




Generation of evidence supporting the content validity of SF-36, FACIT-F, and LupusQoL, and novel patient-reported symptom items for use in patients with systemic lupus erythematosus (SLE) and SLE with lupus nephritis (LN)

Rebecca Williams-Hall,¹ Pamela Berry,² Nicola Williamson ,¹ Melissa Barclay,¹ Anna Roberts,¹ Adam Gater,¹ Chloe Tolley,¹ Helena Bradley,¹ Amy Ward,¹ Elizabeth Hsia,^{3,4} Qing Zuraw,³ Patricia DeLong,² Zahi Touma ,⁵ Vibeke Strand ⁶

To cite: Williams-Hall R, Berry P, Williamson N, *et al.* Generation of evidence supporting the content validity of SF-36, FACIT-F, and LupusQoL, and novel patient-reported symptom items for use in patients with systemic lupus erythematosus (SLE) and SLE with lupus nephritis (LN). *Lupus Science & Medicine* 2022;**9**:e000712. doi:10.1136/lupus-2022-000712

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/lupus-2022-000712>).

Received 13 April 2022
Accepted 1 August 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Nicola Williamson; nicola.williamson@adelphivalues.com

ABSTRACT

Objective SLE and lupus nephritis (LN) have significant impacts on the health-related quality of life of patients living with the condition, which are important to capture from the patient's perspective using patient-reported outcomes (PROs). The objectives of this study were to evaluate the content validity of PROs commonly used in SLE and LN (36-Item Short Form Health Survey (SF-36), Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) and Lupus Quality of Life (LupusQoL), as well as novel PRO symptom severity items measuring skin rash, joint pain, joint stiffness and swelling of the legs and/or feet, in both populations.

Methods Qualitative, semi-structured, cognitive interviews were conducted with 48 participants (SLE=28, LN=20). Understanding and relevance of symptom and impact PRO concepts from existing PROs were assessed, alongside novel PRO symptom severity items with different recall periods (24 hours vs 7 days) and response scales (Numerical Rating Scale (NRS) vs Verbal Rating Scale). Interviews were conducted in multiple rounds to allow for modifications to the novel PRO items. Analysis of verbatim interview transcripts was performed.

Results Symptom and impact concepts assessed by the SF-36, FACIT-F, and LupusQoL were well understood by both participants with SLE and LN (≥90.0%), with most considered relevant by over half of the participants asked (≥51.9%). All participants asked (100%) understood the novel PRO symptom severity items, and the majority (≥90.0%) considered the symptoms relevant. Minor modifications to the novel PRO items were made between rounds to improve clarity based on participant feedback. The selected 7-day recall period and NRS in the final iteration of the PRO items were understood and relevant. No differences in interview findings between the SLE and LN samples were identified.

Conclusions Findings provide evidence of content validity for concepts assessed by the SF-36, FACIT-F, LupusQoL and the novel PRO symptom severity items, supporting use

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Patient-reported outcomes (PROs) including the 36-Item Short Form Health Survey (SF-36), Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) and Lupus Quality of Life (LupusQoL) are commonly used as supportive endpoints in SLE trials to provide unique insights into patient disease experiences and associated functioning that cannot be captured by physician-reported disease activity scores.

WHAT THIS STUDY ADDS

- ⇒ It is important that PROs assess concepts that are relevant to members of the target population and in a manner that is consistently understood by respondents. This study provides content validity evidence of symptom and health-related quality of life (HRQoL) impact concepts assessed by the SF-36, FACIT-F and LupusQoL PROs in lupus nephritis (LN), and further content validity evidence regarding the use of these measures in SLE.
- ⇒ This study adds to the limited qualitative literature exploring the key symptom and HRQoL impact concepts important to patients with LN, confirming overlap in experiences and appropriateness of assessing the same concepts already identified by qualitative research in SLE.
- ⇒ The study also provides content validity evidence for novel PRO symptom severity items designed to assess important concepts (ie, skin rash, joint pain, joint stiffness, and swelling of the legs and/or feet) that are not currently assessed by measures commonly used in both SLE and LN.

of these PROs to comprehensively assess disease impact in future SLE and LN clinical trials.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Findings support the use of existing PROs as well as novel PRO symptom severity items to comprehensively assess disease impact in patients with SLE and LN, for use in clinical trials as supportive endpoints for assessment of treatment efficacy.

INTRODUCTION

SLE is a chronic autoimmune disease characterised by the presence of autoreactive B cells and autoantibodies that result in systemic inflammation and a relapsing, remitting autoimmune disease with multiorgan manifestations.¹ The prevalence of SLE in the USA is estimated to be approximately 450 000 cases.² The clinical presentation of SLE is heterogenous, and patients may experience multiorgan manifestations, which can range from mild to life-threatening. The most common symptoms of SLE are constitutional (eg, fatigue, fever and loss of appetite/weight loss), musculoskeletal (eg, joint/muscle pain, stiffness and swelling) and mucocutaneous (eg, skin rash, oral/nasal ulcers and hair loss).³

Lupus nephritis (LN) is a common and severe complication of SLE which occurs in as many as 50% of patients with SLE.⁴ LN is more common in men than in women,⁵ and often also in Hispanic, African-American and Asian individuals.^{6–8} In addition to SLE symptoms, patients with LN may also complain of symptoms including fatigue; foamy, dark and/or bloody urine; swelling in legs; weight gain; and increased nocturnal urination. LN is a significant risk factor for morbidity and mortality with end-stage renal disease occurring in 10.0%–30.0% of patients with LN. Therefore, SLE and LN have significant impacts on patients' health-related quality of life (HRQoL) including physical functioning, ability to carry out activities of daily living, emotional well-being, work, and participation in family, social and leisure activities.^{9–18}

Capturing information on the symptom presentation and HRQoL impact from the patient perspective provides a unique and clinically relevant view of a medical condition. Accordingly, patient-reported outcomes (PROs) are increasingly recognised as an integral component for the assessment of SLE in randomised controlled trials (RCTs), as they measure additional impacts of disease that cannot be assessed by physician-reported disease activity scores.^{19,20} The US Food and Drug Administration (FDA) Guidance for Industry on PROs and Patient-Focused Drug Development outline best practices for developing and evaluating PROs in a specific context of use to evaluate treatment benefit and ultimately support product label claims. As a first step, it is critical that qualitative evidence is generated in the target population to identify disease-defining concepts that are important and relevant to patients.^{21,22} To effectively evaluate treatment benefit, it is imperative that PROs are fit-for-population (patients with SLE and LN) and purpose, and have strong evidence of content validity (ie, that the content of the measure is appropriate and comprehensive relative to its intended measurement concept and context of use).²¹

The Outcome Measures in Rheumatology (OMERACT) international consensus effort (1998) recommended the assessment of HRQoL using both generic and disease-specific measures as one of five core domains to be assessed in all RCTs and longitudinal observational studies in SLE.²³ Since the publication of OMERACT's recommendations, the majority of RCTs have included one or several PROs. The Medical Outcomes 36-Item Short Form Health Survey (SF-36),²⁴ Lupus Quality of Life (LupusQoL) questionnaire²⁵ and Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) scale²⁶ are generic-specific and disease-specific measures of HRQoL and fatigue (respectively) and commonly used as supportive endpoints in SLE trials.²⁷ Documented evidence regarding the content validity and psychometric measurement properties of these measures supports their suitability for use in SLE as supportive trial endpoints.^{28–31}

Although prior research in SLE and LN has demonstrated that these three PROs assess the majority of key disease impact domains,²⁸ their assessment is either limited to select symptoms (eg, tiredness/fatigue), is non-specific (ie, generalised pain), or focuses on the HRQoL impacts of specific symptoms rather than symptom severity.^{21,32} A qualitative literature and online blog/forum review and qualitative concept elicitation (CE) interviews with patients with SLE and LN exploring the lived experience of each condition identified key symptoms (ie, skin rash, joint pain, joint stiffness and swelling of the legs/feet) and symptom properties (ie, severity) that were not specifically assessed by the SF-36, FACIT-F, and LupusQoL. Therefore, despite each PRO containing sufficient and relevant content for its intended purpose (eg, assessment of the impact of disease (SF-36), impact of SLE on HRQoL (LupusQoL) and the severity and impact of fatigue (FACIT-F)), novel PRO items (questions) assessing the severity of key symptoms have been developed to address measurement gaps in regard to the assessment of disease-specific symptomatology.³³

The overall objective of this study was to conduct qualitative cognitive debriefing (CD) interviews to inform the selection of PROs for use in future SLE and LN RCTs and research studies. The aims of the interviews were first to extend content validity evidence that the symptom and impact concepts assessed by SF-36, FACIT-F, and LupusQoL PROs were consistently understood and relevant to patients with SLE and that no important concepts are missing. Determining whether the same concepts are relevant in a sample of patients with LN was a second aim. A third aim of the interviews was to evaluate the content validity of the four novel PRO symptom severity items developed based on previous qualitative research³³ in a sample of patients with SLE and LN.

METHODS**Study design**

This was a non-interventional, cross-sectional, qualitative interview study in the USA involving CD telephone

- Ethical approval was provided by a centralised independent review board (IRB) prior to any study related activities and fieldwork being conducted.
- A partner recruitment agency worked with rheumatologists and nephrologists to recruit 48 participants (SLE n=28; LN n=20) who met the pre-specified inclusion and exclusion criteria.
- The partner recruitment agency provided eligible patients with the Informed Consent Form (ICF) and Demographic Form for completion via telephone/(e)mail. Recruiting clinicians completed a Case Report form (CRF) to confirm the patients' clinical eligibility.

Conduct of qualitative interviews

- Conduct of telephone 60-minute qualitative interviews across three rounds to evaluate the content validity of the SF-36, FACIT-Fatigue and LupusQoL and/or novel PRO symptom severity items (i.e., skin rash, joint pain, joint stiffness, swelling of legs/feet) with different recall periods (i.e., 24-hour vs 7-day) and response scales (i.e., numerical rating scale vs verbal rating scale).
- All interviews were audio-recorded and transcribed.

Stage 1:

SLE round 1: Combined concept elicitation (CE) and cognitive debriefing (CD) interviews with SLE patients (n=14)

Stage 2:

SLE round 2: CD interviews with SLE patients (n=8)
LN round 1: CD interviews with LN patients (n=10)

Stage 3

SLE round 3: CD interviews with SLE patients (n=8)
LN round 2: CD interviews with LN patients (n=10)

Data analysis

- Qualitative analysis of verbatim interview transcripts using Atlas.ti software.
- Framework analysis was conducted whereby codes were assigned to each concept (existing PRO measure concepts) and item, instructions, response options and recall period (novel PRO symptom severity items) to indicate understanding, relevance and suggested item changes where relevant. As interviews were conducted in rounds, feedback was used to inform modifications to the novel PRO symptom severity items for further testing in subsequent rounds.

Figure 1 Overview of study design. FACIT-F, Functional Assessment of Chronic Illness Therapy–Fatigue; LN, lupus nephritis; LupusQoL, Lupus Quality of Life; PRO, patient-reported outcome; SF-36, 36-Item Short Form Health Survey.

interviews conducted in multiple rounds between February 2019 and December 2019 with a total of 48 participants (SLE=28, LN=20; see [figure 1](#)).

Interviews with participants with SLE were conducted over three rounds, with one round of combined CE/CD interviews and two rounds of purely CD interviews. CD interviews with participants with LN were conducted across two rounds. All interviews included exploration of existing symptom and HRQoL impact concepts assessed by the SF-36, FACIT-F and LupusQoL for relevance and comprehensiveness, while interview rounds 2 and 3 included debriefing of the novel PRO symptom severity items assessing skin rash, joint pain, joint stiffness, and swelling of the legs and/or feet. These qualitative interviews are an industry-standard method to evaluate the content validity of a PRO measure.^{21 34} Findings from the CE component of the SLE interviews (stage 1) are described in a separate publication.³³

To ensure participants were provided with a clear explanation of the study and what was involved, all participant-facing study documents were developed to be understood by individuals with a range of education levels including

low literacy. Participants were assigned a unique identification code and participant names and other identifiable information was removed from all transcripts and other documents.

Study sample

A partner recruitment agency assisted with the recruitment of participants from geographically diverse locations in the USA (New Orleans, Louisiana; Chicago, Illinois; St. Louis, Missouri; Pittsburgh, Pennsylvania; Baltimore, Maryland) via referring rheumatologists and nephrologists.

To be eligible for the study, participants were required to be ≥ 18 years of age and meet the classification criteria of SLE in accordance with Systemic Lupus International Collaborating Clinics classification³⁵ or the American College of Rheumatology criteria.³⁶ To be eligible for the LN sample, participants were also required to have a kidney biopsy documentation within the last 5 years of proliferative nephritis according to the International Society of Nephrology/Renal Pathology Society: class III or class IV (mixed class III/ IV or class IV/V permitted)

and receive standard of care (ie, glucocorticoids and immunosuppressants) determined by the treating physician within the 12 months prior to screening.

Participants were excluded if they were untreated (ie, not in need of immunosuppressive treatment), had widespread pain or fatigue due to fibromyalgia, were currently receiving dialysis treatment or had a previous or planned kidney transplant. Participants were excluded from the SLE sample if they had a diagnosis of cutaneous lupus only.

Age, sex, race, ethnicity, employment status, disease duration, cutaneous subtype (acute, chronic or subacute cutaneous lupus erythematosus (ACLE), or no cutaneous manifestations; SLE sample only) and LN classification (LN sample only) sampling quotas were employed to ensure recruitment of participants with diverse demographic and clinical characteristics (see for online supplemental material table 1 specific targets).

CD interview procedure

At the beginning of the interview, the interviewer explained the aims and purpose of the research, the interview structure, the interviewer's role in the study and the company they work for. The interviewer also outlined the participant's right to withdraw from the study and adverse event reporting procedures, and that the interview would be audio-recorded. Participants were allowed to ask any questions about their participation before, during and following the interview. All interviews were approximately 60 min in duration, conducted via telephone by trained and experienced female qualitative interviewers (authors AW and HB), and were audio-recorded. All participants completed the telephone interview from their home. All interviews involved the participant and the interviewer only.

During the interview, for the debriefing of concepts from existing PROs, participants were asked detailed questions about their understanding of each symptom and HRQoL impact concept and its relevance to their experience of SLE or LN. The five symptom concepts debriefed as assessed in the SF-36, FACIT-F and LupusQoL included pain, tiredness/fatigue, weight gain, hair loss and nausea, while the 16 impact concepts included physical activities, self-care, sleep, emotional functioning, appearance, social activities, family relationships, friend relationships, sexual relationships, work/job, concentration, usual/everyday activities, organising life, treatment, overall health and illness.

For the debriefing of the novel PRO symptom severity items, a 'think-aloud' process was employed whereby participants read and responded to the novel items out aloud, providing thoughts and opinions on each item and instruction (item wording and properties debriefed can be found in online supplemental material table 2).³⁷ Follow-up questioning explored participant understanding and relevance of the item wording, instructions, different recall periods (ie, 24 hours vs 7 days), and response scales (ie, Numerical Rating Scale (NRS)

vs Verbal Rating Scale (VRS)). Interviews were conducted in two rounds to inform modifications. Based on feedback from round 1 of the CD interviews, the instructions, items, response options and item wording where applicable were modified for clarity and retested in a subsequent interview round.

The final novel PRO symptom severity items underwent translatability assessment to ensure linguistic and cultural applicability and appropriateness in Chinese (China, Hong Kong and Taiwan), Spanish (Argentina, Mexico and US Spanish), French, German, Italian, Polish, Russian, Thai and Ukrainian.

Qualitative data analysis

All interview transcripts were pseudonymised and transcribed verbatim. Verbatim transcripts were qualitatively analysed by two members of the Adelphi Values research team using the computer-assisted software Atlas.ti.³⁸ A coding tree was developed prior to analysis of interviews to structure how concepts and subconcepts would be coded (see online supplemental material figure 1). Specific coding schemes were then derived for the interview transcripts using the first two transcripts. Codes were revised and new codes were added iteratively throughout the analysis following the review of subsequent transcripts. Previously coded transcripts were reviewed, and the project leader ensured consistency of coding across all transcripts, after which further revision of codes or new codes were added if relevant. Framework analysis was conducted whereby dichotomous codes were assigned to each item/concept, instruction, response option(s) and recall period(s) to indicate whether it was relevant/not relevant or understood/not understood, where applicable. Codes were also used to indicate whether the novel PRO symptom severity items were easy/difficult to complete, why a specific response option was chosen and how the content of each item applied to the patient experience, as well as to provide suggestions for item, instruction, response option wording or formatting changes and general feedback on items.³⁹ All findings relating to major and minor themes were analysed and reported.

RESULTS

Sample characteristics

Forty-eight participants (SLE=28, LN=20) participated in the study and no participants dropped out during the study. Across both samples, the majority of the sampling quotas were met; however, some quotas for specific categories were not met. For the SLE sample, quotas for ethnicity (Hispanic or Latino) and race (Asian and multiracial) fell below target numbers. However, participants in these categories were still represented in the SLE subsample. For the LN subsample, the only quotas that were not met were also for race (Asian and multiracial).

The demographic and clinical characteristics of each subsample were similar (table 1). The mean age of the overall sample was 42.3 years old (range=22–79); there

Table 1 Demographic and clinical characteristics of the SLE (n=28) and LN (n=20) samples

Characteristic	SLE participants (n=28)	LN participants (n=20)	Total (N=48)
Demographic characteristics			
Age (years)			
Mean (range)	40.1 (22–67)	45.4 (23–79)	42.3 (22–79)
18–30	27.8 (22–30)	26.8 (23–30)	27.4 (22–30)
31–50	38.0 (31–46)	42.4 (32–49)	39.8 (31–49)
51+	55.6 (51–67)	59.1 (51–79)	56.6 (51–79)
Sex, n (%)			
Female	23 (82.1)	17 (85.0)	40 (83.3)
Male	5 (17.9)	3 (15.0)	8 (16.7)
Ethnicity, n (%)			
Non-Hispanic or non-Latino	23 (82.1)	18 (90.0)	41 (85.4)
Hispanic or Latino	5 (17.9)	2 (10.0)	7 (14.6)
Race, n (%)			
Black African/American	13 (46.4)	11 (55.0)	24 (50.0)
White/Caucasian	11 (39.3)	8 (40.0)	19 (39.6)
Multiracial	1 (3.6)	0	1 (2.1)
Asian	1 (3.6)	0	1 (2.1)
Other	2 (7.1)*	1 (5.0)†	3 (6.2)
Work status, n (%)			
Working full-time	14 (50.0)	4 (20.0)	18 (37.5)
Not working due to condition	4 (14.3)	6 (30.0)	10 (20.8)
Working part-time	3 (10.7)	4 (20.0)	7 (14.6)
Full-time homemaker	4 (14.3)	2 (10.0)	6 (12.5)
Retired	1 (3.6)	2 (10.0)	3 (6.3)
Student	1 (3.6)	0	1 (2.1)
Looking for work	0	1 (5.0)	1 (2.1)
Other	1 (3.6)‡	1 (5.0)§	2 (4.2)
Clinical characteristics			
Time since SLE/LN diagnosis (years), n (%)¶			
<2	4 (14.3)	4 (20.0)	8 (16.7)
≥2–5	10 (35.7)	8 (40.0)	18 (37.5)
6–10	6 (21.4)	3 (15.0)	9 (18.7)
11+	8 (28.6)	5 (25.0)	13 (27.1)
Duration of SLE symptoms prior to diagnosis (years), n (%)			
Less than 5	26 (92.9)	14 (70.0)	40 (83.3)
6–10	2 (7.1)	3 (15.0)	5 (10.4)
16 and over	0	1 (5.0)	1 (2.1)
N/A	0	2 (10.0)	2 (4.2)
Duration of LN symptoms prior to diagnosis (years), n (%)			
Less than 5	–	16 (80.0)	16 (33.3)
6–10	–	3 (15.0)	3 (6.2)
16 and over	–	1 (5.0)	1 (2.1)
SLE cutaneous subtype, n (%)			
ACLE	10 (35.7)	3 (15.0)	13 (27.1)
No cutaneous manifestations	9 (32.1)	10 (50.0)	19 (39.6)
SCLE	5 (17.9)	2 (10.0)	7 (14.6)
CCLE	4 (14.3)	5 (25.0)	9 (18.7)

Continued

Table 1 Continued

Characteristic	SLE participants (n=28)	LN participants (n=20)	Total (N=48)
Class of LN diagnosis, n (%)			
III	–	8 (40.0)	8 (16.7)
IV	–	8 (40.0)	8 (16.7)
IV/V	–	4 (20.0)	4 (8.3)

*Hispanic, Puerto Rico.

†Mexican.

‡Reported 'disability' with no further information.

§Commencing school.

¶For three participants, only their diagnosis of SLE/LN occurred on the same day.

ACLE, acute cutaneous lupus erythematosus; CCLE, Chronic Cutaneous Lupus Erythematosus; LN, lupus nephritis; N/A, not applicable; SCLE, Subacute Cutaneous Lupus Erythematosus.

were more women (n=40/48, 83.3%) than men; and most participants interviewed were Black/African-American (n=24/48, 50.0%). Most participants had been clinically diagnosed with SLE or LN between 2 years and 5 years (n=18/48, 37.5%) or 11 years or more years (n=13/48, 27.1%) prior to screening. Most participants had been experiencing SLE symptoms for less than 5 years prior to diagnosis (n=40/48, 83.3%). There were more participants with SLE with the ACLE subtype (n=10/28, 35.7%), while half of participants with LN had no cutaneous manifestations (n=10/20, 50.0%). Participants with LN most commonly had class III (n=8/20, 40.0%) or IV (n=8/20, 40.0%) LN, and most participants had been experiencing LN symptoms for less than 5 years prior to diagnosis (n=16/20, 80.0%).

Debriefing of existing pros

The five symptom concepts (ie, pain, tiredness/fatigue, weight gain, hair loss and nausea) assessed by the SF-36, FACIT-F and LupusQoL measures were each understood by $\geq 90.0\%$ of participants with SLE and all participants with LN (100%) asked. Notably, for one participant, it was unclear if hair loss was understood as an unclear response was provided when asked (SLE=1/26, 3.8%) and two participants did not demonstrate a clear understanding of the symptom nausea (SLE=2/20, 10.0%); one could not explain the term and the other incorrectly described it as tiredness.

The 16 HRQoL impact concepts assessed were each understood by $\geq 96.3\%$ of participants with SLE and $\geq 93.8\%$ of participants with LN asked. Only four impact concepts were not understood by all participants asked, which included 'appearance', as participants provided an unclear response when asked (SLE=1/27, 3.7%) or misunderstood the concept to refer to support provided by family (LN=1/20, 5.0%); 'social activities', as one participant misheard the concept to be 'facial' activities (SLE=1/28, 3.6%); 'self-care activities' (SLE=1/28, 3.6%) and 'organising life' (LN=1/16, 6.2%), as the participants stated that they did not understand the given concept but did not provide further details.

The five symptom concepts were reported to be relevant by the majority of participants with SLE and LN

asked ($\geq 55.0\%$, [figure 2](#)). Fifteen of the 16 HRQoL impact concepts were each reported to be relevant to at least half of the participants with SLE ($\geq 51.9\%$) and participants with LN ($\geq 52.6\%$, [figure 2](#)) asked. The HRQoL impact concept of 'friend relationships' was reported as not relevant by slightly more participants with SLE (55.6%) and LN (52.9%) asked. Participants mostly cited having ample support from their friends as the reason for the lack of relevance of this concept. For both subsamples, at least 10 or more of the 15 impacts were relevant to and experienced by $\geq 70.0\%$ of participants asked. Overall, there were no notable differences between the two subsamples in relation to the understanding and relevance of the symptom and HRQoL impact concepts.

Debriefing of novel PRO symptom severity items

The wording of the novel PRO items was understood consistently by all participants with SLE and LN (100%), and the symptoms assessed were relevant to most ($\geq 90.0\%$, [figure 3](#)). No notable differences were identified between the two participant subsamples in terms of understanding or symptom relevance.

...lupus participants often have joint pain and they're looking to find out how bad the pain is—has been for you in the last day. (Participant in their 50s with LN)

Participants most commonly described experiencing skin rashes on their arms (n=16; SLE=9, LN=7) and face (n=14; SLE=7, LN=7). Joint pain was mostly located in the knees (n=14; SLE=6, LN=8), while descriptions of joint stiffness locations varied, including those in the knees (LN=8), legs (SLE=7), hands (LN=6), and arms (SLE=4). Lastly, participants most commonly described swelling to occur in the feet (n=14; SLE=10, LN=4), legs (n=14; SLE=7, LN=7), and ankles (n=10; SLE=6, LN=4).

Uh, the rashes really are on my legs. They're on my arms. They're really all over on my face. Um, uh, basically they're pretty much all over on my neck and on my arms. (Participant in their 60s with SLE)

In addition, given the potential overlap of the joint pain and joint stiffness concepts, the interviews aimed to explore whether participants considered them to be

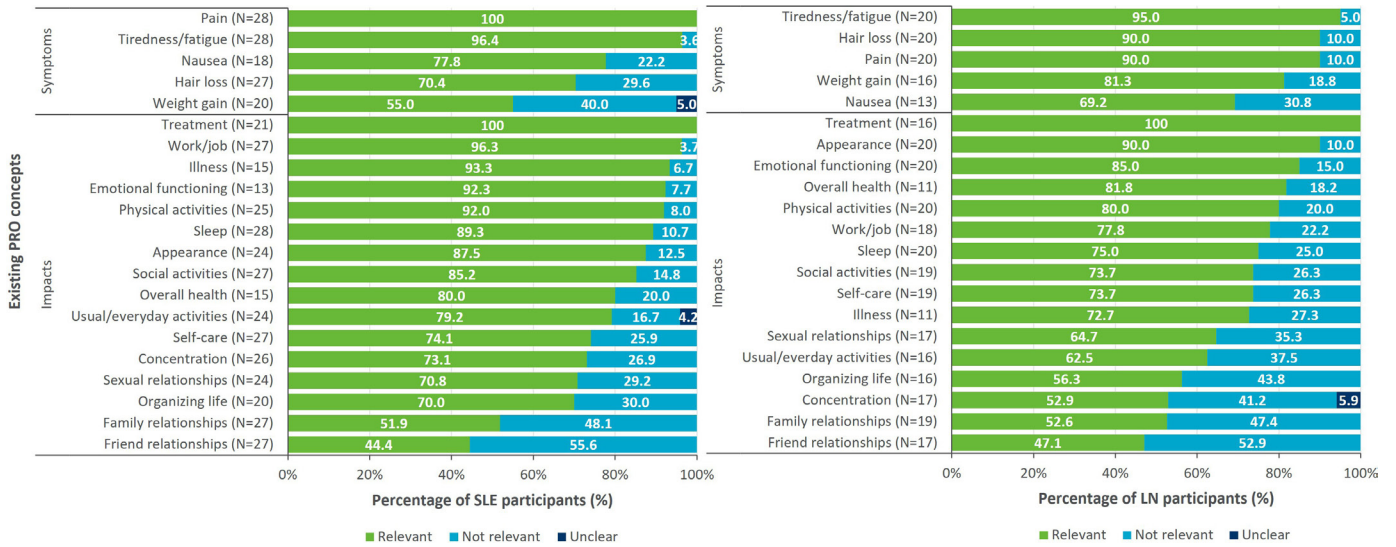


Figure 2 Relevance of symptom and impact concepts assessed by the SF-36, FACIT-F and LupusQoL. FACIT-F, Functional Assessment of Chronic Illness Therapy–Fatigue; LN, lupus nephritis; LupusQoL, Lupus Quality of Life; SF-36, 36-Item Short Form Health Survey.

different conceptings warranting individual assessment or the same concept warranting the inclusion of one item only. Most participants (80.6%; SLE=13/16, LN=16/20) reported that joint pain and joint stiffness were different concepts, thus supporting their inclusion as two separate items:

I feel like you can, you can have pain and not stiffness or you could have stiffness and not necessarily pain. (Participant in their 30s with SLE)

Instructions, recall periods and response options

All participants asked (100%) demonstrated an understanding of each version of the instructions (V.1.0 and V.2.0), the recall periods (24 hours and 7 days) and response options (NRS V.1.0 and V.2.0 and VRS; see online supplemental material table 3 for further details) evaluated. For participants who reported each symptom as relevant to their experience of either SLE or LN, most had experienced the symptom in the past 24 hours

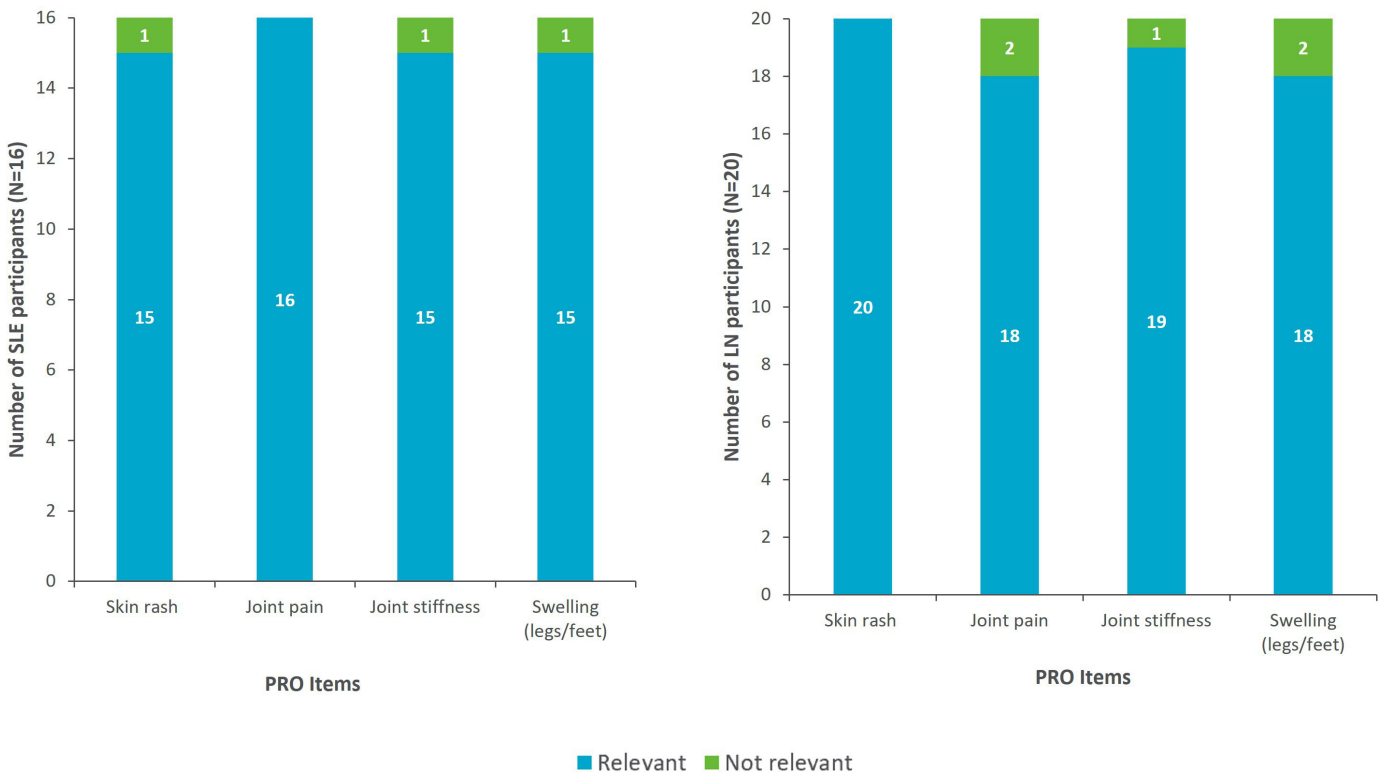


Figure 3 Relevance of novel PRO symptom severity items. PRO, patient-reported outcome.

(SLE $\geq 80.0\%$, LN $\geq 58.8\%$) and 7 days (SLE $\geq 80.0\%$, LN $\geq 60.0\%$; see online supplemental material figure 2 and online supplemental material figure 3 for further information). A small number of participants who reported the symptom as relevant to their condition had not experienced the symptom in the 24 hours prior to the time of the interview. Following evaluation of the recall period and response option data obtained, the 7-day recall period was chosen to provide the optimal balance between recall accuracy and respondent burden, and the NRS response format was chosen as this offered a more granular scale to that of the VRS.

DISCUSSION

The overall objective of this study was to conduct qualitative CD interviews with patients with SLE and LN to inform the selection of PROs for use in future SLE and LN clinical trials. The aims of the interviews were first to extend the content validity evidence of the concepts assessed by SF-36, FACIT-F, and LupusQoL PROs in patients with SLE and, second, to determine whether the same concepts are relevant in patients with LN. The third aim was to evaluate the content validity of four novel PRO symptom severity items³³ in a sample of patients with SLE and LN. Findings are consistent with prior research and provide further evidence supporting the content validity of the symptom and HRQoL impacts assessed by SF-36, FACIT-F and LupusQoL for use in SLE.^{23 27 28 40–42} This study adds to the limited qualitative literature exploring the key symptom and HRQoL impact concepts important to patients with LN, confirming overlap in experiences and appropriateness of assessing the same concepts already identified by qualitative research in SLE.

Content validity evidence was also generated for the novel PRO symptom severity items, supporting the importance of their assessment of the key symptoms in RCTs and observational studies. The final chosen PRO items assess the worst severity of skin rash, joint pain, joint stiffness, and swelling of the legs/feet, using an NRS response scale (0 'no (symptom)' to 10 'severe (symptom)') and a 7-day recall period; see online supplemental material table 4). While all participants understood the NRS and VRS, the NRS was chosen for the final PRO items as it offers greater granularity in responses and therefore greater sensitivity for the assessment of symptom severity. Similarly, despite both the 24-hour and 7-day recall periods being understood by all participants, a 7-day recall period was chosen as more participants had experienced these symptoms intermittently within 7 days and therefore were considered more reflective of the patient experience.

The novel PRO items assess key symptoms (ie, skin rash, joint pain, joint stiffness, and swelling of the legs/feet) identified as not being directly assessed by other PROs commonly used in SLE and LN trials.³³ Specifically, SF-36, LupusQoL and FACIT-F assess general and disease-specific HRQoL and fatigue concepts, respectively, and while they include assessments of some key symptoms

of SLE such as pain and rash, evaluation of these symptoms is generalised and/or limited to the impact of the symptom. For example, LupusQoL assesses the impact of pain on the performance of activities, quality of sleep, mobility and sexual relationships, and the impact of a rash on social situations and feelings of attractiveness. Furthermore, SF-36 assesses generalised pain severity and the impact of pain on the ability to work inside/outside the home. The novel PRO symptom severity items provide greater specificity in the assessment of symptoms and complement the broader concepts assessed by SF-36, FACIT-F, and LupusQoL, providing a comprehensive evaluation of all key symptoms and HRQoL impacts when used for evaluation of the condition and its treatment in the context of an RCT.

Patients with SLE/LN often experience chronic symptoms which demonstrate limited day-to-day variability. Clinical trials in SLE and LN are also lengthy in duration (up to a year in most cases), and participants may be expected to complete assessments frequently during the trial to monitor treatment effect. Taking these factors into consideration, it was determined that a 7-day recall period would be appropriate for the novel PRO symptom severity items and provides an optimal balance between the accuracy of recall and minimisation of respondent burden.⁴³ Additionally, a 7-day recall period, compared with a shorter recall period (eg, 24-hours), provides the opportunity for a holistic assessment of the symptom experience whereby patients' engagement in varied activities (eg, work vs leisure activities) can be captured. The novel PRO items focus on the severity of symptoms during this time frame based on feedback from participants, which indicated that the severity of their symptoms (not the frequency or duration of symptoms) directly affects the impact of their condition on their daily lives. Therefore, assessing changes in the severity of symptoms is an important means of understanding whether a treatment in the context of a clinical trial is having a meaningful benefit to patients. The assessment of those symptoms during the 7-day recall period at their 'worst', which are expected to be salient experiences, is also designed to facilitate recall accuracy. Additionally, this approach is in accordance with the FDA best practice PRO guidance, which typically encourages the assessment of symptom severity.²¹

This was a qualitative study with the limitation of a relatively small sample size in the context of the broader SLE and LN populations generally. However, care was taken to ensure recruitment of a demographic and clinical sample representative of those of the broader disease population, which was achieved via the implementation of recruitment quotas, the majority of which were successfully achieved. Ethnicity and race quotas were narrowly missed (Hispanic/Latino participants, Asian and multiracial participants), meaning that representation is lower than expected in the sample, although there is still some representation in the overall sample. As this study was only conducted with participants from select sites in the

USA, further research with participants from other countries could be conducted to corroborate the findings from this study.

Qualitative research is an iterative process; therefore, there were various iterations of the interview guide throughout the study, and some new questions were introduced to the guide. In addition to associated time constraints of interviews, this meant some questions were only asked in a subset of participants, so sample sizes for certain aspects varied. However, findings were representative of the participants asked and, in most cases, the majority of participants demonstrated understanding and relevance of the concepts. This sample offers insight beyond that presently available in the published literature into the experience of LN, along with content validity evidence for the symptoms and HRQoL impacts assessed by SF-36, FACIT-F, and LupusQoL, as well as novel PRO symptom severity items.

This study provides content validity evidence of the symptom and HRQoL impact concepts assessed by the SF-36, FACIT-F and LupusQoL, and the novel PRO symptom severity items developed to address measurement gaps in SLE and LN. Findings support the use of existing PROs as well as novel PRO symptom severity items to comprehensively assess disease impact in patients with SLE and LN, for use in RCTs as supportive endpoints for assessment of treatment efficacy. Further research is required to establish the measurement properties (particularly reliability and construct validity) of these novel PRO symptom severity items in SLE and LN.

Author affiliations

¹Patient-Centered Outcomes, Adelphi Values, Bollington, UK

²Patient Reported Outcomes, Janssen Global Services LLC, Titusville, Florida, USA

³Immunology Clinical Development, Janssen Research & Development LLC, Spring House, Pennsylvania, USA

⁴Division of Rheumatology, Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

⁵Division of Rheumatology, University of Toronto, Toronto, Ontario, Canada

⁶Division of Immunology/Rheumatology, Stanford University, Palo Alto, California, USA

Acknowledgements The authors wish to acknowledge Kate Burrows for her contribution to the study conception, design and acquisition of data.

Contributors PB, CT, RW-H and AG contributed to the study conception and design and provided supervision. CT, HB and AW contributed to the acquisition of data. RW-H, NW, MB and AR contributed to the data analysis, including interpretation and reporting. VS, ZT, PB, EH, QZ, PDL and ZT also served as clinical and scientific advisors. All authors were involved in drafting the article and/or critically reviewing/ revising it for important intellectual content, and approved the final version to be published. PB is the guarantor.

Funding Research was funded by Janssen Global Services LLC.

Competing interests RW-H, NW, MB, AR, AG, CT, HB and AW are employed by Adelphi Values, a health outcomes research company paid for conducting the research reported in this article. PB and PDL work for and own stocks/shares in Janssen Global Services LLC, which funded the research reported. EH and QZ work for and own stocks in Janssen R&D. ZT received funding from Schroeder Arthritis Institute and Lupus Ontario. VS received funding from numerous clients, including Janssen, which funded this research.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Copernicus Group Independent Review Board (approval number 20183266). All participants provided both verbal and written consent before participating in any study-related activities.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Nicola Williamson <http://orcid.org/0000-0002-9717-8563>

Zahi Touma <http://orcid.org/0000-0001-5177-2076>

Vibeke Strand <http://orcid.org/0000-0003-4978-4072>

REFERENCES

- 1 Tsokos GC. Systemic lupus erythematosus. *N Engl J Med* 2011;365:2110–21.
- 2 Wang Y, Hester L, Lofland J. Prevalence of diagnosed systemic lupus erythematosus (SLE), patient healthcare utilization and characteristics by major health insurance types in the US. *Arthritis & Rheumatology* 2019.
- 3 Cojocaru M, Cojocaru IM, Silosi I, et al. Manifestations of systemic lupus erythematosus. *Maedica* 2011;6:330.
- 4 Liu C-C, Kao AH, Manzi S, et al. Biomarkers in systemic lupus erythematosus: challenges and prospects for the future. *Ther Adv Musculoskelet Dis* 2013;5:210–33.
- 5 Schwartzman-Morris J, Putterman C. Gender differences in the pathogenesis and outcome of lupus and of lupus nephritis. *Clin Dev Immunol* 2012;2012:1–9.
- 6 Korbet SM, Schwartz MM, Evans J, et al. Severe lupus nephritis: racial differences in presentation and outcome. *J Am Soc Nephrol* 2007;18:244–54.
- 7 Tesar V, Hruskova Z. Lupus nephritis: a different disease in European patients? *Kidney Dis* 2015;1:110–8.
- 8 Alarcón GS. Multiethnic lupus cohorts: what have they taught us? *Reumatol Clin* 2011;7:3–6.
- 9 Jump RL, Robinson ME, Armstrong AE, et al. Fatigue in systemic lupus erythematosus: contributions of disease activity, pain, depression, and perceived social support. *J Rheumatol* 2005;32:1699–705.
- 10 McElhone K, Abbott J, Teh LS. A review of health related quality of life in systemic lupus erythematosus. *Lupus* 2006;15:633–43.
- 11 Schmeding A, Schneider M, Fatigue SM. Fatigue, health-related quality of life and other patient-reported outcomes in systemic lupus erythematosus. *Best Pract Res Clin Rheumatol* 2013;27:363–75.
- 12 Katz P, Morris A, Trupin L, et al. Disability in valued life activities among individuals with systemic lupus erythematosus. *Arthritis Rheum* 2008;59:465–73.
- 13 Beckerman NL, Auerbach C, Blanco I. Psychosocial dimensions of SLE: implications for the health care team. *J Multidiscip Healthc* 2011;4:63.
- 14 Almedh K, Carlsten H, Forsblad-d'Elia H. Health-related quality of life in systemic lupus erythematosus and its association with disease and work disability. *Scand J Rheumatol* 2010;39:58–62.
- 15 Hanly JG, O'Keefe AG, Su L, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. *Rheumatology* 2016;55:252–62.
- 16 Dalebout GMN, Berger SP, Broadbent E, et al. Health-related quality of life in patients with systemic lupus erythematosus and proliferative lupus nephritis. *Psychol Health Med* 2011;16:393–404.

- 17 Jolly M, Toloza S, Goker B, *et al.* Disease-specific quality of life in patients with lupus nephritis. *Lupus* 2018;27:257–64.
- 18 Wang Y, Zhao R, Gu C, *et al.* The impact of systemic lupus erythematosus on health-related quality of life assessed using the SF-36: a systematic review and meta-analysis. *Psychol Health Med* 2019;24:978–91.
- 19 Gladman D, Urowitz M, Fortin P, *et al.* Systemic lupus international collaborating clinics conference on assessment of lupus flare and quality of life measures in SLE. systemic lupus international collaborating clinics group. *J Rheumatol* 1996;23:1953–5.
- 20 Strand V, Gladman D, Isenberg D, *et al.* Outcome measures to be used in clinical trials in systemic lupus erythematosus. *J Rheumatol* 1999;26:490–7.
- 21 US Food and Drug Administration (FDA). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims; 2009.
- 22 US Food and Drug Administration (FDA). FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making, 2020. Available: <https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical>
- 23 Strand V, Gladman D, Isenberg D, *et al.* Endpoints: consensus recommendations from OMERACT IV. *Lupus* 2000;9:322–7.
- 24 Ware JE KM, Dewey JE. *How to score version 2 of the SF-36 health survey*. Lincoln, RI: QualityMetric Incorporated, 2000.
- 25 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.
- 26 Yellen SB, Cella DF, Webster K, *et al.* Measuring fatigue and other anemia-related symptoms with the functional assessment of cancer therapy (FACT) measurement system. *J Pain Symptom Manage* 1997;13:63–74.
- 27 Izadi Z, Gandrup J, Katz PP, *et al.* Patient-reported outcome measures for use in clinical trials of SLE: a review. *Lupus Sci Med* 2018;5:e000279.
- 28 Strand V, Simon LS, Meara AS, *et al.* Measurement properties of selected patient-reported outcome measures for use in randomised controlled trials in patients with systemic lupus erythematosus: a systematic review. *Lupus Sci Med* 2020;7:e000373.
- 29 Rendas-Baum R, Baranwal N, Joshi AV, *et al.* Psychometric properties of FACIT-Fatigue in systemic lupus erythematosus: a pooled analysis of three phase 3 randomised, double-blind, parallel-group controlled studies (BLISS-SC, BLISS-52, BLISS-76). *J Patient Rep Outcomes* 2021;5:1–10.
- 30 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.
- 31 Strand V, Berry P, Lin X, *et al.* Long-term impact of belimumab on health-related quality of life and fatigue in patients with systemic lupus erythematosus: six years of treatment. *Arthritis Care Res* 2019;71:829–38.
- 32 US Food Drug Administration (FDA). Guidance for industry on systemic lupus erythematosus: developing medical products for treatment; 2010.
- 33 Berry P, Williams-Hall R, Williamson N. *Qualitative evaluation of patient experience of lupus nephritis*. Manuscript submitted for publication, 2022.
- 34 Patrick DL, Burke LB, Gwaltney CJ, *et al.* Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 2--assessing respondent understanding. *Value Health* 2011;14:978–88.
- 35 Petri M, Orbai A-M, Alarcón GS, *et al.* Derivation and validation of the systemic lupus international collaborating clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2012;64:2677–86.
- 36 Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
- 37 Ericsson KA, Simon HA. Verbal reports as data. *Psychol Rev* 1980;87:215
- 38 ATLAS.ti. *Scientific software development GmbH B, Germany. Atlas.ti. software version 7*, 2013.
- 39 Willis GB. *Analysis of the cognitive interview in questionnaire design*. Oxford University Press, 2015.
- 40 Holloway L, Humphrey L, Heron L, *et al.* Patient-reported outcome measures for systemic lupus erythematosus clinical trials: a review of content validity, face validity and psychometric performance. *Health Qual Life Outcomes* 2014;12:116.
- 41 Mahieu M, Yount S, Ramsey-Goldman R. Patient-reported outcomes in systemic lupus erythematosus. *Rheum Dis Clin North Am* 2016;42:253–63.
- 42 Strand V, Chu AD. Measuring outcomes in systemic lupus erythematosus clinical trials. *Expert Rev Pharmacoecon Outcomes Res* 2011;11:455–68.
- 43 Norquist JM, Girman C, Fehnel S, *et al.* Choice of recall period for patient-reported outcome (pro) measures: criteria for consideration. *Qual Life Res* 2012;21:1013–20.