-5L visual analog scale scores of patients with axial spondyloarthritis at baseline (n=118).

Eur J Rheumatol 2020; 7(2): 71-8

Seng et al. Validity of EQ-5D-5L in axial SpA patients

Table 2. Construct validity of EQ-5D-5L instrument.

	Hypothesis	EQ-5D-5L index score		EQ-5D-5L VAS score	
		Spearman Correlation	Hypothesis Met	Spearman Correlation	Hypothesis Met
SF-36 PCS	Moderate (+)	0.54*	Yes	0.41*	Yes
SF-36 MCS	Moderate (+)	0.41*	Yes	0.60*	Yes
PGA	High (-)	-0.55*	Yes	-0.52*	Yes
HAQ-DI	Moderate (-)	-0.60*	Yes	-0.39*	Yes
BASDAI	Moderate (-)	-0.54*	Yes	-0.47*	Yes
BASFI	Moderate (-)	-0.56*	Yes	-0.44*	Yes
Pain	Moderate (-)	-0.59*	Yes	-0.46*	Yes
WPAI:SpA - absenteeism	Moderate (-)	-0.42*	Yes	-0.42*	Yes
WPAI:SpA - presenteeism	Moderate (-)	-0.59*	Yes	-0.53*	Yes
WPAI:SpA - work productivity loss	Moderate (-)	-0.59*	Yes	-0.54*	Yes
WPAI:SpA - activity impairment	Moderate (-)	-0.65*	Yes	-0.55*	Yes

*p<0.001 as corrected using Bonferroni's correction as 22 a priori hypotheses were tested; (+) and (-) indicate the direction of correlations; 'Hypothesis Met' column indicates whether hypothesis generated prior to analysis about direction of correlation and magnitude was met in the specific variable.
BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; EQ-5D-5L: EuroQOL 5 Dimensions-5 Levels Questionnaire; HAQ-DI: Health Assessment Questionnaire Disability Index; PGA: patient global assessment of disease activity; SF-36 PCS: Short Form-36 Health Survey physical component summary score; SF-36 MCS: Short Form-36 Health Survey mental component summary score; WPAI:SpA: Work Productivity and Activity Impairment Questionnaire: Spondyloarthritis.

Reliability and measurement error

Cronbach's alpha was 0.79, which showed that EQ-5D-5L had excellent internal consistency.

The test-retest reliability of EQ-5D-5L was evaluated in 43 stable patients. Reliability for EQ-5D-5L mobility, self-care, usual activities, and anxiety/depression was good, with a weighted kappa of 0.78 (95% CI: 0.65-0.92), 0.68 (95% CI: 0.13-0.88.), 0.85(95% CI: 0.80-0.0.93) and 0.71 (95% CI: 0.61-0.81), respectively. Similarly, the reliability of the EQ-5D-5L index and VAS scores were excellent with ICCs of 0.92 (95% CI: 0.85-0.96) and 0.99 (95% CI: 0.98-1), respectively. However, the weighted kappa for EQ-5D-5L pain/discomfort was moderate [kappa=0.53 (95% CI=0.41-0.60)]. The MDC_{95} for the EQ-5D-5L index and VAS score were 0.06 and 4.5, respectively.

Discussion

The EQ-5D-5L is a commonly utilized HrQoL instrument studied in clinical research (18). To best of our knowledge, no study has examined the validity and reliability of EQ-5D-5L specifically among patients with axial SpA in Singapore. Our study findings support the utilization of EQ-5D-5L as a valid and reliable instrument in assessing HrQoL among patients with axial SpA.

Our results were consistent with findings from a recent study conducted in Hong Kong among both axial and peripheral patients with SpA, which found that the EQ-5D-5L showed acceptable psychometric properties for the evaluation of HrQoL of patients (8). In addition, the results from this study were comparable to other Asian studies such as Luo et al. (50), which evaluated the psychometric properties of EQ-5D-5L (Chinese language) among a heterogeneous population of rheumatology patients with conditions such as systemic lupus erythematosus, and Leung et al (9), which evaluated its use among psoriatic arthritis patients. It is also noteworthy that Leung et al. (9) found that EQ-5D-5L had a higher patient acceptability compared to other HrQoL instruments such as Short-Form 6D dimensions, while yielding comparable abilities to differentiate varying health states.

However, our results differed from studies conducted in Europe, where EQ-5D scores were found to have inconsistent correlation with different PROMs. In a study conducted among psoriatic patients in Hungary, a weak correlation was shown between EQ-5D scores and Psoriasis Severity Index score, an instrument used to assess disease severity (51). The differing findings could be due to the use of EQ-5D-3L in the study, which permits only three possible responses for the five health items. Studies that evalu-

ated the use of EQ-5D-5L and EQ-5D-3L found that EQ-5D-5L had superior sensitivity and precision for health status measurements (52). Additionally, EQ-5D-3L has been shown to overestimate health-related problems, which may lead to derivation of biased utilities (52).

Overall, the validity of the EQ-5D-5L was demonstrated through the construct validity, of which all 22 a priori hypotheses comparing EQ-5D-5L scores to other PROMs were fulfilled. This indicated that the hypothesized differences in health status quantified by other PROMs existed among patients grouped by their differing responses to EQ-5D-5L items. Our convergent validity analyses also showed that the items in EQ-5D-5L correlated well with all six PROMs evaluated in the study. These results may be explained by the usage of self-administered PROMs that have been shown to have better correlations with HrQoL PROMs, compared to clinician-reported outcome measures (53, 54). Although clinician-reported outcome measures were not evaluated in our study, future studies should consider assessing the construct validity of EQ-5D-5L with other clinician-reported outcome measures.

Our study also showed that WPAI:SpA and HAQ-DI were the two PROMs that had the

Seng et al. Validity of EQ-5D-5L in axial SpA patients

Eur J Rheumatol 2020; 7(2): 71-8

highest correlations with the EQ-5D-5L. One potential reason for this finding may be due to the overlapping constructs of function-based items evaluated in the EQ-5D-5L, WPAI:SpA, and HAQ-DI, especially pertaining to the impact of axial SpA on usual activities and work (17, 21, 25). This in turn reflects the importance of improving patients' physical function, so as to allow them to fulfill their work responsibilities and enjoy better HrQoL (55).

Additionally, the excellent test-retest reliability of the EQ-5D-5L was reflected by weighted kappa of ≥0.61 for four among five items on the EQ-5D-5L, and ICCs exceeding 0.9 for the EQ-5D-5L index and VAS scores. The moderate test-retest reliability of the EQ-5D-5L pain/discomfort (kappa=0.53) was similar to that noted in other studies (50, 56). A potential reason for lower reliability in this dimension could be attributed to the variability in frequency and intensity of pain among patients with axial SpA (50, 56). Consequently, this may have contributed to differences in pain assessment during the test-retest period. Pertaining to the MDC of the EQ-5D-5L index and VAS scores, the results were similar to values utilized in studies performed in ankylosing spondylosis (57) and rheumatoid arthritis patients (58), as well as general reference values described by Walter et al. (59).

This study has several limitations. First, this study only included patients with axial SpA, which limits its generalizability to other subtypes of SpA. Nonetheless, our study findings provide a basis for future research to assess the psychometric properties of EQ-5D-5L among patients with other subtypes of SpA. Second, the range for EQ-5D-5L index scores among patients in the study was relatively narrow and most patients scored between 0.6 and 0.9. Although this was similar to findings of Tsang et al. (8) (mean EQ-5D-5L index score=0.79±0.19), the measurement properties may be less generalizable to patients with EQ-5D-5L scores that are less than 0.6. Owing to the lack of the gold standard comparator for the EQ-5D-5L among patients with axial SpA, criterion validity was not assessed. In addition, as there were less than 5% of patients who were of Malay, Indian, and other ethnicity respectively, the cross-cultural validity of the EQ-5D-5L was not evaluated. Overall, all other measurement properties in the COSMIN checklist

were evaluated and this lends a reasonably comprehensive review on the validity and reliability of the EQ-5D-5L to be used in assessing the HrQoL of patients with axial SpA. Future research is required to achieve a better understanding on how EQ-5D-5L can be used in clinical practice to guide treatment decisions among patients with axial SpA.

Overall, the study findings support the validity and reliability of the EQ-5D-5L for the assessment of HrQoL among patients with axial SpA in Singapore.

References

1. Kotsis K, Voulgari PV, Drosos AA, Carvalho AF, Hyphantis T. Health-related quality of life in patients with ankylosing spondylitis: A comprehensive review. Expert Rev Pharmacoecon Outcomes Res 2014; 14: 857-72. [Crossref]

2. Reveille JD, Weisman MH. The epidemiology of back pain, axial spondyloarthritis and HLA-B27 in the United States. Am J Med Sci 2013; 345: 431-6. [Crossref]

3. Dean LE, Jones GT, MacDonald AG, Downham C, Sturrock RD, Macfarlane GJ. Global prevalence of ankylosing spondylitis. Rheumatology (Oxford) 2014; 53: 650-7. [Crossref]

4. Sieper J, Rudwaleit M, Khan MA, Braun J. Concepts and epidemiology of spondyloarthritis. Best Pract Res Clin Rheumatol 2006; 20: 401-17. [Crossref]

5. Braun J, Kiltz U, Sarholz M, Heldmann F, Regel A, Baraliakos X. Monitoring ankylosing

spondylitis: Clinically useful markers and prediction of clinical outcomes. Expert Rev Clin Immunol 2015; 11: 935-46. [Crossref]

6. Seng JJB, Kwan YH, Low LL, Thumboo J, Fong WSW. Role of neutrophil to lymphocyte ratio (NLR), platelet to lymphocyte ratio (PLR) and mean platelet volume (MPV) in assessing disease control in Asian patients with axial spondyloarthritis. Biomarkers 2018; 23: 335-8. [Crossref]

7. Bullinger M, Quitmann J. Quality of life as patient-reported outcomes: Principles of assessment. Dialogues Clin Neurosci 2014; 16: 137-45.

8. Tsang HHL, Cheung JPY, Wong CKH, Cheung PWH, Lau CS, Chung HY. Psychometric validation of the EuroQoL 5-dimension (EQ-5D) questionnaire in patients with spondyloarthritis. Arthritis Res Ther 2019; 21: 41. [Crossref]

9. Leung YY, Png ME, Wee HL, Thumboo J. Comparison of EuroQol-5D and short form-6D utility scores in multiethnic Asian patients with psoriatic arthritis: A cross-sectional study. J Rheumatol 2013; 40: 859-65. [Crossref]

10. Hermann J. [Spondyloarthritis and quality of life]. Z Rheumatol 2010; 69: 213-9. [Crossref]

11. Wallman JK, Kapetanovic MC, Petersson IF, Geborek P, Kristensen LE. Comparison of non-radiographic axial spondyloarthritis and ankylosing spondylitis patients - baseline characteristics, treatment adherence, and development of clinical variables during three years of anti-TNF therapy in clinical practice. Arthritis Res Ther 2015; 17: 378. [Crossref]

12. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol 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 2010; 19: 539-49. [Crossref]

13. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of assessment of SpondyloArthritis international society classification criteria for axial spondyloarthritis (part II): Validation and final selection. Ann Rheum Dis 2009; 68: 777-83. [Crossref]

14. van der Linden S, Akkoc N, Brown MA, Robinson PC, Khan MA. The ASAS criteria for axial Spondyloarthritis: Strengths, weaknesses, and proposals for a way forward. Curr Rheumatol Rep 2015; 17: 62. [Crossref]

15. Jacquemin C, Molto A, Servy H, Sellam J, Foltz V, Gandjbakhch F, et al. Flares assessed weekly in patients with rheumatoid arthritis or axial spondyloarthritis and relationship with physical activity measured using a connected activity tracker: A 3-month study. RMD Open 2017; 3: e000434. [Crossref]

16. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. Control Clin Trials 1991; 12: 142s-58s. [Crossref]

Eur J Rheumatol 2020; 7(2): 71-8

Seng et al. Validity of EQ-5D-5L in axial SpA patients

17. Brooks R. EuroQol: The current state of play. Health Policy 1996; 37: 53-72. [Crossref]

18. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; 35: 1095-108. [Crossref]

19. Thumboo J, Wu Y, Tai ES, Gandek B, Lee J, Ma S, et al. Reliability and validity of the English (Singapore) and Chinese (Singapore) versions of the Short-Form 36 version 2 in a multi-ethnic urban Asian population in Singapore. Qual Life Res 2013; 22: 2501-8. [Crossref]

20. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med 2001; 33: 350-7. [Crossref]

21. Bruce B, Fries JF. The Stanford health assessment questionnaire: Dimensions and practical applications. Health Qual Life Outcomes 2003; 1: 20. [Crossref]

22. Kwan YH, Fong W, Lui NL, Yong ST, Cheung YB, Malhotra R, et al. Validity and reliability of the Health Assessment Questionnaire among patients with spondyloarthritis in Singapore. Int J Rheum Dis 2018; 21: 699-704. [Crossref]

23. Quinzanos I, Luong PT, Bobba S, Steuart Richards J, Majithia V, Davis LA, et al. Validation of disease activity and functional status questionnaires in spondyloarthritis. Clin Exp Rheumatol 2015; 33: 146-52.

24. Wang CTM, Fong W, Kwan YH, Phang JK, Lui NL, Leung YY, et al. A cross-sectional study on factors associated with patient-physician discordance in global assessment of patients with axial spondyloarthritis: An Asian perspective. Int J Rheum Dis 2018; 21: 1436-42. [Crossref]

25. Reilly MC, Gooch KL, Wong RL, Kupper H, van der Heijde D. Validity, reliability and responsiveness of the Work Productivity and Activity Impairment questionnaire in ankylosing spondylitis. Rheumatology (Oxford) 2010; 49: 812-9. [Crossref]

26. Luo N, Wang Y, How CH, Tay EG, Thumboo J, Herdman M. Interpretation and use of the 5-level EQ-5D response labels varied with survey language among Asians in Singapore. J Clin Epidemiol 2015; 68: 1195-204. [Crossref]

27. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: Are available health status surveys adequate?. Qual Life Res 1995; 4: 293-307. [Crossref]

28. EMGO+. Questionnaires: Selecting, translating and validating. Institute for health and Care Research. 2010.

29. COSMIN. COSMIN methodology for assessing the content validity of PROMs. 2018.

30. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: Theory and application. Am J Med 2006; 119: 166.e7-16. [Crossref]

31. Grimm LG, Yarnold PR. Reading and Understanding More Multivariate Statistics. Washington DC: American Psychological Association; 2000.

32. Kiltz U, van der Heijde D, Boonen A, Akkoc N, Bautista-Molano W, Burgos-Vargas R, et al. Measurement properties of the ASAS Health Index: Results of a global study in patients with axial and peripheral spondyloarthritis. Ann Rheum Dis 2018; 77: 1311-7. [Crossref]

33. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018; 27: 1147-57. [Crossref]

34. Brazier J, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. Qual Life Res 1993; 2: 169-80. [Crossref]

35. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: Validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997; 36: 551-9. [Crossref]

36. Shikiar R, Willian MK, Okun MM, Thompson CS, Revicki DA. The validity and responsiveness of three quality of life measures in the assessment of psoriasis patients: Results of a phase II study. Health Qual Life Outcomes 2006; 4: 71. [Crossref]

37. Wailoo A, Hernandez M, Philips C, Brophy S, Siebert S. Modeling health state utility values in ankylosing spondylitis: Comparisons of direct and indirect methods. Value Health 2015; 18: 425-31. [Crossref]

38. Adams R, Walsh C, Veale D, Bresnihan B, FitzGerald O, Barry M. Understanding the relationship between the EQ-5D, SF-6D, HAQ and disease activity in inflammatory arthritis. Pharmacoeconomics 2010; 28: 477-87. [Crossref]

39. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. Rheumatology (Oxford). 2013; 52: 944-50. [Crossref]

40. de Hooge M, Ramonda R, Lorenzin M, Frallonardo P, Punzi L, Ortolan A, et al. Work productivity is associated with disease activity and functional ability in Italian patients with early axial spondyloarthritis: An observational study from the SPACE cohort. Arthritis Res Ther 2016; 18: 265. [Crossref]

41. King Jr JT, Tsevat J, Roberts MS. Measuring preference-based quality of life using the EuroQol EQ-5D in patients with cerebral aneurysms. Neurosurgery 2009; 65: 565-72; discussion 72-3. [Crossref]

42. Savoia E, Fantini MP, Pandolfi PP, Dallolio L, Collina N. Assessing the construct validity of the Italian version of the EQ-5D: Preliminary results from a cross-sectional study in North Italy. Health Qual Life Outcomes 2006; 4: 47. [Crossref]

43. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes 2007; 5: 70. [Crossref]

44. Bland JM, Altman DG. Cronbach's alpha. BMJ. 1997; 314: 572. [Crossref]

45. Morera OF, Stokes SM. Coefficient alpha as a measure of test score reliability: Review of 3 popular misconceptions. Am J Public Health 2016; 106: 458-61. [Crossref]

46. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979; 86: 420-8. [Crossref]

47. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-74. [Crossref]

48. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. BMC Med Res Methodol 2010; 10: 22. [Crossref]

49. King MT. A point of minimal important difference (MID): A critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res 2011; 11: 171-84. [Crossref]

50. Luo N, Chew LH, Fong KY, Koh DR, Ng SC, Yoon KH, et al. Validity and reliability of the EQ-5D self-report questionnaire in English-speaking Asian patients with rheumatic diseases in Singapore. Qual Life Res 2003; 12: 87-92. [Crossref]

51. Brodszky V, Péntek M, Bálint PV, Géher P, Hajdu O, Hodinka L, et al. Comparison of the Psoriatic Arthritis Quality of Life (PsAQoL) questionnaire, the functional status (HAQ) and utility (EQ-5D) measures in psoriatic arthritis: Results from a cross-sectional survey. Scand J Rheumatol 2010; 39: 303-9. [Crossref]

52. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. Pharmacoeconomics 2018; 36: 675-97. [Crossref]

53. El Miedany Y. Adopting patient-centered care in standard practice: PROMs moving toward disease-specific era. Clin Exp Rheumatol 2014; 32: S-40-6. [Crossref]

54. Lubelski D, Alvin MD, Nesterenko S, Sundar SJ, Thompson NR, Benzel EC, et al. Correlation of quality of life and functional outcome measures for cervical spondylotic myelopathy. J Neurosurg Spine 2016; 24: 483-9. [Crossref]

55. Huang J-C, Qian B-P, Qiu Y, Wang B, Yu Y, Zhu Z-Z, et al. Quality of life and correlation with clinical and radiographic variables in patients with ankylosing spondylitis: A retrospective case series study. BMC Musculoskelet Disord 2017; 18: 352. [Crossref]

56. Conner-Spady BL, Marshall DA, Bohm E, Dunbar MJ, Loucks L, Al Khudairy A, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. Qual Life Res 2015; 24: 1775-84. [Crossref]

57. Braun J, McHugh N, Singh A, Wajdula JS, Sato R. Improvement in patient-reported out-

Seng et al. Validity of EQ-5D-5L in axial SpA patients

Eur J Rheumatol 2020; 7(2): 71-8

comes for patients with ankylosing spondylitis treated with etanercept 50 mg once-weekly and 25 mg twice-weekly. Rheumatology (Oxford) 2007; 46: 999-1004. [Crossref]

58. Mian AN, Ibrahim F, Scott DL, Galloway J. Optimal responses in disease activity scores to treatment in rheumatoid arthritis: Is a DAS28 reduction of >1.2 sufficient? Arthritis Res Ther 2016; 18: 142. [Crossref]

59. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 2005; 14: 1523-32. [Crossref]