

# Psychometric Properties of the EuroQol-5D and Short Form-6D in Patients with Systemic Lupus Erythematosus

ROHIT AGGARWAL, CAITLYN T. WILKE, A. SIMON PICKARD, VIKRANT VATS, RACHEL MIKOLAITIS, LEWIS FOGG, JOEL A. BLOCK, and MEENAKSHI JOLLY

**ABSTRACT.** *Objective.* Health related quality of life (HRQOL) is an important patient-reported outcome in systemic lupus erythematosus (SLE). We evaluated the psychometric properties of 2 widely used preference-based generic HRQOL measures, EuroQol-5D (EQ-5D) and Short Form-6D (SF-6D), among United States patients with SLE.

*Methods.* Patients with SLE enrolled at an academic institution were assessed for self-reported generic HRQOL (EQ-5D, Medical Outcomes Study SF-36), disease activity, and disease damage SF-6D. Physical Component Score (PCS) and Mental Component Score (MCS) were calculated from SF-36. Criterion validity, convergent validity, and known-groups comparisons were evaluated for EQ-5D and SF-6D. Sensitivity to change (t tests, effect size) was evaluated in a subset of the cohort followed longitudinally.

*Results.* One hundred sixty-seven patients with SLE were enrolled. Related domains on the EQ-5D and SF-36 correlated strongly, e.g., mobility and physical functioning ( $r = 0.60$ ), whereas unrelated domains showed weak to moderate correlation. EQ-5D index, EQ-5D visual analog scale, and SF-6D score correlated strongly among each other as well as with most domains of SF-36. Both EQ-5D and SF-6D indices differentiated among patients of varied disease severity. EQ-5D and SF-6D were found to be sensitive to self-reported change in health but insensitive to change in disease activity longitudinally. Disease activity and damage showed weak correlation with HRQOL measures.

*Conclusion.* The SF-6D and EQ-5D exhibited satisfactory psychometric properties for use among US patients with SLE. Measures of disease activity and damage were weakly correlated with HRQOL, suggesting that HRQOL is an important complementary source of information about patients with SLE. (J Rheumatol First Release April 15 2009; doi:10.3899/jrheum.081022)

*Key Indexing Terms:*

EUROQoL-5D                      SHORT FORM-6D                      HEALTH RELATED QUALITY OF LIFE  
SYSTEMIC LUPUS ERYTHEMATOSUS

Systemic lupus erythematosus (SLE) is a chronic, disabling disease affecting physical, mental, and social aspects of life<sup>1</sup>. With the advent of new treatments and better understanding of the disease, survival has improved in recent decades<sup>2</sup>; however, this has not translated into improvement in health related quality of life (HRQOL)<sup>3,4</sup>.

Indices such as the SLE Disease Activity Index (Selena-SLEDAI)<sup>5</sup> and the Systemic Lupus International Collaborating Clinics (SLICC) Damage Index (SDI)<sup>6</sup> are used to measure disease activity and damage, respectively. These instruments focus only on the physical and physiological effect of disease; they do not assess effect on other well-being domains pertinent to patients with SLE. Also, these indices do not reflect patients' perception of their health in day-to-day life, thus resulting in wide discrepancies observed between physicians' and patients' perceptions of SLE activity and global health<sup>7-9</sup>. For these reasons, HRQOL has assumed increased importance as an outcome measure in clinical research in SLE<sup>4,10</sup>. Hence, both patient-reported outcomes such as HRQOL and physician-assessed outcomes such as disease activity or damage scores provide important yet distinct information, and both should be used to maximize our understanding of the physical, mental, and social health of patients with SLE<sup>4,10</sup>.

Although disease-specific HRQOL instruments such as the LupusQoL<sup>11,12</sup> provide detailed information regarding

---

*From the Section of Rheumatology, Rush University Medical Center; Center for Pharmacoeconomics and Outcomes Research, University of Illinois at Chicago; and College of Nursing, Rush University, Chicago, Illinois, USA.*

*Supported in part by the Rush University Committee on Research.*

*R. Aggarwal, MD, Section of Rheumatology, Rush University Medical Center; C.T. Wilke, MS; A.S. Pickard, PhD; V. Vats, PhD, Center for Pharmacoeconomics and Outcomes Research, University of Illinois at Chicago; R. Mikolaitis, BSc, Section of Rheumatology, Rush University Medical Center; L. Fogg, PhD, College of Nursing, Rush University; J.A. Block, MD; M. Jolly, MD, Section of Rheumatology, Rush University Medical Center.*

*Address reprint requests to Dr. R. Aggarwal, Section of Rheumatology, Rush University Medical Center, 1725 W. Harrison Street, Room 1017, Chicago, IL 60612. E-mail: rohit\_aggarwal@rush.edu*

*Accepted for publication January 23, 2009.*

aspects of HRQOL shown to be important to patients with SLE, generic instruments are often useful to facilitate comparisons with other chronic diseases. Generic HRQOL measures, such as the Medical Outcomes Study Short-Form 36 (SF-36) are widely used to assess functioning and well-being in patients with SLE<sup>10,13</sup>. The SF-36 provides a profile of the effects of disease on various aspects of HRQOL and generates 2 summary score of overall HRQOL. Such a feature characterizes indirect preference-based measures, where an overall index-based score is generated based on societal preference weights for responses to a health state classifier. A single summary (index) score thus generated is typically anchored at 1 for perfect health and 0 for death, with the possibility of negative scores assigned to health states considered worse than death. The most widely used generic preference-based measures include the EuroQoL instrument (EQ-5D)<sup>14</sup> or the use of selected items of the SF-36 to form the Short Form-6D (SF-6D)<sup>15</sup>. The EQ-5D is a widely used generic HRQOL measure with an index-based score that was originally developed for application to economic evaluations of healthcare interventions, providing the utilities necessary to generate quality-adjusted life-years used in the denominator of comparative effectiveness analyses<sup>14,16-18</sup>. The SF-6D was developed subsequent to the introduction of the SF-36 and provides a method to transform responses from a subset of items from 6 domains of the SF-36<sup>15</sup>. Both tools provide preference-based assessments of HRQOL in the form of single numeric values, which may be useful to follow outcomes in patients with SLE as well as to compare the economic effects of different treatments<sup>14,15</sup>.

The validation studies of both the SF-6D and the EQ-5D have not been performed on ethnically heterogeneous patients such as the SLE population in the United States<sup>18</sup>, and, to our knowledge, neither the SF-6D nor the EQ-5D has been systematically evaluated for use in patients with SLE in the US<sup>16,18,19</sup>, although 1 study evaluated the SF-6D for disease damage and mortality endpoints in American patients with SLE<sup>19</sup>. The primary aim of our study was to determine the psychometric properties of the EQ-5D and SF-6D among US patients with SLE. A secondary objective was to compare EQ-5D and SF-6D to the SF-36 Physical and Mental Component Scores (PCS, MCS) as HRQOL in patients with SLE. We also examined whether SLEDAI and SDI are useful clinical anchors to evaluate HRQOL measures. We proposed the following hypotheses based on previous studies on EQ-5D and SF-6D in other diseases. First, the related domains between SF-36 and EQ-5D will be strongly correlated and nonrelated domains will be weakly correlated. Second, EQ-5D and SF-6D scores will show strong correlation with each other and with the subscales of SF-36. Third, EQ-5D and SF-6D will detect meaningful changes (e.g., effect size > 0.50) in disease activity, damage, or health status. And fourth, we expect a poor correlation between disease activity and damage, and HRQOL meas-

ures, indicating a lack of a good clinical anchor for responsiveness assessment<sup>20</sup>.

## MATERIALS AND METHODS

**Subjects.** Our study was part of an ongoing longitudinal evaluation of a newly developed disease-specific HRQOL measure in SLE at Rush University Medical Center. After approval by the Institutional Review Board, consecutive, consenting patients fulfilling American College of Rheumatology (ACR) criteria for SLE<sup>21</sup> and receiving care at Rush during 2006 and 2007 were eligible for enrollment. Demographic, clinical, and laboratory information pertaining to their disease was obtained by retrospective chart review. Disease activity (Selena-SLEDAI), disease damage (SDI), and the SF-36 and EQ-5D were assessed prospectively. Available longitudinal data on 66 subjects with SLE were used to determine the responsiveness of the EQ-5D and SF-6D, described below.

**Measures.** Disease status measures: The SDI is a validated physician-rated index that assesses cumulative organ damage due to the disease or therapy<sup>6</sup>. The total SDI score ranges from 0 (no damage) to 47 (maximum damage)<sup>6</sup>. The Selena-SLEDAI uses a validated weighting system to evaluate SLE disease activity in 9 organ systems<sup>5</sup>. The total SLEDAI score can range from 0 (no activity) to 105 (maximum activity)<sup>5</sup>.

HRQOL measures: A standard version of the EQ-5D self-report instrument was completed by patients<sup>14</sup>. The EQ-5D health state classifier consists of 5 single-item dimensions — mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD) — with 3 levels of response for no, some, or extreme problems in each dimension. In addition to the health state classifier, patients rated their current health on a 20-cm visual analog scale (EQ-5D VAS) ranging from 0 (worst possible health state) to 100 (best possible health state)<sup>14</sup>. The US preference-based algorithm<sup>17</sup> was used to convert patient responses to the health state classifier into a single index, which produces scores anchored at 0 for death and ranges from -0.11 to 1. For both the EQ-5D index and the VAS, a higher score denotes better health.

A standard 4-week recall version of the SF-36 was used<sup>13</sup>. The SF-36 has several components including physical functioning (PF), role limitations physical (RP), bodily pain (BP), mental health (MH), role limitations emotional (RE), social functioning (SF), vitality (V), and general health perceptions (GH). In addition, the 2 summary scores, physical component scores (PCS), and mental component scores (MCS), were assessed. The SF-6D was calculated from SF-36 based on the standard algorithm to create a weighted index score ranging from 1.0 [no difficulty in any dimensions (or perfect health)] to 0.296 (severely impaired levels in all dimensions), with death anchored at 0<sup>15</sup>. The SF-6D is composed of 6 dimensions, which can be used as a stand-alone instrument or can be derived using a transformation of items on the SF-36: PF (3 SF-36 items), role limitation [(RL) 2 items], SF (1 item), BP (2 items), MH (1 item), and V (1 item); the GH items are not included and 2 scales measuring role limitations (RP and RE) are collapsed into single dimension — RL<sup>15</sup>.

**Evaluation of psychometric properties and statistics.** The EQ-5D and SF-6D were tested for validity, reliability, and responsiveness<sup>22</sup>. After calculation of summary scores, the SLEDAI, SDI, SF-6D, EQ-5D index, and EQ-5D VAS were assessed for normality using Q-Q plots. Nonparametric statistics were used when the distribution of the data was non-normal, and  $p < 0.05$  was chosen as the level of statistical significance for all tests. Correlations were classified as very strong ( $r > 0.8$ ), strong ( $0.6 \leq r < 0.8$ ), moderate ( $0.4 \leq r < 0.6$ ), weak ( $0.2 \leq r < 0.4$ ), or absent ( $0 \leq r < 0.2$ )<sup>23</sup>.

**Criterion validity:** Criterion validity involves the assessment of an instrument against an external reference of the true value, or against some other standard that is accepted as providing an indication of the true values for the measurements. As there is no gold standard measure of HRQOL, criterion validity was evaluated by examining the strength of relationship between (1) the EQ-5D index score and SF-6D score and (2) the EQ-5D-VAS, SF-36, SLEDAI, and SDI using Spearman's ( $r$ ) or Pearson's correlation coefficient, as appropriate to the scale.

Convergent and divergent validity: To measure convergent and divergent validity (concurrent validity), a corresponding new instrument is usually compared with an established questionnaire. To establish the convergent and divergent validity of the EQ-5D health state classifier, corresponding dimensions of the EQ-5D and domains of the SF-36 were compared using Spearman's rho correlation ( $\rho$ ). Since the SF-6D is based on the same measure as the SF-36 and would naturally be expected to correlate with the domains of that measure, tests of convergent and divergent validity were performed only for the EQ-5D.

Discriminant (known-groups) construct validity: Known-groups construct validity evaluates whether a measure is able to identify differences between patient subgroups stratified based on an external anchor of health. In our study, individuals were dichotomized into higher versus lower disease activity and damage subgroups, based on the median SLEDAI and SDI score. Patients above the median score of SLEDAI and SDI were defined as having higher disease activity and higher disease damage, respectively, whereas patients whose scores were below the median were considered to have lower disease activity and damage. The subgroups thus generated were consistent with the general clinical consensus concerning higher and lower disease activity and damage in SLE<sup>20</sup>. Student's *t* tests were used to compare HRQOL indices associated with lower and higher disease activity/damage subgroups for the EQ-5D and SF-6D.

Responsiveness: The proportion of subjects with floor and ceiling effects (percentages of respondents scoring at the lowest and highest possible scale level) were calculated. The sensitivity of the EQ-5D and SF-6D was evaluated for patients who were measured during 2 separate visits and additionally had improved or declined in health based on disease activity or self-reported health (by SF-36 item 2 and EQ-5D VAS). Patients with a SLEDAI increase of more than 3 points between the visits were considered to have had health decline (greater disease activity), whereas a SLEDAI decrease of more than 3 points was considered a meaningful improvement; the threshold was set at more than 3 to conform with the Selena-SLEDAI definition of mild to moderate flare<sup>5</sup>. A response of "much better or somewhat better" to SF-36 item 2 was considered improvement in quality of health, and responses of "somewhat worse" or "much worse" were characterized as a decline. In addition, we used EQ-5D VAS as patient global assessment of disease and the percentage improvement or deterioration over baseline was calculated using the formula:  $[(\text{visit 2 EQ-5D VAS} - \text{visit 1 EQ-5D VAS}) / \text{visit 1 EQ-5D VAS}] \times 100$ . Patients were classified as "better" if the patient global had increased by  $\geq 20\%$ , and "worse" if the patient global had decreased by  $\geq 20\%$ , which is consistent with guidelines of response from ACR20 criteria<sup>24</sup>. Paired *t* tests were used to identify the degree of change in EQ-5D and SF-6D indices for improved and declined subgroups in SLEDAI and self-reported health external anchors (SF-36 item 2 and EQ-5D VAS). Effect sizes, calculated as mean change score divided by the standard deviation of 2 groups, were additionally calculated as a measure of responsiveness. Cohen's guidelines for effect size interpretation were used: "small effect size,  $d \geq 0.2 < 0.5$ ," "medium effect size,  $d \geq 0.5 < 0.8$ ," and "large effect size,  $d \geq 0.8$ "<sup>23</sup>.

## RESULTS

Demographic characteristics of study subjects are summarized in Table 1. Of the 167 patients with SLE recruited into our study, 93.5% were women, 56.3% were African American, and the mean age at study entry was 42.5 years [standard deviation (SD) 13.0; Table 1].

The mean ( $\pm$  SD) [median (interquartile range, IQR)] of SLEDAI and SDI scores were 6.2 ( $\pm$  5.7) [5.0 (IQR 2–10)] and 2.0 ( $\pm$  2.0) [2.0 (IQR 1–3)], respectively (Table 1). Ninety-four percent (157 of 167) of subjects fully completed the EQ-5D questionnaire, whereas only 76% (127 of 167) of subjects fully completed the SF-36. Eighty-nine percent

Table 1. Demographics and clinical characteristics.

Clinical Features/Indices	Mean ( $\pm$ SD), or n (%)
Age at presentation, yrs	42.5 $\pm$ 13.0
Age at diagnosis, yrs	33.2 $\pm$ 12.9
Duration of illness, yrs	9.3 $\pm$ 8.8
Race (%)	
African American	94/167 (53.5)
Caucasian	42/167 (30.2)
Hispanic	21/167 (12.5)
Asian	10/167 (6.0)
Sex (%)	
Female	156/167 (93.5)
Male	11/167 (6.5)
Disease activity and damage	
SLEDAI, mean	6.2 $\pm$ 5.7
SDI, mean	2.0 $\pm$ 2.0
SF-36 subscales (0–100)	
Physical function	55.4 $\pm$ 30.4
Role-physical	38.2 $\pm$ 41.3
Bodily pain	51.5 $\pm$ 25.9
General health	43.1 $\pm$ 20.6
Vitality	45.1 $\pm$ 22.7
Social functioning	59.6 $\pm$ 27.1
Role emotional	57.7 $\pm$ 45.45
Mental health	68.1 $\pm$ 21.21
MCS	51.9 $\pm$ 8.9
PCS	35.6 $\pm$ 10.8
EQ-5D index (0–1)	0.72 $\pm$ 0.19
EQ-5D VAS (0–100)	68.7 $\pm$ 20.46
SF-6D (0.296–1)	0.64 $\pm$ 0.14

SLEDAI: Systemic Lupus Erythematosus Disease Activity Index; SDI: Systemic Lupus International Collaborating Clinics (SLICC) Damage Index; SF-36: Medical Outcomes Study Short-Form 36; MCS/PCS: Mental/Physical Component Scores; EQ-5D: EuroQol; SF-6D: Short Form-6D; VAS: Visual analog scale.

(149 of 167) of subjects completed all necessary questions required to calculate SF-6D scores. The mean ( $\pm$  SD) SF-6D and EQ-5D scores were 0.64 ( $\pm$  0.14) and 0.72 ( $\pm$  0.19), respectively (Table 1). The most frequent problem reported for the EQ-5D was pain/discomfort (77.6%), followed by inability to engage in usual activities (60.0%); the least reported problems were with self-care (23.7%; Figure 1).

**Criterion validity.** The EQ-5D index appeared to be most strongly correlated with the SF-6D ( $r = 0.62$ ). Moderate correlations of EQ-5D with PCS ( $r = 0.52$ ) and VAS ( $r = 0.50$ ) were also observed. However, the correlation with MCS ( $r = 0.3$ ) was weak (Table 2). The EQ-5D index correlated moderately to strongly with most domains of the SF-36: PH ( $r = 0.54$ ), RP ( $r = 0.43$ ), BP ( $r = 0.62$ ), GH ( $r = 0.42$ ), V ( $r = 0.46$ ), SF ( $r = 0.55$ ), RE ( $r = 0.5$ ), and MH ( $r = 0.56$ ). The SF-6D similarly showed evidence of criterion validity, with strong correlations with PCS ( $r = 0.72$ ) and EQ-5D VAS ( $r = 0.62$ ), but only a weak correlation with the MCS ( $r = 0.30$ ). SF-6D index strongly correlated with all domains of the SF-36 (Table 2). Correlations of EQ-5D and SF-6D with SLEDAI and SDI were weak: EQ-5D and SLEDAI ( $r =$

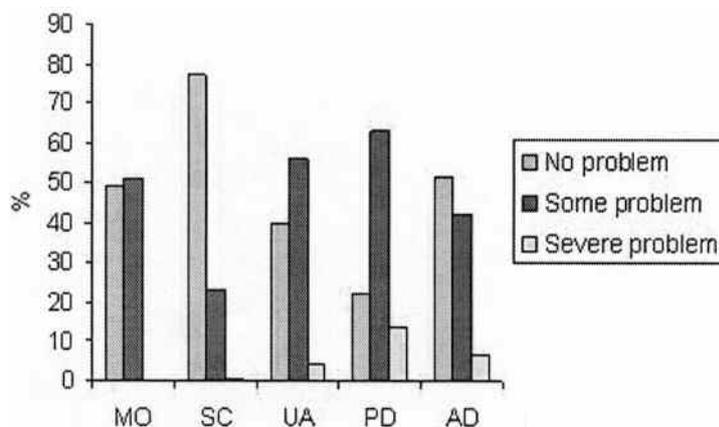


Figure 1. Percentage of subjects reporting problems, by EQ-5D dimension. MO: mobility; SC: self-care; UA: usual activity; PD: pain and disability; AD: anxiety and depression.

-0.21), EQ-5D and SDI ( $r = -0.20$ ), SF-6D and SLEDAI ( $r = -0.23$ ), and SF-6D and SDI ( $r = -0.22$ ). SF-36 scores (PCS, MCS) showed very weak correlation with SLEDAI and SDI (Table 2). EQ-5D dimensions also showed no correlation with SLEDAI (MO,  $r = 0.01$ ; SC,  $r = 0.15$ ; UA,  $r = 0.09$ ; PD,  $r = 0.07$ ; AD,  $r = 0.16$ ) and weak to no correlation with SDI (MO,  $r = 0.21$ ; SC,  $r = 0.12$ ; UA,  $r = 0.09$ ; PD,  $r = 0.06$ ; AD,  $r = 0.18$ ).

**Convergent/divergent validity.** As evident from Table 3, corresponding domains of the SF-36 and the EQ-5D were highly correlated. Strong correlations were observed between related constructs of the EQ-5D and SF-36: MO and PF; UA

and PF, RP, and SF; PD and BP; and AD and MH. The strong correlations between these domains support convergent validity of EQ-5D and SF-36 among subjects with SLE. The weak to moderate correlation between noncorresponding domains of EQ-5D and SF-36 supports divergent validity among subjects with SLE (Table 3).

**Discriminant validity.** The median (IQR) score of SLEDAI and SDI were 5 (2–10) and 2 (1–3), respectively. Patients were defined as having higher disease activity if SLEDAI was  $> 5$  and lower disease activity if SLEDAI  $\leq 5$ , and higher disease damage if SDI  $> 2$  and lower disease damage if  $\leq 2$ . Both EQ-5D and SF-6D detected higher HRQOL in the

Table 2. Spearman correlations ( $r$ ) for the EQ-5D and SF-6D with SF-36 domains, summary scores, SLE disease activity (SLEDAI), and damage scores (SDI).

	SF-6D Score	SF-36 PCS	SF-36 MCS	SF-36 PF	SF-36 RP	SF-36 BP	SF-36 GH	SF-36 V	SF-36 SF	SF-36 RE	SF-36 MH	SLEDAI	SDI
EQ-5D index	0.62	0.52	0.30	0.54	0.43	0.62	0.42	0.46	0.55	0.5	0.56	-0.22	0.20
SF-6D	NA	0.72	0.30	0.72	0.72	0.69	0.57	0.73	0.76	0.61	0.66	-0.23	-0.22
EQ-5D VAS	0.62	0.52	0.22	0.52	0.53	0.47	0.49	0.55	0.54	0.33	0.53	-0.21	-0.20
SLEDAI	-0.23	-0.13	-0.10	-0.18	-0.15	-0.22	-0.22	-0.16	-0.13	-0.18	-0.19		
SDI	-0.22	-0.06	-0.09	-0.21	-0.09	-0.08	-0.1	-0.09	-0.13	-0.16	-0.21		

PF: physical functioning; RP: role limitations physical; BP: bodily pain; GH: general health perceptions; V: vitality; SF: social functioning; RE: role limitations emotional; MH: mental health; NA: not applicable. For other abbreviations, see Table 1.

Table 3. Spearman correlations ( $r$ ) between EQ-5D and SF-36 domains\*.

EQ-5D	SF-36 PF	SF-36 RP	SF-36 BP	SF-36 GH	SF-36 V	SF-36 SF	SF-36 RE	SF-36 MH
Mobility	-0.62	-0.42	-0.46	-0.31	-0.36	-0.46	-0.37	-0.42
Self-care	-0.39	-0.34	-0.41	-0.32	-0.29	-0.37	-0.31	-0.24
Usual activities	-0.55	-0.58	-0.49	-0.43	-0.44	-0.58	-0.42	-0.38
Pain/discomfort	-0.39	-0.30	-0.66	-0.27	-0.31	-0.42	-0.39	-0.39
Anxiety/depression	-0.38	-0.29	-0.29	-0.45	-0.48	-0.43	-0.47	-0.69

\* All  $p$  value  $< 0.01$ . For abbreviations, see Table 2.

lower disease activity subgroup and both were able to significantly discriminate between with lower and higher disease activity subgroups (Table 4). With regard to disease damage, however, the EQ-5D was not able to significantly differentiate between patients with lower and higher damage groups, but the SF-6D was able to detect significantly higher HRQOL in patients with lower disease damage. In addition, both EQ-5D and SF-6D indices were more discriminative between disease activity and damage levels than the commonly used SF-36 PCS and MCS summary scores (Table 4). There were no differences in terms of domains of SF-36 in patients with higher and lower disease activity.

**Responsiveness.** Of the subjects completing the EQ-5D and SF-6D, 12.7% (20/157) and 2.6% (4/149) were at the ceiling of the EQ-5D and SF-6D, respectively. No floor effects were seen for EQ-5D score, while only 0.67% (1/167) was at the floor of the SF-6D. Of the 66 patients measured at multiple timepoints, 15 had declined (SLEDAI increase > 3) and 25 had improved based on SLEDAI scores (SLE decrease > 3; Table 5); similarly, 15 declined and 28 improved based on the self-reported change item. Ten patients declined in SLEDAI out of 15 patients who reported decline in health status. Similarly, 14 patients improved their SLEDAI out of 28 patients who reported improvement in their health status. Overall percentage agreement was 65%. The sensitivity of the EQ-5D and SF-6D to improvement or decline in SLEDAI scores was poor, with small to absent effect size (Table 5), which is consistent with the

weak correlation between SLEDAI and all HRQOL measures (EQ-5D, SF-6D, and SF-36). The EQ-5D measured significant improvement in HRQOL for patients self-reporting somewhat or much better health (mean EQ-5D change from 0.70 to 0.77) with small to medium effect sizes. The SF-6D was sensitive to self-reported improvement in health (SF-6D increase from 0.63 to 0.68) with small to medium effect sizes. Neither the EQ-5D nor the SF-6D was sensitive to self-reported decline in health, showing small to absent effects. The SF-6D was sensitive to EQ-5D VAS improvement, but not to decline, because of the small number of patients who declined. Similarly, EQ-5D showed a trend towards improvement, with increased EQ-5D VAS score during 2 visits, but not with decline, for the same reason. Both SF-6D and EQ-5D showed small to medium effect sizes for changes in EQ-5D VAS (Table 5). The magnitude of effect sizes for changes in VAS approximated a clinically meaningful difference, i.e., effect size of 0.5.

## DISCUSSION

Our study assessed the psychometric properties of EQ-5D and SF-6D in multiethnic US subjects with SLE. To our knowledge, it represents the first attempt to evaluate the validity of the EQ-5D in patients with SLE in the US, and provides complementary information to the only previous study of SF-6D validity in the US<sup>19</sup>.

The demographic distribution of the sample cohort closely resembles the American SLE population as a whole<sup>25</sup>.

Table 4. EQ-5D, SF-6D, SF-36 domains, and PCS and MCS discrimination between patients with lower and higher disease activity and damage subgroup.

	HRQOL Score		p	HRQOL Score		p
	SLEDAI ≤ 5, n = 83	SLEDAI > 5, n = 84		SDI ≤ 2, n = 101	SDI > 2, n = 54	
EQ-5D	0.75 (± 0.18)	0.69 (± 0.19)	0.043	0.74 (± 0.18)	0.68 (± 2.1)	0.071
SF-6D	0.66 (± 0.14)	0.61 (± 0.13)	0.025	0.65 (± 0.14)	0.6 (± 0.13)	0.043
PCS	37.03 (± 11.4)	33.99 (± 10.0)	0.115	36.17 (± 11.4)	34.86 (± 9.0)	0.532
MCS	51.75 (± 8.8)	52.13 (± 9.0)	0.832	52.13 (± 9.1)	51.19 (± 9.8)	0.573
PF	55 (IQR 32.5–90)	50 (IQR 30–75)	NS	55 (IQR 35–90)	50 (IQR 30–70)	NS
RP	25 (IQR 0–100)	25 (IQR 0–75)	NS	25 (IQR 0–100)	25 (IQR 0–50)	NS
BP	51 (IQR 41–74)	41 (IQR 22–62)	NS	51 (IQR 31–73)	42 (IQR 31–68)	NS
GH	42 (IQR 27–57)	42 (IQR 22–50)	NS	42 (IQR 27–57)	37 (IQR 25–57)	NS
V	30 (IQR 15–50)	45 (IQR 22.5–55)	NS	45 (IQR 30–60)	50 (IQR 20–60)	NS
SF	62.5 (IQR 50–87.5)	62.5 (IQR 37.5–75)	NS	56.3 (IQR 37.5–87.5)	62.5 (IQR 37.5–75)	NS
RE	100 (IQR 0–100)	66.7 (IQR 0–100)	NS	100 (IQR 0–100)	33.3 (IQR 0–100)	NS
MH	68 (IQR 52–80)	62.5 (IQR 37.5–75)	NS	72 (IQR 56–88)	64 (IQR 48–80)	NS

HRQOL: health related quality of life; IQR: interquartile range; NS: not significant. For other abbreviations, see Tables 1 and 2.

Table 5. Sensitivity of EQ-5D and SF-6D for patients changing according to external anchors.

	n	Index Change (± SD)*	p	Effect Size
SF-6D				
EQ-VAS improved	12	0.065 (± 0.092)	0.034	0.436
EQ-VAS declined	8	-0.038 (± 0.17)	0.54 (NS)	0.372
Self-reported improvement	28	0.045 (± 0.013)	0.002	0.321
Self-reported decline	15	-0.002 (± 0.036)	0.957 (NS)	0.074
SLEDAI improved	25	0.028 (± 0.017)	0.103 (NS)	0.200
SLEDAI declined	15	-0.006 (± 0.034)	0.864 (NS)	0.042
EQ-5D				
EQ-VAS improved	12	0.071 (± 0.158)	0.114 (NS)	0.428
EQ-VAS declined	9	-0.08 (± 0.02)	0.274 (NS)	0.27
Self-reported improvement	28	0.069 (± 0.033)	0.048	0.353
Self-reported decline	15	-0.008 (± 0.043)	0.854 (NS)	0.080
SLEDAI improved	25	0.024 (± 0.036)	0.500 (NS)	0.12
SLEDAI declined	15	-0.024 (± 0.029)	0.537 (NS)	0.012

\* Change in index signifies the difference in summary score between visits (visit 2 minus visit 1). Average time between visits = 8 months. NS: not significant. For abbreviations, see Table 1.

The results of the EQ-5D are consistent with those of a Canadian study by Wang, *et al*, despite that study being limited by its small sample size (55 patients), lack of multi-ethnicity (Caucasian > 83%), and lack of correlative information concerning disease activity and damage<sup>18</sup>. In addition, these data augment the recent report that SF-6D is a valid instrument in US-based patients with SLE, as it independently predicts damage<sup>19</sup>.

In the selection of an appropriate HRQOL tool for SLE, several issues have been identified that should be considered<sup>10</sup>. First, does the instrument measure domains relevant to patients with SLE? Second, does it have good psychometric properties: reliability and validity? Third, is it accepted by patients and by researchers and clinicians using it? Fourth, is the measure widely available and validated across multiple languages and in different countries? Last, is the measure sensitive to longitudinal changes in disease?

The results of our study support the validity and reliability of the SF-6D and EQ-5D for use in patients with SLE in the US, and highlight the disparity between HRQOL and disease activity and damage measures for patients with SLE, consistent with previous studies that have also failed to identify a significant relationship between SLE disease activity and patient-reported health status<sup>8,20</sup>. Most generic HRQOL instruments, including the widely used SF-36, are not sensitive to change in SLE disease activity over time, primarily due to the well described disconnect between clinical measures and patient self-reported perceptions of health in SLE<sup>20,26</sup>. For this reason, SLEDAI use as a clinical anchor to determine the responsiveness of EQ-5D or SF-6D in our study was thought to be problematic. The longitudinal analysis of 66 individuals suggested that both the EQ-5D and the SF-6D were insensitive to change in SLEDAI, but were sensitive to improvement of self-reported health by both SF-36 and EQ-5D VAS. While the SLEDAI identifies

a component of disease activity that is meaningful to clinicians and patient management, the lack of association with HRQOL suggests a disconnection between the clinical symptoms and effect of disease activity on the patient. This observation draws attention to the potential role of patient-reported outcome measure in the evaluation of healthcare interventions and the clinical management of patients with SLE, as these instruments provide information on the patient that is not given by provider-based assessment tools.

Although single summary scores by index-based measures such as the EQ-5D and SF-6D provide less information than multidimensional profile measures such as SF-36, they have several advantages. These measures tend to be brief and simple for the patient to self-complete. For instance, EQ-5D contains only 5 items plus a VAS and takes only a couple of minutes to complete. Additionally, a single summary score is simpler than multiple measures for clinicians and researchers to understand and to follow longitudinally, and can be used to convey burden of illness relative to other conditions. EQ-5D and SF-6D provide preference-based scores that can be used to facilitate the calculation of quality-adjusted life-years to evaluate the comparative effectiveness of treatments, both within SLE and across multiple disease states.

We observed higher survey completion rates for the EQ-5D and SF-6D than for the SF-36, which is consistent with prior reports that the EQ-5D and SF-6D have good patient acceptability and low survey burden<sup>18,19</sup>. However, higher completion rates may also be due in part to the sequential order in which the surveys are presented to patients, with both survey burden and perceived similarity of questions playing a role. EQ-5D and SF-6D were strongly correlated with each other, with EQ-5D more likely to have ceiling effects<sup>27,28</sup>. The limitations of each measure should be considered during instrument selection, and *a priori*

knowledge of the anticipated HRQOL of the target population may be helpful in making the selection.

At present, EQ-5D and SF-6D are widely used internationally to evaluate HRQOL in nonrheumatological<sup>29,30</sup> and rheumatological conditions<sup>31-33</sup>. Further, both the EQ-5D and SF-6D have been translated into most major languages<sup>34,35</sup>. The wide use of the EQ-5D and SF-6D makes it possible to compare HRQOL across conditions, and to do economic evaluations for resource allocation decisions in healthcare.

One limitation of our study was the lack of ability to evaluate test-retest reliability due to a relatively small sample. The optimal duration to assess test-retest reliability is to perform retests within 2–4 weeks to ensure that clinical and HRQOL characteristics of the subjects are relatively stable. In our study, however, the repeat assessments were only performed after several months, and therefore we did not evaluate test-retest reliability. The poor correlation between SLEDAI and HRQOL limited the possibility of establishing the minimal clinically important difference for the EQ-5D and SF-6D using an external clinically-based anchor. Consequently, responsiveness was evaluated using patient-reported change in HRQOL, for which both the EQ-5D and SF-6D were sensitive to improvement. A final limitation was the use of the SF-36 as the comparator for the validation of the EQ-5D and the SF-6D in SLE, in the absence of a validated US-based disease-specific measure of HRQOL for SLE at the time the study was undertaken<sup>11,26</sup>. Since then, LupusQoL has been validated for use among US patients<sup>12</sup>. We employed the SF-36 because it is the most widely used and accepted generic, psychometric-based measure of HRQOL, but it may not identify all HRQOL domains that are important to patients with SLE, including sleep and sexual functioning. A larger validation study would be desirable to further support the psychometric properties of these measures, particularly test-retest reliability and responsiveness.

Our findings provide information on the validity and responsiveness of EQ-5D and SF-6D in an ethnically diverse US-based population with SLE, and also highlight the disparity between disease activity, damage, and HRQOL for these patients. While neither instrument was sensitive to change in disease activity, both are sensitive to improvement in self-reported health status. Our results suggest that EQ-5D and SF-6D have acceptable properties among respondents, regardless of their demographic characteristics or their disease severity. Hence, the EQ-5D and SF-6D provide meaningful information about patients with SLE using a single summary score of HRQOL that can be compared within groups of patients with SLE and across disease states.

## REFERENCES

1. Siegel M, Lee SL. The epidemiology of systemic lupus erythematosus. *Semin Arthritis Rheum* 1973;3:1-54.
2. Urowitz MB, Gladman DD, Abu-Shakra M, Farewell VT. Mortality studies in systemic lupus erythematosus. Results from a single center. III. Improved survival over 24 years. *J Rheumatol* 1997;24:1061-5.
3. Jolly M. How does quality of life of patients with systemic lupus erythematosus compare with that of other common chronic illnesses? *J Rheumatol* 2005;32:1706-8.
4. McElhone K, Abbott J, Teh LS. A review of health related quality of life in systemic lupus erythematosus. *Lupus* 2006;15:633-43.
5. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. *Lupus* 1999;8:685-91.
6. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996;39:363-9.
7. Gladman DD, Urowitz MB, Ong A, Gough J, MacKinnon A. Lack of correlation among the 3 outcomes describing SLE: disease activity, damage and quality of life. *Clin Exp Rheumatol* 1996;14:305-8.
8. Jolly M, Utset TO. Can disease specific measures for systemic lupus erythematosus predict patients health related quality of life? *Lupus* 2004;13:924-6.
9. Strand V, Gladman D, Isenberg D, Petri M, Smolen J, Tugwell P. Outcome measures to be used in clinical trials in systemic lupus erythematosus. *J Rheumatol* 1999;26:490-7.
10. Thumboo J, Strand V. Health-related quality of life in patients with systemic lupus erythematosus: an update. *Ann Acad Med Singapore* 2007;36:115-22.
11. McElhone K, Abbott J, Shelmerdine J, et al. Development and validation of a disease-specific health-related quality of life measure, the LupusQoL, for adults with systemic lupus erythematosus. *Arthritis Rheum* 2007;57:972-9.
12. Jolly M, Pickard AS, Wilke C, et al. Lupus specific health outcome measure for US patients: The LupusQoL-US Version(C). *Ann Rheum Dis* 2009 Jan 6. [Epub ahead of print]
13. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
14. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53-72.
15. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271-92.
16. Ariza-Ariza R, Hernandez-Cruz B, Navarro-Sarabia F. EuroQol is a useful instrument for assessing the health-related quality of life of the patients with systemic lupus erythematosus. *Lupus* 2005;14:334-5.
17. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43:203-20.
18. Wang C, Mayo NE, Fortin PR. The relationship between health related quality of life and disease activity and damage in systemic lupus erythematosus. *J Rheumatol* 2001;28:525-32.
19. Fernandez M, Alarcon GS, McGwin G Jr, et al. Using the Short Form 6D, as an overall measure of health, to predict damage accrual and mortality in patients with systemic lupus erythematosus: XLVII, results from a multiethnic US cohort. *Arthritis Rheum* 2007;57:986-92.
20. Kuriya B, Gladman DD, Ibanez D, Urowitz MB. Quality of life over time in patients with systemic lupus erythematosus. *Arthritis Rheum* 2008;59:181-5.
21. Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271-7.
22. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res*

- 1993;2:441-9.
23. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 1988.
  24. Guidelines for the management of rheumatoid arthritis: 2002 Update. *Arthritis Rheum* 2002;46:328-46.
  25. Alarcon GS, McGwin G Jr, Uribe A, et al. Systemic lupus erythematosus in a multiethnic lupus cohort (LUMINA). XVII. Predictors of self-reported health-related quality of life early in the disease course. *Arthritis Rheum* 2004;51:465-74.
  26. Leong KP, Kong KO, Thong BY, et al. Development and preliminary validation of a systemic lupus erythematosus-specific quality-of-life instrument (SLEQOL). *Rheumatology* 2005;44:1267-76.
  27. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004;13:873-84.
  28. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Econ* 2005;14:1169-89.
  29. Kortt MA, Clarke PM. Estimating utility values for health states of overweight and obese individuals using the SF-36. *Qual Life Res* 2005;14:2177-85.
  30. Pickard AS, Johnson JA, Feeny DH. Responsiveness of generic health-related quality of life measures in stroke. *Qual Life Res* 2005;14:207-19.
  31. Boonen A, van der Heijde D, Landewe R, et al. How do the EQ-5D, SF-6D and the well-being rating scale compare in patients with ankylosing spondylitis? *Ann Rheum Dis* 2007;66:771-7.
  32. Bansback N, Marra C, Tsuchiya A, et al. Using the Health Assessment Questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. *Arthritis Rheum* 2007;57:963-71.
  33. Harrison MJ, Davies LM, Bansback NJ, Ingram M, Anis AH, Symmons DP. The validity and responsiveness of generic utility measures in rheumatoid arthritis: a review. *J Rheumatol* 2008;35:592-602.
  34. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33:337-43.
  35. Lam CL, Brazier J, McGhee SM. Valuation of the SF-6D health states is feasible, acceptable, reliable, and valid in a Chinese population. *Value Health* 2008;11:295-303.