LUPUS SCIENCE & MEDICINE™

# Patient-reported outcome measures for use in clinical trials of SLE: a review

Zara Izadi,[1] Julie Gandrup,[2] Patricia P Katz,[2] Jinoos Yazdany[2]

**To cite:** Izadi Z, Gandrup J, Katz PP, *et al*. Patient-reported outcome measures for use in clinical trials of SLE: a review. *Lupus Science & Medicine* 2018;**5**:e000279. doi:10.1136/lupus-2018-000279

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/lupus-2018-000279).

Received 28 May 2018
Revised 19 July 2018
Accepted 21 July 2018



[1]Department of Epidemiology and Biostatistics, University of California, San Francisco, California, USA
[2]Division of Rheumatology, University of California, San Francisco, California, USA

**Correspondence to**
Dr Zara Izadi; zara.izadi@ucsf.edu

## ABSTRACT

Inclusion of patient-reported outcomes is important in SLE clinical trials as they allow capture of the benefits of a proposed intervention in areas deemed pertinent by patients. We aimed to compare the measurement properties of health-related quality of life (HRQoL) measures used in adults with SLE and to evaluate their responsiveness to interventions in randomised controlled trials (RCTs). A systematic review was undertaken using full original papers in English identified from three databases: MEDLINE, EMBASE and PubMed. Studies describing the validation of HRQoL measures in English-speaking adult patients with SLE and SLE drug RCTs that used an HRQoL measure were retrieved. Twenty-five validation papers and 26 RCTs were included in the indepth review evaluating the measurement properties of 4 generic (Medical Outcomes Study Short-Form 36 (SF36), Patient Reported Outcomes Measurement Information System (PROMIS) item-bank, EuroQol-5D, and Functional Assessment of Chronic Illness Therapy-Fatigue) and 3 disease-specific (Lupus Quality of Life (LupusQoL), Lupus Patient Reported Outcomes, Lupus Impact Tracker (LIT)) instruments. All measures had good convergent and discriminant validity. PROMIS provided the strongest evidence for known-group validity and reliability among generic instruments; however, data on its responsiveness have not been published. Across measures, standardised response means were generally indicative of poor-moderate sensitivity to longitudinal change. In RCTs, clinically important improvements were reported in SF36 scores from baseline; however, between-arm differences were frequently non-significant and non-important. SF36, PROMIS, LupusQoL and LIT had the strongest evidence for acceptable measurement properties, but few measures aside from the SF36 have been incorporated into clinical trials. This review highlights the importance of incorporating a broader range of SLE-specific HRQoL measures in RCTs and warrants further research that focuses on longitudinal responsiveness of newer instruments.

## INTRODUCTION

SLE is a chronic inflammatory autoimmune disorder with variable multisystem involvement, an unpredictable relapsing–remitting course, an early onset and a significant impact on health-related quality of life (HRQoL).[1] Previous research has shown poor correlation between HRQoL and physician assessments of disease activity and damage, highlighting the distinct contribution of HRQoL data to understanding patient trajectories and supporting the need for its assessment in SLE.[2] Further, HRQoL has been found to be an important determinant of adherence and healthcare utilisation in patients with SLE and may facilitate justifying the considerable costs of new therapies.[3] Therefore, both the US Food and Drug Administration and the European Medicines Agency advocate use of patient-reported instruments such as those measuring HRQoL in clinical trials (guidelines available at fda.gov and ema.europa.eu, respectively).

Patient-reported HRQoL measures, in the form of questionnaires, have been either developed exclusively for use in SLE (disease-specific measures) or have been used in patients with SLE but developed for any disease state or healthy individuals (generic measures). Patient-reported outcome evaluation has been incorporated in drug clinical trials in SLE; however, it has not been a consistent practice,[4] and it is not clear whether sensitivity to change over time has been observed.[5 6] Knowledge of acceptable measurement standards, responsiveness to change, generalisability and cultural adaptability would help determine the adequacy of the HRQoL measure for clinical research.

The aims of this systematic review were (1) to compare the measurement properties of published HRQoL measures that have been developed and/or evaluated for use in adults with SLE and (2) to evaluate the responsiveness of validated HRQoL measures used in SLE randomised controlled trials (RCTs) to date. Our goal was to provide a comprehensive review of these outcome measures to inform future selection of these tools in SLE clinical trials.

## METHODS

### Search strategy

Literature searches were conducted in MEDLINE, EMBASE and PubMed, limited to humans, English language and articles published between inception and 1 April

2018. Journal articles (excluding conference abstracts, letters to editor, dissertations and book chapters) containing the keywords in the title and/or abstract were included (search terms are available in online supplementary material 1).

For our first aim, we included papers that described the methodology of the development and validation of HRQoL measures in SLE, and papers that described the evaluation of an existing HRQoL measure or its translated/adapted version for patients with SLE. Exclusion criteria were inadequate numbers of patients with SLE (<50% of the study population) and patients <18 years old. For our second aim, we included drug RCTs (pilot studies, phase I, II and III) in patients with SLE with published HRQoL data. Exclusion criteria were transplantation or plasma exchange RCTs, cutaneous lupus RCTs and RCTs in patients <18 years old.

The selected articles were categorised into (1) validation studies of extensively published HRQoL instruments (defined as having >3 validation studies in English-speaking SLE populations) and HRQoL instruments that had been used in an RCT, and (2) RCTs that used a validated HRQoL instrument.

## Outcomes

### Measurement properties

To assess measurement properties, we evaluated floor and ceiling effects, construct validity, test–retest reliability, internal consistency, and responsiveness. The instrument was considered to have floor or ceiling effects if >15% of the respondents scored at the extreme ends of the scale.[7] Construct validity was determined using convergent and discriminant validity and known-group validity. Convergent validity was judged to be adequately demonstrated if there were high (>0.6) positive correlations between scales and discriminant validity, if correlations were low (<0.3) or if they were negative.[8] Known-group validity was adequate if group means differed by ≥0.5 SD.[9] Test–retest reliability was gauged by the intraclass correlation coefficient (ICC) and considered adequate if ICC was >0.7.[10] The acceptable statistical value for internal consistency was a Cronbach's α >0.7.[10] Responsiveness was compared using standardised response means (SRMs) and considered poor if SRMs were <0.5, moderate if SRMs were ≥0.5 and high if SRMs were ≥0.8.[11] The generalisability of the instrument was assessed by establishing if the study population was adequately described to help investigators extrapolate the results to their study cohorts. For each measure, we also determined if estimates of SLE-specific minimally important differences (MID) were available. While our review focused on English-speaking populations, we noted the availability of validation studies in non-English-speaking populations for each measure.

### Responsiveness to interventions in RCTs

The results of any intervention during RCTs were interpreted in the context of the MID value of the instrument being used. We first determined if the direction of change in HRQoL scale was consistent with clinical changes measured by disease activity, damage or flare indices. We then determined if between-arm differences in HRQoL were ≥MID. Finally, we determined if HRQoL changes from baseline were ≥MID.

## Screening process and data extraction

Screening and data extraction were performed by two independent researchers using predesigned templates and in line with the Centre for Reviews and Dissemination guidance, available at york.ac.uk/crd. Any disagreements were discussed and resolved based on consensus opinion. In addition to the outcomes, demographics, clinical data and information on instrument characteristics were extracted.

After removal of duplicates, screening of titles and abstracts, and full-text review, 23 validation studies in English-speaking patients with SLE that met the inclusion criteria were identified and were selected for the indepth review (figure 1). Of the 23 studies, 21 focused on 5 HRQoL instruments (the Medical Outcomes Study Short-Form 36 (SF36), the Patient Reported Outcomes Measurement Information System item-bank (PROMIS), the Lupus Quality of Life (LupusQoL), the Lupus Patient Reported Outcomes and the Lupus Impact Tracker (LIT)) and 2 papers described the validation of the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) and EuroQoL-5D (EQ5D). The references of the selected 23 papers were also screened for additional relevant papers and 2 further papers were identified and included in the review. We excluded 25 studies that described the measurement properties of 22 additional HRQoL instruments not used in an RCT setting, with 1–3 studies per instrument (list available in online supplementary material 1).

After removal of duplicates, screening of title and abstracts, and full-text review, 26 papers describing an RCT in SLE with HRQoL data were identified. HRQoL instruments used included the SF36 (25 papers), EQ5D (1 paper) and FACIT-Fatigue (3 papers). The references of the selected 26 papers were also screened for additional relevant papers and no further papers were identified.

# RESULTS

Table 1 provides information on demographics and clinical characteristics of participants with SLE in patient-reported outcome validation studies. Information on instrument characteristics is provided in table 2. Measurement properties are summarised in table 3, and data on measure responsiveness from RCTs are summarised in table 4.

## Medical Outcomes Study Short-Form 36

We found only one study that reported the frequency of maximum and minimum obtainable scores; role emotional and role physical domains were found to have significant floor and ceiling effects.[12] The instrument had good internal consistency; however, its test–retest

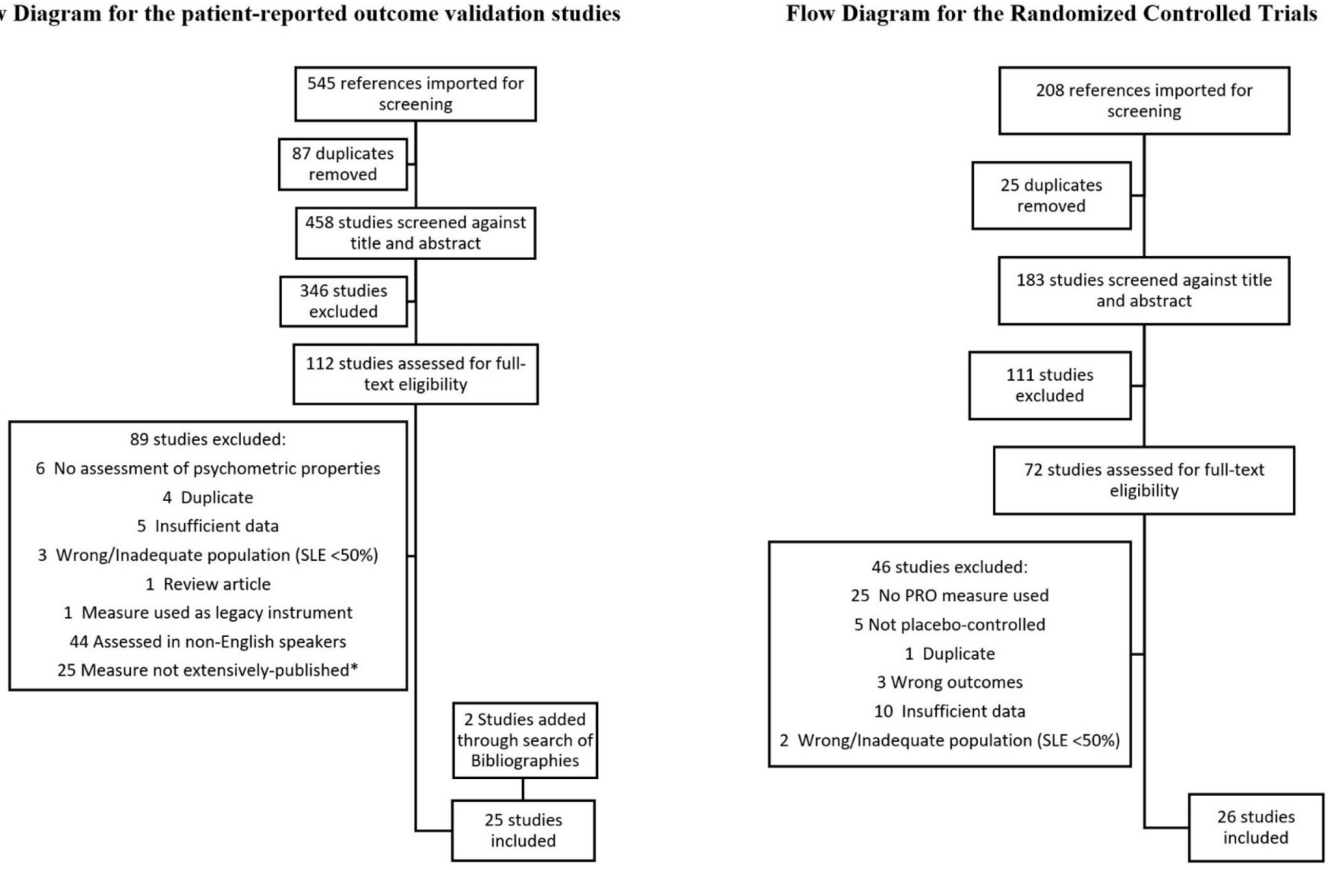


**Figure 1** Of the 25 validation papers identified, 8 assessed the measurement properties of Short-Form 36, 1 assessed EuroQoL-5D, 7 assessed Lupus Quality of Life questionnaire, 5 assessed Patient Reported Outcomes Measurement Information System item-bank, 3 assessed Lupus Patient Reported Outcomes questionnaire, 6 assessed Lupus Impact Tracker, and 1 assessed Functional Assessment of Chronic Illness Therapy-Fatigue. Some studies assessed multiple quality of life instruments. *Extensively published defined as having >3 validation studies in English-speaking SLE populations or having been used in an RCT. Of the 26 RCT papers identified, 25 used Short-Form 36, 3 used Functional Assessment of Chronic Illness Therapy-Fatigue and 1 used EuroQol-5D. Some studies used >2 quality of life instruments. PRO, patient-reported outcome; RCT, randomised controlled trial.

reliability using ICC is currently unknown.[13 14] The Health Assessment Questionnaire scores correlated strongly with the SF36 physical function scores (r=0.75) and moderately with role physical, bodily pain and vitality scores (r=0.41–0.48).[15] Weak-moderate correlations (r≤0.41) were reported with various disease activity or damage indices.[12–15] The mean SF36 scores differed significantly across categories of disease activity; however, the effect size was not reported.[13 16] Studies examining responsiveness were inconsistent in the methodology but seem to suggest that the measure is not particularly responsive in SLE. SRMs were poor in most domains and inconsistent (poor to moderate) across different anchors.[17 18] SLE-specific MIDs have been reported using anchor-based and distributions-based methods. Anchor-based MIDs for improvement ranged from 2.8 to 10.9 for domains and from 2.1 to 2.4 for summary scores, which is consistent with literature-reported estimates from other rheumatological conditions (5–10 points for domains and 2.5–5 for summary scores).[17 18] Validation studies using the Chinese, French and Turkish versions of the SF36 have shown results that are comparable with the English version.[19–22]

In RCTs, we assessed responsiveness in two ways. First, we determined if between-arm differences met or exceeded the commonly accepted MID (5–10 points for domains and 2.5–5 for summary scores). Then we determined if improvements over time (from baseline) were ≥MID. We found that although SF36 scores generally improved over time, between-arm differences were clinically non-important, implying that the SF36 is not responsive to interventions.[23–48] Among 10 studies that met the primary efficacy endpoint and had sufficient data for analysis, 6 reported improvements in SF36 scores from baseline that were ≥MID (table 4), while only 2 reported between-arm differences in SF36 scores that were ≥MID in all or most domains. Within-arm improvements from baseline were not limited to RCTs that achieved the primary efficacy endpoint. Improvements in SF36 scores from baseline were also ≥MID in 6 of 10 RCTs that reported a statistically non-significant clinical improvement.

**Table 1** Demographics and clinical characteristics of participants with SLE in patient-reported outcome validation studies

| Measure | First author and year | Citation | Patients in the study (n) | Female, n (%) | Age in years, mean (SD) | SLE disease duration in years, mean (SD) | Disease activity, median (range) | Disease damage, median (range) |
|---|---|---|---|---|---|---|---|---|
| SF36 | Nantes 2018[12] | 9 | 78 | NS (90) | 39.2 (13.3) | 11.9 (7.6) | SLEDAI 9.7 (4.8)* | SDI 1.7 (1.6)* |
| | Stoll 1997[13] | 10 | 150 | 143 (95) | 40 (11.0) | 10 (7) | BILAG 5 (3–7) | SLICC 1 (0–2) |
| | Thumboo 1999[14] | 11 | 118 | 112 (94.5) | NS | 3.61 (0.01–16.1) | BILAG 2 (0–15) | SLICC 0 (0–8) |
| | Fortin 1998[15] | 12 | 96 | NS (90) | 36 (28–46)† | 15 (10–19)† | SLEDAI 6.2 (6.4)* | SLICC 1.29 (1.62)* |
| | Saba 2003[16] | 13 | 219 | NS (94.4) | 45 (NS) | NS | m-SLAM 4 (0–17)‡ | NS |
| | Touma 2011[59] | 14 | 41 | 37 (NS) | 45.3 (13.2) | 14.8 (10.3) | SLEDAI 2.59 (2.41) | NS |
| | McElhone 2016[17] | 15 | 101 | 95 (94) | 40.9 (12.8) | 9.3 (8.1) | BILAG 16.4 (8.1)* | SLICC 0.56 (1.18)* |
| | Colangelo 2009[18] | 16 | 202 | NS (94) | 50 (14) | 10 (7.4) | SLEDAI 6.1 (4.7)* | SLICC 1.4 (1.7)* |
| EQ5D | Aggarwal 2009[55] | 54 | 167 | 156 (93.5) | 42.5 (13.0) | 9.3 (8.8) | SLEDAI 6.2 (5.7)* | SLICC 2 (2) |
| FACIT-Fatigue | Lai 2011[54] | 52 | 254 | 232 (90.9) | 40.3 (11.9) | NS | NS | NS |
| PROMIS | Kasturi 2017[50] | 48 | 204 | NS (93) | 40.0 (13.2) | 12.2 (8.8) | SLEDAI 4.2 (3.5)* | SDI 1.2 (1.7)* |
| | Kasturi 2018[51] | 49 | 204 | NS (93) | 40.0 (13.2) | 12.2 (8.8) | SLEDAI 4.2 (3.5)* | SDI 1.2 (1.7)* |
| | Katz 2017[49] | 47 | 240 | 219 (91) | 58 (13) | 22 (12) | SLAQ 10[8]* | NS |
| | Lai 2017[53] | 51 | 333 | 305 (91.6) | 44.7 (12.5) | 12.3 (9.6) | NS | NS |
| | Mahieu 2016 | 50 | 123 | 116 (94.3) | 45.3 (10.8) | NS | SLEDAI 2.3 (2.8)* | SLICC 1.7 (2.2)* |
| LupusQoL | McElhone 2007[58] | 57 | 160 | 152 (95) | 45.3 (13.9) | NS | NS | NS |
| | McElhone 2010[61] | 60 | 322 | 257 (NS) | 45.0 (13.4) | 10.5 (8.6) | NS | SLICC 0.75 (1.25)* |
| | Nantes 2018[12] | 9 | 78 | NS (90) | 39.2 (13.3) | 11.9 (7.6) | SLEDAI 9.7 (4.8)* | SDI 1.7 (1.6)* |
| | Devilliers 2017[57] | 56 | 295 | NS (90) | 42 (17–75)‡ | NS | SLEDAI 4 (2–6) | NS |
| | Touma 2011[59] | 14 | 41 | 37 (NS) | 45.3 (13.2) | 14.8 (10.3) | SLEDAI 2.59 (2.41) | NS |
| | McElhone 2016[17] | 15 | 101 | 95 (94) | 40.9 (12.8) | 9.3 (8.1) | BILAG 16.4 (8.1)* | SLICC 0.56 (1.18)* |
| | Jolly 2010[60] | 59 | 186 | NS (94) | 42.5 (12.9) | NS | SLEDAI 6.2 (5.8)* | SLICC 2.0 (2.1)* |
| LupusPRO | Jolly 2012[71] | 71 | 323 | NS (93) | 45.1 (12.3) | 9.3 (7.6) | SLEDAI 3.9 (3.8)* | SDI 0.7 (1.1)* |
| | Azizoddin 2017[73] | 72 | 137 | 124 (90.5) | 40.4 (14.1) | 8.13 (6.84) | SLEDAI 4.89(4.43)* | SDI 0.61 (1.05)* |
| | Bourré-Tessier 2014[74] | 73 | 123 | NS (94) | 47.7 (14.8) | NS | SLEDAI 4.0 (6.0)† | SDI 1.0 (3.0)† |
| LIT | Jolly 2014[84] | 83 | 308 | NS (93) | 43.3 (13.3) | NS | SLEDAI 3.9 (3.8)* | NS |
| | Jolly 2016[83] | 82 | 325 | NS (90) | 41.9 (13) | NS | SLEDAI 4.28 (3.8)* | SLICC 2.17 (1.59)* |
| | Giangreco 2015[87] | 86 | 182 | 170 (93) | 43.5 (13.2) | NS | SLEDAI 6.4 (7.3)* | NS |
| | Antony 2017[85] | 84 | 73 | NS (90) | 41 (19–74)‡ | NS | SLEDAI 4 (0–16)‡ | SDI 1 (0–6)‡ |
| | Brandt 2017[86] | 85 | 860 | 807 (93.8) | 45.8 (13.4) | 13.4 (9.2) | SLAQ 17.2 (9.4)* | NS |
| | Devilliers 2017[57] | 56 | 295 | NS (90) | 42 (17–75)‡ | NS | SLEDAI 4 (2–6) | NS |

Continued

Table 1 Continued

| Measure | First author and year | Citation | Patients in the study (n) | Female, n (%) | Age in years, mean (SD) | SLE disease duration in years, mean (SD) | Disease activity, median (range) | Disease damage, median (range) |
|---|---|---|---|---|---|---|---|---|

*mean (SD).
†median (IQR).
‡mean (range).
BILAG, British Isles Lupus Assessment Group Disease Activity Index; EQ5D, EuroQol-5D; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; LIT, Lupus Impact Tracker; LupusPRO, Lupus Patient Reported Outcomes; LupusQoL, Lupus Quality of Life questionnaire; mSLAM, modified Systemic Lupus Activity Measure; NS, not shown; PROMIS, Patient Reported Outcomes Measurement Information System item-bank; SDI, The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (ACR) Damage Index; SF36, Medical Outcomes Study Short-Form 36; SLAQ, Systemic Lupus Activity Questionnaire; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; SLICC, Systemic Lupus International Collaborating Clinics.

### Patient Reported Outcomes Measurement Information System

Significant ceiling effects were reported in most domains of the 29-item profile,[49] while no floor or ceiling effects were observed among 14 Computer Adaptive Tests (CATs).[50] PROMIS measures had good internal consistency and test–retest reliability, although the internal consistency of the PROMIS CATs remains to be established. PROMIS scores correlated strongly (r>0.6) with other HRQoL instruments across comparable domains and weakly to moderately (r≤0.6) across divergent domains.[49–52] Correlations with disease activity indices, physician global assessment, damage and physical activity (using an accelerometer) were mostly weak (<0.3).[50–52] Patients with SLE scored 0.5 SD or worse than the general population across most domains.[49–53] Longitudinal responsiveness and MIDs have not been published. While PROMIS measures have been translated into many other languages including Spanish and Chinese, additional studies are needed to validate PROMIS measures in non-English-speaking patients with SLE.[50] PROMIS measures remain to be used in RCTs.

### Functional Assessment of Chronic Illness Therapy-Fatigue

We found only one study that reported the measurement properties of FACIT-Fatigue in patients with SLE. The measure was found to have good internal consistency; however, the test–retest reliability and the floor and ceiling effects are currently unknown.[54] FACIT-Fatigue had moderate-high correlations (r=0.5–0.8) with SF36, brief pain inventory and patient global assessment, but poor correlations with disease activity and physician global assessments (r=0.1–0.3). Cross sectionally, FACIT-Fatigue has good discrimination between remission-mild versus moderate disease activity (0.52 SDs) but not between moderate versus severe disease activity (0.24 SDs). The measure was responsive to clinical improvements (SRM=0.69) but not clinical deteriorations. The measure was responsive to improvements (SRM=0.82) and deteriorations (SRM=0.53) in patient global assessment. Distribution and anchor-based estimates suggested an MID range of 3–6 points, which is consistent with literature reports of 3–4 points for patients with rheumatoid arthritis or cancer.

We identified three papers that used FACIT-Fatigue in an RCT setting (table 4). In all three, the intervention led to statistically non-significant improvements in disease activity. Two studies reported a change in FACIT-Fatigue scores from baseline that was ≥MID (4 points); however, only one study reported between-arm differences that were ≥MID.

### EuroQoL-5D

We found only one study that reported the measurement properties of EQ5D in patients with SLE.[55] No floor or ceiling effects were observed. Related domains on the EQ5D and SF36 correlated strongly (r=0.60), whereas unrelated domains showed weak-moderate correlation. Disease activity and damage showed weak correlation

**Table 2** Summary of characteristics of instruments used to measure patient-reported outcomes in SLE

| Measure | Instrument type | Items | Domains | Response format | Range of scores | Score interpretation | Mode of administration | Recall period | Normative data |
|---|---|---|---|---|---|---|---|---|---|
| SF36 | Generic | 36 | Physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, mental health and health change. The domains can be further summarised into two component scores: the physical component summary and the mental component summary. | Likert scales (range: 3–6) and yes/no answers. | 0–100 | Higher scores indicate better health. | Paper and electronic. | Standard form: 4 weeks; acute form: 1 week. | Available for both summary scores in several populations. |
| EQ5D | Generic | 6 | Health profile: mobility, self-care, usual activities, pain/discomfort, anxiety/depression; health-state rating scale: current health. | Health profile items: 3-point Likert scale; health-state rating scale: 20 cm VAS. | Health profile items are aggregated to give a score that ranges from –0.11 to 1; VAS ranges from 0 to 100. | Health profile items: 1 is perfect health, 0 implies death and negative values mean health states worse than death; VAS: higher scores denote better health. | Paper. | Same day (current state). | Available in several populations including the US general population. |
| FACIT-Fatigue | Generic | 13 | Fatigue. | 5-point Likert scale. | 0–52 | Higher scores indicate less fatigue. | Paper. | 1 week; a 4-week recall period has shown to be comparable from a measurement perspective. | Available through a link between FACIT-Fatigue and PROMIS fatigue item-bank. |
| PROMIS Short Forms | Generic | 4–10 items per domain | Cover one domain (various available spanning global, physical, social and mental health). | All items except pain intensity: 5-point Likert scale; pain intensity: 11-point rating from 0 to 10. | 1–100 | The higher the score, the more of that domain the patient experiences. | Paper and electronic. | All domains except physical function and social health: 7 days; physical function and social health: no time-frame. | Available for the US general population using a T-score metric. |
| PROMIS Profiles | | 4, 6 or 8 items per domain except pain intensity which is a single item | Physical function, pain interference, fatigue, sleep disturbance, satisfaction with participation in social roles, depression and anxiety, plus a single pain intensity item. | | | | | | |
| PROMIS CATs | | 4–12 items per CAT | Cover one domain (various available spanning global, physical, social and mental health). | | | | | | |
| LupusQoL | Disease-specific | 34 | Physical health, emotional health, body image, pain, planning, fatigue, intimate relationships and burden to others. | 5-point Likert scale. | 0–100 | Higher scores indicate better health. | Paper and electronic. | 4 weeks. | Not available. |
| LupusPRO | Disease-specific | 43 | HRQoL: SLE symptoms, cognition, SLE medications, physical health, pain vitality, body image, emotional health and procreation; non-HRQoL: desires/goals, relationship/social support, coping and satisfaction with medical care. | 5-point Likert scale. | 1– 100 | Higher scores indicate better health. | Paper. | 4 weeks. | Not available. |

Continued

**Table 2** Continued

| Measure | Instrument type | Items | Domains | Response format | Range of scores | Score interpretation | Mode of administration | Recall period | Normative data |
|---|---|---|---|---|---|---|---|---|---|
| LIT | Disease-specific | 10 | Concentration, medication side effects, fulfilling family responsibilities, feelings of being worn out on waking, bodily pain and aching, limitation of activities due to pain/fatigue, anxiety, depression, self-consciousness related to physical appearance, and ability to plan activities. | 5-point Likert scale. | 0–100 | Higher scores indicate greater negative impact of SLE. | Paper. | 4 weeks. | Not available. |

CAT, Computer Adaptive Tests; EQ5D, EuroQol-5D; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; HRQoL, health-related quality of life; LIT, Lupus Impact Tracker; LupusPRO, Lupus Patient Reported Outcomes; LupusQoL, Lupus Quality of Life questionnaire; PROMIS, Patient Reported Outcomes Measurement Information System item-bank; SF36, Medical Outcomes Study Short-Form 36; VAS, Visual Analogue Scale.

with EQ5D domains (r<0.22). The mean scores differed significantly across categories of disease activity but not damage. The measure showed poor responsiveness to self-reported change in health (SRMs ranged from 0.08 to 0.27 in patients who deteriorated and from 0.35 to 0.43 in patients who improved) but was not responsive to longitudinal changes in disease activity (SRM=0.01 in patients who deteriorated and 0.12 in patients who improved). SLE-specific MIDs have not been reported. EQ5D was shown to have good construct and criterion validity in a group of Chinese-speaking patients with SLE.[56]

EQ5D was used in one RCT that met our inclusion criteria. In this study, the intervention led to statistically non-significant improvements in disease activity. While changes in EQ5D from baseline were ≥MID in one of the intervention arms, between-arm differences did not reach MID (table 4).

### Lupus Quality of Life

Some domains were found to have significant floor and ceiling effects, including intimate relationships and planning.[12 57] The measure was found to have good internal consistency and test–retest reliability. LupusQoL had strong correlations with SF36 across comparable domains (r>0.6) and weak correlations with age, disease duration, disease activity and damage across all domains (r<0.30).[12 17 58–61] Scores differed significantly across categories of disease activity and damage in all domains except fatigue and intimate relationships.[58 60] The effect size has not been reported. SRMs were poor in most domains and inconsistent (poor to moderate) across different anchors.[57 59] SLE-specific MIDs derived using the anchor-based approach ranged from 2.4 to 8.7 for deteriorations and from 3.5 to 7.3 for improvements.[17] MIDs using distribution-based approaches based on 0.5 SD ranged from 12.9 to 16.7. Measurement properties of the LupusQoL have been examined and published in Chinese-speaking, Farsi-speaking, French-speaking, Italian-speaking, Spanish-speaking and Turkish-speaking populations.[62–70] A version adapted and validated for a US population is also available.[13]

We identified one RCT that used LupusQoL; however, the results have not been published.[39]

### Lupus Patient Reported Outcomes

The measure was found to have significant floor effects in satisfaction with medical care; data on ceiling effects were inconclusive, with one study reporting no ceiling effects and another reporting significant ceiling effects in all domains except coping.[71 72] Good internal consistency and test–retest reliability were reported in most domains.[71–73] Procreation and satisfaction with care had the lowest ICCs. Moderate-strong correlations (r≥0.5) were reported with the SF36 across comparable domains, while correlations with disease activity, physician global assessment, damage and flare were weak-moderate (r≤0.50).[71–73] In cross-sectional analyses, significant associations were reported with categories of patient-reported

**Table 3** Summary of measurement properties of HRQoL instruments used in SLE

| Instrument | Floor/ceiling effect | Internal consistency | Test–retest reliability | Convergent validity | Discriminant validity | Known-group validity | Responsiveness | MID | Generalisability | Cross-cultural validation | Responsiveness in RCTs |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SF36 | Present in some domains. | Cronbach's $\alpha \geq 0.71$. | Not assessed. | Valid. | Valid. | Some evidence but effect size not reported. | Poor in most domains and inconsistent across anchors. | Available. | Yes. | Chinese, French and Turkish versions. | Changes from baseline were ≥MID in most studies but between-arm differences did not reach MID. |
| EQ5D | None reported. | Not assessed. | Not assessed. | Valid. | Valid. | Inconclusive evidence; effect size not reported. | Not responsive. | Not available. | Yes. | Chinese version. | Current evidence is limited and inconclusive. |
| FACIT-Fatigue | Not assessed. | Cronbach's $\alpha \geq 0.95$. | Not assessed. | Valid. | Valid. | Partially. | Medium responsiveness to clinical improvements but not deteriorations; responsive to patient-reported changes in health. | Available. | No indication of clinical disease activity, damage or disease duration. | None. | Current evidence is limited and inconclusive. |
| PROMIS item-bank | Ceiling effects reported for 29-item profile but not PROMIS CATs. | Cronbach's $\alpha \geq 0.88$.* | ICC>0.7. | Valid. | Valid. | Demonstrated. | Not reported. | Not available. | Yes. | None. | Not used in RCTs. |
| LupusQoL | Present in some domains. | Cronbach's $\alpha \geq 0.85$. | ICC≥0.68. | Valid. | Valid. | Some evidence but effect size not reported. | Poor in most domains and inconsistent across anchors. | Available. | Yes. | Italian, French, Spanish, Chinese, Turkish and Farsi versions. | Used in one RCT but data not reported. |
| LupusPRO | Present in some domains. | Cronbach's $\alpha \geq 0.60$. | ICC≥0.55. | Valid. | Valid. | Some evidence but effect size not reported. | Scores changed significantly in response to longitudinal changes in health but effect size has not been reported. | Not available. | Yes. | Tagalog, Turkish, Spanish, French, Italian, Japanese, and Chinese versions. | Not used in RCTs. |
| LIT | Not assessed. | Cronbach's $\alpha \geq 0.84$. | ICC≥0.83. | Valid. | Valid. | Some evidence but effect size not reported. | Some evidence for medium responsiveness to clinical improvements but not deteriorations. | Available. | Yes. | German, Italian, Spanish, Swedish and French versions. | Not used in RCTs. |

*Internal consistency of PROMIS CATs remains to be established.
CAT, Computer Adaptive Tests; EQ5D, EuroQol-5D; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; ICC, intraclass correlation; LIT, Lupus Impact Tracker; LupusPRO, Lupus Patient Reported Outcomes; LupusQoL, Lupus Quality of Life questionnaire; MID, minimally important differences; RCTs, Patient Reported Outcomes Measurement Information System item-bank; RCTs, randomised controlled trials; SF36, Medical Outcomes Study Short-Form 36.

**Table 4** Summary of measure responsiveness from RCTs

| Reference | Citation | n | HRQoL Instrument | Clinical disease activity indicator | Time of endpoint assessment (from baseline) | Between-arm clinical improvement | Between-arm HRQoL improvement | Between-arm differences in HRQoL ≥MID | Within-arm HRQoL change from baseline (active arm or pooled) ≥MID | Notes |
|---|---|---|---|---|---|---|---|---|---|---|
| Galbraith *et al*, 2014[23] | 21 | 19 | SF36 | Renal or major non-renal SLE relapse | 13 weeks | Yes | Unclear | Yes | Unclear | |
| Uppal *et al*2009[24] | 22 | 27 | SF36 | SLEDAI | 6 months | Yes | Yes | Partially | Yes | |
| Nordmark *et al*, 2005[25] | 23 | 37 | SF36 | mSLEDAI | 6 months | Yes, but not SS | Yes | Partially | Yes | |
| Dussán *et al*, 2008[26] | 24 | 47 | SF36 | SRI | 6 months | Not reported | Unclear | Partially | Yes | |
| Arriens *et al*, 2015[27] | 25 | 50 | SF36 | SLEDAI | 6 months | Yes, but not SS | Yes | Yes | Yes | |
| Danowski *et al*, 2006[28] | 26 | 50 | SF36 | Patient-reported SLE activity | 4 weeks | Yes, but not SS | Yes | No | Not reported | |
| Hartkamp *et al*, 2010[29] | 27 | 60 | SF36 | SLEDAI | 5 months | Yes, but not SS | Yes | Not reported | Not reported | |
| Furie *et al*, 2008[30] | 28 | 70 | SF36 | SLEDAI | Day 20 | Yes, but not SS | Unclear | No | Not reported | |
| Gordon *et al*, 2008[31] | 29 | 71 | SF36 | SLEDAI | 12 weeks | Yes, but not SS | Yes | Partially | Partially | |
| Grootscholten *et al*, 2007[32] | 30 | 87 | SF36 | SLEDAI | 1 year | No | No | No | Yes | |
| Griffiths *et al*, 2010[33] | 31 | 90 | SF36 | BILAG | 12 months | No | Yes | Partially^ | Unclear” | ^In MCS but not PCS; "analyses are adjusted at 12 months. |
| Strand 2014 | 32 | 90 | SF36 | BILAG | 12 weeks | Yes | Yes | Yes | Yes | |
| Wallace *et al*, 2016[35] | 33 | 113 | SF36 | SLEDAI-2K | 108 weeks | Yes | Yes | Not reported* | Yes | This is an analysis of continuation phase of the active arm.* |
| Merrill 2010 | 34 | 118 | SF36 | BILAG | 12 months | Yes | Yes | Partially* | Yes | In PCS but not MCS.* |
| Kiani *et al*, 2013[37] | 35 | 200 | SF36 | SLEDAI | 24 months | Not reported | Unclear | No | No | |
| Strand *et al*, 2003[38] | 36 | 230 | SF36 | Time to renal flare | 16 weeks | Yes | Yes | Partially | Partially | |
| Merrill 2010 | 37 | 257 | SF36 | BILAG | 52 weeks | Yes, but not SS | Yes | Yes | Yes | |
| Strand *et al*, 2005[41] | 39 | 298 | SF36 | Clinical responder vs non-responder analysis | 6 months | Yes | Yes | Partially | Partially | |
| Cardiel *et al*, 2008[42] | 40 | 317 | SF36 | SLEDAI | 48 weeks | Yes | No | No | No | |
| Wallace *et al*, 2009[43] | 41 | 449 | SF36 | SLEDAI | 52 weeks | Yes, but not SS | Yes | No | Yes | |
| Furie *et al*, 2011[44] | 42 | 819 | SF36 | SLEDAI | 52 weeks | Yes | Yes | No | No | |
| Strand 2014 | 38 | 865 | SF36 | SRI | 52 weeks | Yes | Yes | Partially | Yes | |

Continued

**Table 4** Continued

| Reference | Citation | n | HRQoL Instrument | Clinical disease activity indicator | Time of endpoint assessment (from baseline) | Between-arm clinical improvement | Between-arm HRQoL improvement | Between-arm differences in HRQoL ≥MID | Within-arm HRQoL change from baseline (active arm or pooled) ≥MID | Notes |
|---|---|---|---|---|---|---|---|---|---|---|
| Furie *et al*, 2014[45] | 43 | 1684 | SF36 | SLEDAI | 52 weeks | Yes, but not SS | Yes | Yes | Yes* | Among SRI responders vs non-responders.* |
| Wallace *et al*, 2017[46] | 44 | 183 | SF36 | SRI | 24 weeks | Yes, but not SS | Yes | Partially | Yes | |
| van Vollenhoven *et al*, 2012[47] | 45 | 1684 | SF36 | SRI | 52 weeks | Yes | Yes | No | Yes | |
| Wallace *et al*, 2017[46] | 44 | 183 | EQ5D | SRI | 25 weeks | Yes, but not SS | Yes | No | Yes | |
| Furie *et al*, 2014[45] | 43 | 1684 | FACIT-Fatigue | SRI | 52 weeks | Yes, but not SS | Yes | Yes | Yes | |
| Wallace *et al*, 2017[46] | 44 | 183 | FACIT-Fatigue | SRI | 26 weeks | Yes, but not SS | Yes | No | No | |
| Petri *et al*, 2017[48] | 46 | 547 | FACIT-Fatigue | SRI | 52 weeks | Yes, but not SS | Yes | No | Yes* | Among SRI responders vs non-responders.* |

BILAG, British Isles Lupus Assessment Group Disease Activity Index; EQ5D, EuroQol-5D; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; MCS, mental component summary; MID, minimally important differences; mSLEDAI, modified Systemic Lupus Erythematosus Disease Activity Index; PCS, physical component summary; RCT, randomised controlled trial; SF36, Medical Outcomes Study Short-Form 36; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; SRI, Systemic Lupus Erythematosus Responder Index; SS, statistically significant; partially implies in some domains; HRQoL, health-related quality of life.

health status across all domains except social support, coping, satisfaction with medical care and procreation. Lupus symptom scores and pain/vitality scores differentiated among patients with flare/active disease and those without; however, estimates of effect size were not reported.[72 73] Scores changed significantly in response to longitudinal changes in patient-reported health (across seven domains), physician global assessment (across six domains) and flare (across five domains); however, SRMs have not been reported.[71] No data are currently available on MIDs. The instrument has been validated in several languages, including Chinese, French, Italian, Japanese, Spanish, Tagalog and Turkish.[74–82] The instrument has not been used in an RCT.

### Lupus Impact Tracker

The instrument had good internal consistency and test–retest reliability[83–85]; however, no data are available on floor and ceiling effects. Moderate-strong correlations (r≥0.40) were reported with patient-reported HRQoL measures, and correlations with disease activity, damage and physician global assessment were mostly weak (r≤0.31).[83 84 86] The mean scores differed significantly across dichotomised categories of disease activity, disability, socioeconomic status, age, race, education and marital status[83–86]; estimates of effect size have not been reported. Scores changed significantly in response to longitudinal changes in patient-reported outcomes or disease activity.[83–86] Data on SRMs were contradictory and further research is warranted. One study suggested that the measure was responsive to clinical improvements (SRM=0.69) but not to clinical deteriorations (SRM=0.20).[87] MIDs range from 2 to 4 for clinical deteriorations and an MID of 28 points has been reported for clinical improvements.[83 87] The instrument has been validated in several languages, including German, Italian, Spanish, Swedish and French.[88] The instrument has not been used in an RCT setting.

## DISCUSSION

In this review, we compared the measurement properties of 7 patient-reported HRQoL instruments in English-speaking adult patients with SLE using data from 25 validation studies and 26 drug RCTs. Overall, we found comparable measurement properties between the disease-specific measures, the SF36 and the PROMIS measures, but few measures aside from the SF36 have been incorporated into clinical trials. In general, instruments had good validity but poor-moderate responsiveness to change over time. Cultural adaptability and responsiveness of the PROMIS measures remain to be reported. In RCTs, clinically important improvements were reported in SF36 scores from baseline; however, between-arm differences were frequently non-significant and non-important, implying the SF36 is not responsive to interventions.

Despite the validation of the PROMIS item-bank and the disease-specific instruments in SLE, SF36 frequently has been the only patient-reported outcome in RCTs. Several prior publications have called for standardisation of instruments to measure HRQoL in SLE research to enable comparison between studies and encouraged SF36 use as it is internationally recognised and well-validated across multiple conditions.[89] The 1995 Systemic Lupus International Collaborating Clinics Workshop recommended SF36 for measuring HRQoL in patients with SLE.[16] It was also recommended by Outcome Measures in Rheumatology IV for assessment in RCTs and longitudinal observational studies in SLE.[17] Our findings do not support the use of SF36 as the key measure moving forward and show that the SF36 is not particularly responsive in SLE. Despite its extensive validation, the measure's test–retest reliability and known-group validity using effect size remain to be reported and may provide further insight into the measure's responsiveness. Generally, instruments found to discriminate among clinically distinct groups are also found to be responsive to change.[90] PROMIS measures provided the strongest evidence for known-group validity, indicating they may be more sensitive to change over time; however, this remains to be tested. Our findings also demonstrate that the disease-specific measures had good validity and reliability and were equally or more responsive to change than the SF36. As the field of clinical trials in SLE evolves, guidelines should be revised to encourage use of a broader range of validated HRQoL measures in clinical research to improve study designs. Incorporating disease-specific HRQoL measures as endpoints is also important in providing patient-centric care to improve outcomes pertinent to patients with SLE.[4]

Consistent with a prior review of patient-reported outcomes in lupus clinical trials,[4] data from RCTs that used the SF36 show that longitudinal changes were clinically important regardless of assignment to pharmacological intervention. In contrast, between-arm differences were mostly non-important and non-significant. One interpretation of this finding is that non-pharmacological interventions associated with RCTs (such as routine monitoring of adverse events, improved access to health services, provision of multidisciplinary care, use of background medication, provision of health-related educational material and improved patient–physician dialogue) may have a greater impact on constructs measured by the SF36 than pharmacological interventions that specifically target clinical outcomes. The observation that non-pharmacological approaches can improve HRQoL is supported by prior research[89] and suggests that the combinations of pharmacological and non-pharmacological therapies may have an additive (or perhaps synergistic) effect on improving HRQoL. Selection bias may be another plausible explanation for clinically important improvements in HRQoL among patients assigned to placebo in RCTs. Strict inclusion criteria often mean that patients enrolled in RCTs are generally healthier and better-informed than the general SLE population and more likely to experience further improvements in self-perceived health. Finally, there are observations that the MID determined through anchor-based methods seen in validation studies may differ from MIDs seen in RCTs.[90] As evidence accumulates in RCTs, the observed changes in HRQoL measures based on effective treatments provide a valuable source of data on responsiveness and MIDs. Therefore, it is important that clinical trial literature in SLE is reviewed for older instruments such as the SF36 and synthesised for newer instruments to further support the evidence base on responsiveness and MID for interpreting HRQoL data.

While the validity of this literature review is strengthened by the inclusion of validation studies and RCTs, this study has some limitations. First, an assessment of the quality of the studies identified from the literature search was not conducted, so as not to limit our search. Second, HRQoL measures that had few publications were not prioritised and therefore not included in the indepth review. Third, we did not evaluate measurement properties in non-English-speaking SLE populations.

In conclusion, SLE is a condition associated with high unmet need and considerable burden to patients. SF36, PROMIS, LupusQoL and LIT have the strongest evidence for validity and as such are suitable for use in SLE RCTs; however, few measures aside from the SF36 have been incorporated into clinical trials. SRMs were inconsistent across different anchors and generally poor in all instruments with data for analysis. In RCTs, between-arm differences in SF36 scores were frequently non-significant and non-important. This review highlights the importance of incorporating a broader range of SLE-specific HRQoL measures in RCTs and warrants further research that focuses on longitudinal responsiveness and cultural adaptability of newer instruments such as the PROMIS item-bank.

**Contributors** ZI is the primary author responsible for planning, conduct and reporting of the work described in the article. JG is an independent reviewer and has contributed to the conduct and reporting of the work described in the article. PPK and JY have provided significant input into the planning and reviewed the overall work.

**Funding** JY is supported by the Robert L Kroc Chair in Rheumatic and Connective Tissue Diseases, NIH/NIAMS P30 AR070155 and the Russell/Engleman Medical Research Center for Arthritis. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Competing interests** None declared.

**Patient consent** Not required.

**Provenance and peer review** Commissioned; externally peer reviewed.

**Data statement** All clinical trial data were retrieved from published manuscripts. Patient-level data or unpublished data were not obtained for this review. As such no further data is available for sharing beyond the content of this review.

## REFERENCES


1. Archenholtz B, Burckhardt CS, Segesten K. Quality of life of women with systemic lupus erythematosus or rheumatoid arthritis: domains of importance and dissatisfaction. *Qual Life Res* 1999;8:411–6.
2. Jolly M, Katz P. Systemic lupus erythematosus: The promise of PROMIS - is it ready for prime time in SLE? *Nat Rev Rheumatol* 2017;13:453–4.
3. Mahieu M, Yount S, Ramsey-Goldman R. Patient-reported outcomes in systemic lupus erythematosus. *Rheum Dis Clin North Am* 2016;42:253–63.
4. Annapureddy N, Devilliers H, Jolly M. Patient-reported outcomes in lupus clinical trials with biologics. *Lupus* 2016;25:1111–21.
5. McElhone K, Abbott J, Teh LS. A review of health related quality of life in systemic lupus erythematosus. *Lupus* 2006;15:633–43.
6. Gordon C, Bertsias G, Ioannidis JP, *et al*. EULAR points to consider for conducting clinical trials in systemic lupus erythematosus. *Ann Rheum Dis* 2009;68:470–6.
7. Castelino M, Abbott J, McElhone K, *et al*. Comparison of the psychometric properties of health-related quality of life measures used in adults with systemic lupus erythematosus: a review of the literature. *Rheumatology* 2013;52:684–96.
8. Oude Voshaar MA, ten Klooster PM, Taal E, *et al*. Measurement properties of physical function scales validated for use in patients with rheumatoid arthritis: a systematic review of the literature. *Health Qual Life Outcomes* 2011;9:99.
9. Arat S, Lenaerts JL, De Langhe E, *et al*. Illness representations of systemic lupus erythematosus and systemic sclerosis: a comparison of patients, their rheumatologists and their general practitioners. *Lupus Sci Med* 2017;4:e000232.
10. Terwee CB, Bot SD, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
11. Guyatt GH, Deyo RA, Charlson M, *et al*. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol* 1989;42:403–8.
12. Nantes SG, Strand V, Su J, *et al*. Comparison of the sensitivity to change of the 36-item short form health survey and the lupus quality of life measure using various definitions of minimum clinically important differences in patients with active systemic lupus erythematosus. *Arthritis Care Res* 2018;70:125–33.
13. Stoll T, Gordon C, Seifert B, *et al*. Consistency and validity of patient administered assessment of quality of life by the MOS SF-36; its association with disease activity and damage in patients with systemic lupus erythematosus. *J Rheumatol* 1997;24:1608–14.
14. Thumboo J, Fong KY, Ng TP, *et al*. Validation of the MOS SF-36 for quality of life assessment of patients with systemic lupus erythematosus in Singapore. *J Rheumatol* 1999;26:97–102.
15. Fortin PR, Abrahamowicz M, Neville C, *et al*. Impact of disease activity and cumulative damage on the health of lupus patients. *Lupus* 1998;7:101–7.
16. Saba J, Quinet RJ, Davis WE, *et al*. Inverse correlation of each functional status scale of the SF-36 with degree of disease activity in systemic lupus erythematosus (m-SLAM). *Joint Bone Spine* 2003;70:348–51.
17. McElhone K, Abbott J, Sutton C, *et al*. Sensitivity to change and minimal important differences of the lupusqol in patients with systemic lupus erythematosus. *Arthritis Care Res* 2016;68:1505–13.
18. Colangelo KJ, Pope JE, Peschken C. The minimally important difference for patient reported outcomes in systemic lupus erythematosus including the HAQ-DI, pain, fatigue, and SF-36. *J Rheumatol* 2009;36:2231–7.
19. Thumboo J, Fong KY, Chan SP, *et al*. A prospective study of factors affecting quality of life in systemic lupus erythematosus. *J Rheumatol* 2000;27:1414–20.
20. Thumboo J, Fong KY, Ng TP, *et al*. Initial construct cross-cultural validation of the Short Form 36 for quality of life assessment of systemic lupus erythematosus patients in Singapore. *Ann Acad Med Singapore* 1997;26:282–4.
21. Devilliers H, Amoura Z, Besancenot JF, *et al*. RespoNsiveness of the 36-item short form health survey and the lupus quality of life questionnaire in sle. *Rheumatology* 2015;54:940–9.
22. Yilmaz-Oner S, Oner C, Dogukan FM, *et al*. Health-related quality of life assessed by LupusQoL questionnaire and SF-36 in Turkish patients with systemic lupus erythematosus. *Clin Rheumatol* 2016;35:617–22.
23. Galbraith L, Manns B, Hemmelgarn B, *et al*. The Steroids In the Maintenance of remission of Proliferative Lupus nephritis (SIMPL) pilot trial. *Can J Kidney Health Dis* 2014;1:30.
24. Uppal SS, Hayat SJ, Raghupathy R. Efficacy and safety of infliximab in active SLE: a pilot study. *Lupus* 2009;18:690–7.
25. Nordmark G, Bengtsson C, Larsson A, *et al*. Effects of dehydroepiandrosterone supplement on health-related quality of life in glucocorticoid treated female patients with systemic lupus erythematosus. *Autoimmunity* 2005;38:531–40.
26. Dussán KB, Magder L, Brodsky RA, *et al*. High dose cyclophosphamide performs better than monthly dose cyclophosphamide in quality of life measures. *Lupus* 2008;17:1079–85.
27. Arriens C, Hynan LS, Lerman RH, *et al*. Placebo-controlled randomized clinical trial of fish oil's impact on fatigue, quality of life, and disease activity in Systemic Lupus Erythematosus. *Nutr J* 2015;14:82.
28. Danowski A, Magder L, Petri M. Flares in lupus: Outcome assessment trial (float), a comparison between oral methylprednisolone and intramuscular triamcinolone. *J Rheumatol* 2006;33:57–60.
29. Hartkamp A, Geenen R, Godaert GL, *et al*. Effects of dehydroepiandrosterone on fatigue and well-being in women with quiescent systemic lupus erythematosus: a randomised controlled trial. *Ann Rheum Dis* 2010;69:1144–7.
30. Furie R, Stohl W, Ginzler EM, *et al*. Biologic activity and safety of belimumab, a neutralizing anti-B-lymphocyte stimulator (BLyS) monoclonal antibody: a phase I trial in patients with systemic lupus erythematosus. *Arthritis Res Ther* 2008;10:R109.
31. Gordon C, Wallace DJ, Shinada S, *et al*. Testosterone patches in the management of patients with mild/moderate systemic lupus erythematosus. *Rheumatology* 2008;47:334–8.
32. Grootscholten C, Snoek FJ, Bijl M, *et al*. Health-related quality of life and treatment burden in patients with proliferative lupus nephritis treated with cyclophosphamide or azathioprine/ methylprednisolone in a randomized controlled trial. *J Rheumatol* 2007;34:1699–707.
33. Griffiths B, Emery P, Ryan V, *et al*. The BILAG multi-centre open randomized controlled trial comparing ciclosporin vs azathioprine in patients with severe SLE. *Rheumatology* 2010;49:723–32.
34. Strand V, Levy RA, Cervera R, *et al*. Improvements in health-related quality of life with belimumab, a B-lymphocyte stimulator-specific inhibitor, in patients with autoantibody-positive systemic lupus erythematosus from the randomised controlled BLISS trials. *Ann Rheum Dis* 2014;73:838–44.
35. Wallace DJ, Hobbs K, Clowse ME, *et al*. Long-term safety and efficacy of epratuzumab in the treatment of moderate-to- severe systemic lupus erythematosus: results from an open-label extension study. *Arthritis Care Res* 2016;68:534–43.
36. Merrill JT, Burgos-Vargas R, Westhovens R, *et al*. The efficacy and safety of abatacept in patients with non-life-threatening manifestations of systemic lupus erythematosus: results of a twelve-month, multicenter, exploratory, phase IIb, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2010;62:3077–87.
37. Kiani AN, Strand V, Fang H, *et al*. Predictors of self-reported health-related quality of life in systemic lupus erythematosus. *Rheumatology* 2013;52:1651–7.
38. Strand V, Aranow C, Cardiel MH, *et al*. Improvement in health-related quality of life in systemic lupus erythematosus patients enrolled in a randomized clinical trial comparing LJP 394 treatment with placebo. *Lupus* 2003;12:677–86.
39. Merrill JT, Neuwelt CM, Wallace DJ, *et al*. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. *Arthritis Rheum* 2010;62:222–33.
40. Strand V, Petri M, Kalunian K, *et al*. Epratuzumab for patients with moderate to severe flaring SLE: health-related quality of life outcomes and corticosteroid use in the randomized controlled ALLEVIATE trials and extension study SL0006. *Rheumatology* 2014;53:502–11.
41. Strand V, Crawford B. Improvement in health-related quality of life in patients with SLE following sustained reductions in anti-dsDNA antibodies. *Expert Rev Pharmacoecon Outcomes Res* 2005;5:317–26.
42. Cardiel MH, Tumlin JA, Furie RA, *et al*. Abetimus sodium for renal flare in systemic lupus erythematosus: Results of a randomized, controlled phase III trial. *Arthritis & Rheumatism* 2008;58:2470–80.
43. Wallace DJ, Stohl W, Furie RA, *et al*. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. *Arthritis & Rheumatism* 2009;61:1168–78.
44. Furie R, Petri M, Zamani O, *et al*. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. *Arthritis & Rheumatism* 2011;63:3918–30.

45. Furie R, Petri MA, Strand V, *et al*. Clinical, laboratory and health-related quality of life correlates of Systemic Lupus Erythematosus Responder Index response: a post hoc analysis of the phase 3 belimumab trials. *Lupus Sci Med* 2014;1:e000031.
46. Wallace DJ, Strand V, Merrill JT, *et al*. Efficacy and safety of an interleukin 6 monoclonal antibody for the treatment of systemic lupus erythematosus: a phase II dose-ranging randomised controlled trial. *Ann Rheum Dis* 2017;76:534–42.
47. van Vollenhoven RF, Petri MA, Cervera R, *et al*. Belimumab in the treatment of systemic lupus erythematosus: high disease activity predictors of response. *Ann Rheum Dis* 2012;71:1343–9.
48. Petri MA, Martin RS, Scheinberg MA, *et al*. Assessments of fatigue and disease activity in patients with systemic lupus erythematosus enrolled in the Phase 2 clinical trial with blisibimod. *Lupus* 2017;26:27–37.
49. Katz P, Pedro S, Michaud K. Performance of the patient-reported outcomes measurement information system 29-item profile in rheumatoid arthritis, osteoarthritis, fibromyalgia, and systemic lupus erythematosus. *Arthritis Care Res* 2017;69:1312–21.
50. Kasturi S, Szymonifka J, Burket JC, *et al*. Validity and reliability of patient reported outcomes measurement information system computerized adaptive tests in systemic lupus erythematosus. *J Rheumatol* 2017;44:1024–31.
51. Kasturi S, Szymonifka J, Burket JC, *et al*. Feasibility, validity, and reliability of the 10-item patient reported outcomes measurement information system global health short form in outpatients with systemic lupus erythematosus. *J Rheumatol* 2018;45:397–404.
52. Mahieu MA, Ahn GE, Chmiel JS, *et al*. Fatigue, patient reported outcomes, and objective measurement of physical activity in systemic lupus erythematosus. *Lupus* 2016;25:1190–9.
53. Lai J-S, Beaumont JL, Jensen SE, *et al*. An evaluation of health-related quality of life in patients with systemic lupus erythematosus using PROMIS and Neuro-QoL. *Clin Rheumatol* 2017;36:555–62.
54. Lai JS, Beaumont JL, Ogale S, *et al*. Validation of the functional assessment of chronic illness therapy-fatigue scale in patients with moderately to severely active systemic lupus erythematosus, participating in a clinical trial. *J Rheumatol* 2011;38:672–9.
55. Aggarwal R, Wilke CT, Pickard AS, *et al*. Psychometric Properties of the EuroQol-5D and Short Form-6D in Patients with Systemic Lupus Erythematosus. *J Rheumatol* 2009;36:1209–16.
56. Wang SL, Wu B, Zhu LA, *et al*. Construct and criterion validity of the Euro Qol-5D in patients with systemic lupus erythematosus. *PLoS One* 2014;9:e98883.
57. Devilliers H, Bonithon-Kopp C, Jolly M. The lupus impact tracker is responsive to changes in clinical activity measured by the systemic lupus erythematosus responder index. *Lupus* 2017;26:396–402.
58. 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 & Rheumatism* 2007;57:972–9.
59. Touma Z, Gladman DD, Ibanez D, *et al*. Is there an advantage over SF-36 with a quality of life measure that is specific to systemic lupus erythematosus? *J Rheumatol* 2011;38:1898–905.
60. Jolly M, Pickard AS, Wilke C, *et al*. Lupus-specific health outcome measure for US patients: the LupusQoL-US version. *Ann Rheum Dis* 2010;69:29–33.
61. McELHONE K, Castelino M, Abbott J, *et al*. The LupusQoL and associations with demographics and clinical measurements in patients with systemic lupus erythematosus. *J Rheumatol* 2010;37:2273–9.
62. Devilliers H, Amoura Z, Besancenot J-F, *et al*. LupusQoL-FR is valid to assess quality of life in patients with systemic lupus erythematosus. *Rheumatology* 2012;51:1906–15.
63. Conti F, Perricone C, Reboldi G, *et al*. Validation of a disease-specific health-related quality of life measure in adult Italian patients with systemic lupus erythematosus: LupusQoL-IT. *Lupus* 2014;23:743–51.
64. González-Rodríguez V, Peralta-Ramírez MI, Navarrete-Navarrete N, *et al*. [Adaptation and validation of the Spanish version of a disease-specific quality of life measure in patients with systemic lupus erythematosus: the Lupus quality of life]. *Med Clin* 2010;134:13–16.
65. Handa ML, McKown S, Gawlicki M. PSY39 linguistic validation of the lupus qol questionnaire into 13 languages. *Value in Health* 2009;12:A137.
67. Hosseini N, Bonakdar ZS, Gholamrezaei A, *et al*. Linguistic validation of the lupusqol for the assessment of quality of life in iranian patients with systemic lupus erythematosus. *Int J Rheumatol* 2014;2014:1–5.
67. Isbagio H. Validity and reliability of lupus quality of life questionnaire in patients with systemic lupus erythematosus in Cipto Mangunkusumo hospital. *Int J Rheum Dis* 2016;19:256.
68. Machado Escobar MA, Yacuzzi MS, Martinez RN, *et al*. Validation of an Argentine version of Lupus Quality of Life questionnaire. *Lupus* 2016;25:1615–22.
69. Pamuk ON, Onat AM, Donmez S, *et al*. Validity and reliability of the Lupus QoL index in Turkish systemic lupus erythematosus patients. *Lupus* 2015;24:816–21.
70. Wang S-li, Wu B, Leng L, *et al*. Validity of LupusQoL-China for the assessment of health related quality of life in Chinese patients with systemic lupus erythematosus. *PLoS One* 2013;8:e63795.
71. Jolly M, Pickard AS, Block JA, *et al*. Disease-specific patient reported outcome tools for systemic lupus erythematosus. *Semin Arthritis Rheum* 2012;42:56–65.
72. Bourre-Tessier J, Clarke AE, Mikolaitis-Preuss RA, *et al*. Cross-cultural validation of a disease-specific patient-reported outcome measure for systemic lupus erythematosus in Canada. *J Rheumatol* 2013;40:1327–33.
73. Azizoddin DR, Weinberg S, Gandhi N. Validation of the LupusPRO version 1.8: an update to a disease-specific patient-reported outcome tool for systemic lupus erythematosus. *Lupus* 2017:961203317739128.
74. Bourré-Tessier J, Clarke AE, Kosinski M, *et al*. The French-Canadian validation of a disease-specific, patient-reported outcome measure for lupus. *Lupus* 2014;23:1452–9.
75. Devilliers H, Jolly M, Samson M, *et al*. Validation of the french version of lupus patient reported outcome (LupusPRO), a disease-specific patient reported outcome for lupus patients. *Arthritis and Rheumatology* 2016;68:1776–8.
76. Inoue M, Shiozawa K, Yoshihara R, *et al*. The Japanese LupusPRO: A cross-cultural validation of an outcome measure for lupus. *Lupus* 2017;26:849–56.
77. Jolly M, Toloza S, Block J, *et al*. Spanish LupusPRO: cross-cultural validation study for lupus. *Lupus* 2013;22:431–6.
78. Josiane BT, Clarke AE, Mikolaitis RA, *et al*. Cross cultural validation of english and french versions of a disease specific patient reported outcome measure for lupus in canada. *Arthritis Rheum* 2012;64:S947.
79. Kaya A, Goker B, Cura ES, *et al*. Turkish lupusPRO: cross-cultural validation study for lupus. *Clin Rheumatol* 2014;33:1079–84.
80. Mok CC, Kosinski M, Ho LY, *et al*. Validation of the LupusPRO in Chinese patients from Hong Kong with systemic lupus erythematosus. *Arthritis Care Res* 2015;67:297–304.
81. Navarra S, Mikolaitis RA, Block JA, *et al*. LupusPRO: Cross cultural validation study for lupus in the philippines. *Arthritis Rheum* 2011;63.
82. Salvador M, Tananguan RM, Jolly M, *et al*. Validation of the filipino version of lupuspro-questionnaire among filipino patients with systemic lupus erythematosus. *Int J Rheum Dis* 2013;16:82.
83. Jolly M, Kosinski M, Garris CP, *et al*. Prospective validation of the lupus impact tracker: a patient-completed tool for clinical practice to evaluate the impact of systemic lupus erythematosus. *Arthritis Rheumatol* 2016;68:1422–31.
84. Jolly M, Garris CP, Mikolaitis RA, *et al*. Development and validation of the Lupus Impact Tracker: a patient-completed tool for clinical practice to assess and monitor the impact of systemic lupus erythematosus. *Arthritis Care Res* 2014;66:1542–50.
85. Antony A, Kandane-Rathnayake RK, Ko T, *et al*. Validation of the Lupus impact tracker in an Australian patient cohort. *Lupus* 2017;26:98–105.
86. Brandt JE, Drenkard C, Kan H, *et al*. External validation of the Lupus impact tracker in a Southeastern US longitudinal cohort with systemic lupus erythematosus. *Arthritis Care Res* 2017;69:842–8.
87. Giangreco D, Devilliers H, Annapureddy N, *et al*. Lupus Impact Tracker is responsive to physician and patient assessed changes in systemic lupus erythematosus. *Lupus* 2015;24:1486–91.
88. Schneider M, Mosca M, Pego-Reigosa J-M, *et al*. Cross-cultural validation of Lupus Impact Tracker in five European clinical practice settings. *Rheumatology* 2017;56:818–28.
89. 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.
90. Revicki D, Hays RD, Cella D, *et al*. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.