Validation and evaluation of the German version of the Systemic Lupus Activity Questionnaire (SLAQ)


Policlinic of Rheumatology, Heinrich-Heine-University, Düsseldorf, Germany.

Abstract

Objective

Disease activity accounts for damage, overall mortality and co-morbidities in SLE and should frequently be assessed to adapt therapeutic patient management. The Systemic Lupus Activity Questionnaire (SLAQ) is a patient-reported instrument for the assessment of disease activity derived from the Systemic Lupus Erythematosus Activity Measure (SLAM) and was originally developed in English. Our aim was to validate the SLAQ in German and evaluate its use in a large cohort.

Methods

We translated and adapted the SLAQ questionnaire in German. It was applied to SLE outpatients at a tertiary centre (n=328) and compared to the SLAM and other SLE outcome parameters. Internal consistency, criterion validity, inter-rater and test-retest reliability as well as construct validity were examined. Correlation, Cronbach’s alpha, Mann-Whitney U-test or the Kruskal-Wallis one-way analysis of variance test were ascertained where appropriate. Levels of statistical significance were defined at 5% (p<0.05). All reported p-values are two-tailed.

Results

The German SLAQ showed a comparable strong correlation with the SLAM (r=0.632, p<0.0001) as the original version of the SLAQ and presented a good to excellent internal consistency reliability (Cronbach’s alpha=0.89). Accumulated damage as well as low disease activity are factors possibly influencing the score. Amongst others, scores were higher in patients with more reported flares, lower self-reported overall health, lower functional status and higher daily doses of prednisolone.

Conclusion

Our German version of the SLAQ shows a comparable validity as the original SLAQ and is a promising instrument to survey disease activity in clinical routine as well as in clinical and epidemiological studies. Possible interacting factors need to be considered when applying.

Key words

systemic lupus erythematosus, patient-reported instrument, disease activity, SLAQ, German
Introduction

Evaluation of disease activity is one of the most important assessments in care of chronic and especially rheumatic diseases. Disease activity accounts for accrual of damage and complicating co-morbidities resulting in an increase of morbidity and mortality (1-4). In systemic lupus erythematous (SLE) disease activity causes damage, coronary artery disease as well as overall mortality (3-5). Careful monitoring of disease activity in SLE and adapting therapeutic management has remarkably improved mortality over the past decades (4, 6). With respect to this result frequent assessment of disease activity in patients with SLE is recommended by the European League Against Rheumatism (EULAR) (7) using a validated index at each patient visit. Multiple instruments have been developed over the last decades to assess disease activity in SLE, e.g. British Isles Lupus Assessment Group index (BILAG), Systemic Lupus Erythematous Disease Activity Index (SLEDAI), Systemic Lupus Erythematous Activity Measure (SLAM), European Consensus Lupus Activity Measurement (ECLAM) and Systemic Lupus Activity Questionnaire (SLAQ) (8-13). Most of these instruments retrieve information about examined disease characteristics and laboratory values to calculate an overall activity score and hence require physicians’ assessments. However, in epidemiologic or clinical research and routine clinical setting it is often more suitable to use patient administered questionnaires with regard to time and cost saving. Furthermore, patient-reported outcomes (PROs) represent self-assessments that are not being interpreted or weighted by a clinician. PROs have even gained recognition in regulatory agencies like the Food and Drug Administration (FDA) (14). With regard to this Karlson et al. developed the SLAQ (12) published first in 2003, that is based on the SLAM (9, 15). The SLAQ omits laboratory results and focuses on disease characteristics that are amenable for self-reporting. The SLAQ contains 24 items representing symptoms of 9 different organ systems. Symptoms that appeared in the past three months are weighted from mild, to moderate and severe. The SLAQ score showed a moderately high correlation with the SLAM (without laboratory results) and a positive predictive value ranging from 56 to 89% for detecting clinically significant disease activity in a cohort of 93 SLE patients (12). A further validation study of Yazdany et al. demonstrated adequate reliability, construct validity, and responsiveness in a larger, community-based cohort (16). To our knowledge, there has only been one validated translation of the SLAQ until now: the Japanese version showed acceptable reliability and validity among 246 Japanese patients with SLE with only a weak correlation to the SLEDAI (17).

To provide the instrument for German-speaking populations and further assessment of its usefulness, we validated the SLAQ in German.

Materials and methods

The methodology of translation was oriented to previously published recommendations (18). The SLAQ questionnaire was translated by a qualified translator. The results were discussed by a committee of four rheumatologists and a medical documentalist and pre-tested by 15 SLE patients. Minor cross-cultural adaptations and corrections were necessary. A professional back translation was again reviewed by the committee to release the final questionnaire and to ensure content validity. The final version of the German SLAQ was adapted to a computerised form in the local web-based patient documentation system (DocuMed.rh) used in outpatient clinics and private practice throughout Germany.

The questionnaire was applied to consecutive outpatients (HHUD cohort; Heinrich-Heine-University Düsseldorf cohort) diagnosed with SLE according to the revised classification criteria of the American College of Rheumatology from 1997 (19, 20). The questionnaire was completed by every patient who visited our clinic over a time period of 24 months and consented to participate. The patients processed the questionnaire directly at a computer terminal. Only for physically impaired patients...
Validation and evaluation of the German SLAQ / G. Chehab et al.

unable to use the computer or for patients inexperienced in the use of computers input assistance was offered. Beyond that no further support was given. After the appointment the SLAM was completed by the attending physicians blinded for the SLAQ information. During the same outpatient visit further disease parameters were recorded, amongst others patient’s evaluation of overall health on a Numerical Rating Scale (NRS 0 best – 10 worst), patient’s and physician’s assessment of disease activity in the past 3 months (NRS, 0 none – 10 highest), the occurrence of lupus flares in the past 3 months (patient-reported; NRS 0 no, 1 mild, 2 moderate or 3 severe flare), the Hannover Functional Questionnaire (FFbH), a common outcome measure for physical functioning in German speaking countries, which is comparable to the well-known Health Assessment Questionnaire (HAQ) though the scale is inverted and a patient-reported damage questionnaire (Brief Index Of Lupus Damage) (21, 22).

For test-retest validity assessment, a paper copy of the SLAQ was provided to 75 consecutive patients to be completed within 24 to 48 hours after their visit and to be mailed back to the study site. Our study collected cross-sectional data and was approved by our local ethics committee.

The SLAQ scores were calculated using the formula proposed by Karlson (12). Similarly we calculated the SLAQ symptom score which represents only the count of positive responses without the weighting of the individual SLAQ items. Laboratory values were omitted for the calculation of the SLAM (referred to as SLAM$_{lab}$).

Criterion validity and inter-rater reliability are represented by the comparison of the SLAQ and the SLAM$_{lab}$ score as well as between single items (correlation). Intra-rater reliability was not addressed in this study, because data collection took part during clinical routine. Internal consistency was measured by calculating Cronbach’s alpha and test–retest reliability by evaluating the correlation between the patient responses at two time points and Cronbach’s alpha. A Cronbach’s alpha >0.70 was considered acceptable. To assess construct validity we compared other disease characteristics and scores with the SLAQ and its quartiles in our cohort. To ascertain additional effects on the SLAQ we examined differences in the SLAQ/SLAM$_{lab}$ correlation between our raters and between patients with lower and higher disease activity, represented by SLAM$_{lab}$ terciles (SLAM$_{lab}$ <2, 3-5 and >5). Differences in characteristics were tested using the Mann-Whitney U-test or the Kruskal–Wallis one-way analysis of variance where appropriate. Pearson’s and Spearman’s rank correlations were both carried out. Spearman correlation is only reported when considerably different. Levels of statistical significance were defined at 5% (p<0.05). All reported p-values are two-tailed.

Results

We evaluated 328 SLAQ questionnaires in our cohort. Completion rate was good with only one SLAQ item response rate lower than 99% (‘swollen glands (nodes) in the neck’). The SLAQ presented a good to excellent internal consistency with a Cronbach’s alpha scoring 0.89.

The outpatients had a mean age of 43.6 and mean disease duration of 13.4 years. Further demographic details and results of self-reported disease activity, assessments of general health and physical functioning are listed in Table I.

Both the SLAQ and the SLAM$_{lab}$ show a clear skewness to the right (SLAQ 0.96 and SLAM$_{lab}$ 1.14) with 17.7% zero-scored SLAM$_{lab}$ and 8.4% SLAQ questionnaires respectively. The median is 3 in SLAM$_{lab}$ and 7 in the SLAQ and the upper quartile is 6 in the SLAM$_{lab}$ and 14.0 in the SLAQ. Pearson’s r depicted a strong correlation of 0.632 between the SLAM$_{lab}$ and the SLAQ (Pearson’s rho 0.579; both p<0.0001). The SLAQ symptom score presented an inferior correlation with the SLAM$_{lab}$ (r=0.586, p<0.001). Correlation with other reported disease and demographic parameters scored lower as depicted in Table II. Neither the SLAQ nor the SLAM$_{lab}$ demonstrated a significant correlation with disease duration.

The correlations between the corre-
Validation and evaluation of the German SLAQ / G. Chehab et al.

Table II. Correlation of SLAQ/SLAM<sub>total</sub> with other disease parameters.

<table>
<thead>
<tr>
<th></th>
<th>SLAQ</th>
<th>SLAM&lt;sub&gt;total&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLAM&lt;sub&gt;total&lt;/sub&gt;</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td>SLAQ</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td>SLAQ symptom score</td>
<td>0.945</td>
<td>0.586</td>
</tr>
<tr>
<td>FFbH</td>
<td>-0.558</td>
<td>-0.498</td>
</tr>
<tr>
<td>HAQ</td>
<td>0.558</td>
<td>0.499</td>
</tr>
<tr>
<td>Disease activity (physician, NRS 0-10)</td>
<td>0.473</td>
<td>0.550</td>
</tr>
<tr>
<td>Disease activity (patient, NRS 0-10)</td>
<td>0.708</td>
<td>0.522</td>
</tr>
<tr>
<td>Overall health (patient, NRS 0-10)</td>
<td>0.661</td>
<td>0.549</td>
</tr>
<tr>
<td>Lupus flare (patient, 0-3)</td>
<td>0.534</td>
<td>0.398</td>
</tr>
<tr>
<td>Prednisolone equivalent dose (mg)</td>
<td>0.277</td>
<td>0.227</td>
</tr>
<tr>
<td>BILD</td>
<td>0.270</td>
<td>0.250</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.169</td>
<td>0.168</td>
</tr>
<tr>
<td>Disease duration (years)</td>
<td>0.068&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.074&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

All values significant p<0.005 except<sup>a</sup> (not significant).

Table III. Description of the HHUD cohort split by SLAQ quartiles.

<table>
<thead>
<tr>
<th></th>
<th>SLAQ Score</th>
<th>p-value&lt;sup&gt;3&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-3 (n=76)</td>
<td>4-7 (n=57)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>39.1</td>
<td>42.9</td>
</tr>
<tr>
<td>Lupus flare (patient, 0-3)</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>Disease activity (patient, NRS 0-10)</td>
<td>0.8</td>
<td>1.8</td>
</tr>
<tr>
<td>Disease activity (physician, NRS 0-10)</td>
<td>1.4</td>
<td>1.8</td>
</tr>
<tr>
<td>Overall health (patient, NRS 0 best - 10 worst)</td>
<td>1.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Hannover Functional Questionnaire (FFbH)</td>
<td>96.8</td>
<td>93.0</td>
</tr>
<tr>
<td>Health Assessment Questionnaire (HAQ)</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>BILD</td>
<td>0.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Prednisolone equivalent daily dose (mean mg)</td>
<td>2.6</td>
<td>4.2</td>
</tr>
</tbody>
</table>

<sup>3</sup>calculated by the Kruskal-Wallis one-way analysis of variance.

**Discussion**

Our study assessed the reliability and validity of the German Systemic Lupus Activity Questionnaire in a tertiary centre cohort. The acceptance of the questionnaire was good with a high response rate and only low missing value numbers of single items which confirms the comprehensibility of our German questionnaire. The SLAQ presented a good to excellent internal consistency. We observed a strong correlation with the SLAM<sub>total</sub>, representing criterion validity and inter-rater reliability, that is identical with the correlation observed by Karlson <em>et al.</em> (Pearson’s r=0.62, p<0.0001) (12). On the contrary, Okamoto <em>et al.</em> observed only a weak correlation between the SLEDAI-2K-nolab and their Japanese SLAQ questionnaire (Spearman’s rho=0.18) (17). Irrespective of the quality of the translation or cultural adaptation, the choice of the SLEDAI as the comparative standard might have caused this low value. The SLAM implies a high number of items that are subjective and rely on patient-reported symptoms. Therefore and because of its derivation from the SLAM a higher correlation is to be expected between the SLAQ and SLAM<sub>total</sub> compared to other less patient-derived disease activity measures. This was also discussed by Fortin <em>et al.</em> who experienced similar differences in the evaluation of the SLEDAI and SLAM concerning the HAQ and SF-36 and its subscales (23). Other patient characteristics that are not directly associated to disease activity might also influence the scoring of the SLAQ. In a mixed cohort of Hispanic and White SLE patients Carr <em>et al.</em> showed that depression and ethnicity independently correlated with the SLAQ (24).

Correlation of the SLAQ with the patient-reported damage instrument (BILD) was considerably lower as to be expected regarding divergent construct validity. However we observed higher damage scores in the upper than in the lower SLAQ quartile. This indicates that higher damage scores might affect the SLAQ. Wang <em>et al.</em> investigated the relationship between health related quality of life, disease activity and damage (25). Thereby, they depict-
Validation and evaluation of the German SLAQ / G. Chehab et al.

ed a significant correlation between the Systemic Lupus International Collaborating Clinics Damage Index and the SLAM ($r=0.36, p<0.05$) in a cohort with a comparable mean disease duration of 13 years. So even in physician reported outcomes some overlap between damage and disease activity is observed and does not contradict the effectiveness of the instruments. Other outcome parameters that are expected to be linked to disease activity, represented by physical functioning, overall health and patient-reported flares correlated comparably well with the SLAQ and SLAM and thereby support external convergent construct validity. The SLAQ benefits from its derivation of the SLAM, as other disease activity measures that collect less patient-derived information like the SLEDAI or the BILAG do not demonstrate comparable associations with general health status or health related quality of life (23, 26).

In addition, we were able to prove external convergent construct validity. Sociodemographic and outcome parameters developed correspondingly with increasing SLAQ quartiles in our cohort. Conflicting seems the increasing age in higher SLAQ quartiles as younger patients in SLE are considered to present with a more active disease. Alarcón et al. reported that higher age was even negatively associated with high disease activity (though barely) (27). In our study other influencing factors like accumulated damage or conflicting co-morbidities might be responsible for this observation.

We detected that the correlation between the SLAQ and SLAM was inferior in patients with lower disease activity (SLAM >5) compared to patients with higher disease activity (SLAM >5). This was particularly evident for a subgroup of patients all visiting a specialist clinic on care before, during or after a planned pregnancy. Though the SLAM scores in these patients depict a low disease activity the corresponding SLAQ scores are disproportional high. This suggests that the scaling of the SLAQ is limited in low disease activity. An impact might have mild to moderate subjective and unpecific items (e.g. fatigue, shortness of breath, forgetfulness, depression, muscle pain, joint pain) that are not correctly attributed to SLE by the patients, an observation that was likewise published by Karlson et al. (12). However limiting the questionnaire to more specific symptoms would reduce its sensitivity. With regard to this, discordance between physician and patient evaluation is an important factor that needs attention when evaluating patient-reported outcomes. In addition, patients’ evaluation of lupus activity might be biased by their psychological status (28) or – like in rheumatoid arthritis – by accompanying pain or fatigue (29, 30). Both above mentioned aspects as well as low positively reported frequency of some items have to be considered when evaluating the correlation of the single items (SLAQ/SLAM). Despite of this the resulting overall score of the SLAQ remains comparable to the SLAM.

Study limitations are the lack of intrarater-reliability analysis, which we were not able to implement in clinical routine and the focussing on the SLAM as comparator. Additional studies are needed to assess the agreement between the German SLAQ and less patient-derived instruments (e.g. SLEDAI, ECLAM). Furthermore we studied patients at a single tertiary centre. Further analysis is eligible, focusing on the longitudinal use of the SLAQ in phases of accurately defined SLE flares as well as remission to assess its correct scalability especially in early phases of the disease. Besides the role of accrued damage and co-morbidities in patients’ evaluation of disease activity needs further investigation in cross-sectional and longitudinal studies.

An interchangeable longitudinal use of the SLAQ and SLAM was not advis- able. Due to validity of the SLAQ the exclusion of laboratory values would be prejudicial. Moreover, the scores, calibrated to their respective educated rater, result higher in the SLAQ than the SLAM.

Conclusion

The German version of the SLAQ shows a comparable validity as the original SLAQ in assessing disease activity. The German SLAQ represents a feasible way to survey disease activity in clinical routine as well as in clinical or epidemiological studies. The reliability of the questionnaire as well as validity in longitudinal use needs further validation.

References


