Why and how should we measure disease activity and damage in lupus?

Joy Feld1, David Isenberg2

1. Carmel and Lin medical centres, Rheumatology Unit, Haifa, Israel
2. Centre for Rheumatology, Department of Medicine, University Collage London, London, United Kingdom

Available online: 29 April 2014

Summary

The assessment of disease activity and flare and differentiating them from permanent damage in patients with SLE is challenging. The SLEDAI, SLEDAI-2K and SELENA-SLEDAI measure global disease activity. The BILAG measures organ-specific activity. The BILAG better captures the change in the different organs at the expense of complexity. The SRI is a composite index incorporating both BILAG and SLEDAI indices and a physician’s global assessment. It has been used in the most recent clinical trials. Damage correlates with prognosis; it is assessed by the SLICC/SDI index. This index scores damage whatever the cause, disease or treatment related, or the consequence of concomitant disease. The disease activity and damage indices do not correlate well with the patient’s health related quality of life (HRQoL), the degree of disability or the impact of disease. The impact of the patients’ joint disease on their HRQoL is assessed via the HAQ questionnaire and the global health status via the SF-36 index, or one of the more recently described lupus specific quality of life indices [Lupus QoL]. The global assessment instruments and the BILAG index can also be used in children and adolescents with SLE. However, a modified paediatric version of the SLICC/SDI damage index is advised. Many advances have been achieved in disease activity and damage measurement in the past 20 years but the problem of how best to capture flare accurately remains.

Systemic lupus erythematosus (SLE) is an autoimmune rheumatic disease which is heterogeneous in its presentation, course and outcome. The disease can present with diverse manifestations and severity. When active and flaring, it can cause life threatening manifestations and, in some, irreversible permanent end organ scaring referred to as damage. The differentiation between disease activity and damage is crucial for patient management and treatment strategies.

Accurate assessment of disease activity and damage in SLE is needed to: assess the disease longitudinally in observational and clinical trials, differentiate patterns of disease involvement, evaluate responses to new drugs and evaluate outcome. Several validated, in two cases updated,
activity assessment instruments have been available since the 1980s. More recent studies (described later on in this review) have focussed on ascertaining reliability and validity for classifying and monitoring groups of patients in both the clinic and research setting.

For the purpose of this review, we have selected those indices that have shown the strongest evidence of validity when used by investigators from different countries in large studies of patients with SLE.

Assessing disease activity

Assessment of disease activity is crucial for forming the basis of treatment decisions in routine clinical practice and research. Therefore, it is necessary to be able to quantify the change in disease activity in a way that includes all the possible manifestations.

Two cardinal features of SLE have challenged investigators when assessing disease activity:

- the complex multi-system nature of this disease with levels of disease activity which may fluctuate in different body systems, varying between patients and within the same patient over time;
- the absence of a “gold standard” for comparing the instruments with an objective measure.

Several assessment systems have been developed and validated, including: the British Isles Lupus Assessment Group (BILAG) [1,2], Systemic Lupus Activity Measure (SLAM) [3], European Community Lupus Activity Measure (ECLAM) [4,5], and the SLE Disease Activity Index (SLEDAI) [6,7]. These systems have been developed to capture disease activity at a given time. Clearly patients may have no active disease, or if they are active this may be persistent, improving or deteriorating.

The SLEDAI, SLAM and ECLAM are global indices providing an overall measure of activity. In contrast, the classic BILAG [1,2], established on the principle of the physician’s intention to treat, provides a more comprehensive “at a glance” overview of activity in eight organs/systems. In all activity indices, the recorded clinical data should only be entered if the physician is sure that the feature is due to SLE. Assessment in the BILAG index distinguishes clinical features that are improving from those that are getting worse, staying the same, or are new or recurrent. Instead of giving a single score covering all systems, the BILAG index gives individual scores (from A to E, where A represents the highest disease activity) for eight different systems. There were problems, however, with the classic BILAG index, which incorporated a small number of items that were more clearly due to damage rather than to disease activity and failed to capture adequately disease activity in the gastrointestinal or ophthalmic systems. The substantially revised version, the BILAG 2004 index [2], has now been validated [8], shown to be reliable [9] and sensitive to change [10].

The SLEDAI, SLAM, and BILAG have performed in an effective and reliable manner in many studies and have been shown to correlate well with one another, despite their different origins [11–13]. The BILAG assessment tool takes longer to complete than SLEDAI or SLAM, but all of the indices work optimally with training.

The global indices have the advantage of simplicity, in that the clinical features in each organ/system are assigned numerical scores that are summated to give a total score for disease activity. The main problem with these scoring systems is that points are awarded for clinical features if present, but do not distinguish those features that are improving from those that are deteriorating or those that are unchanged. The original SLEDAI version was introduced in 1985; it was revised in 2002 [6,7,14] to reflect persistent active disease in those descriptors that had previously considered new or recurrent occurrences (SLEDAI-2K). The SLEDAI-2K measures only complete recovery in active descriptors on follow-up visits. For this reason the SLEDAI-2K Responder Index-50 (SZE-RI-50) was developed [15]. This index is able to capture partial improvement ≥ 50%, in each of the active descriptors at subsequent visits. However, although this index captures partial improvement, it does not capture deterioration in SLE symptoms [16].

There is ongoing debate about how best to capture flare in patients with lupus. This clearly represents disease which is becoming more active. However, in SLE the accurate and uniform acceptance of flare definition has been challenging. For example, can a patient with increasing disease activity in just the kidney be judged to be suffering a worse flare than a patient with increasing disease activity in the skin, lung or joints? Should flare be determined by the treatments likely to be offered to a patient?

A “flare” of disease activity has been utilised in previous and ongoing studies. The critical question of how best to define a flare of SLE remains problematic. Using the BILAG 2004 index, a flare can be defined in terms of the number of systems scoring A or B based on items recorded as new or worse. On this basis, one might define a severe flare as occurring in a patient with an A score in any system, or a moderate flare as a B score in at least two systems. In the global score indices the presence of a flare is defined according to a pre-specified increase in this total score. Therefore, if a patient was to increase her\his score by, say, 4 points; a flare might be deemed to have occurred.

The SELENA-SLEDAI flare index was developed for use in clinical trials by the Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) group with the intention of distinguishing severe flares from those that are only mild or moderate [17,18]. The SELENA group has recently devised a more comprehensive instrument that distinguished mild from moderate flares. It provides separate analysis of flares in different organ systems, and collects treatment data as part of the evaluation. The revised SELENA flare index is organ-system based, and is not linked to the
SLEDAI. For each organ system, suggested clinical manifestations are given but the categorisation is dictated by the treatment decision. In particular, a “mild flare” is assigned if there is either no treatment, or if there is initiation of hydroxychloroquine, prednisone 7.5 mg per day or less or a non-immunosuppressive therapy. Definition of a “moderate flare” requires the use of prednisone greater than 7.5 mg per day but less than 0.5 mg/kg per day, or immunosuppressive therapy (other than cyclophosphamide), and “severe flare” is defined as prednisone (or equivalent) 0.5 mg/kg per day or greater, cyclophosphamide, biological treatment, warfarin\heparin initiation or hospitalisation. The SELENA flare instrument thus shares concepts with the BILAG index, which as stated above predicated on the notion that disease activity is based on the principle of the “physician’s intent to treat” in individual organs or systems [19].

A working party organised by the Lupus Foundation of America (LFA) recently defined a flare as a measurable increase in disease activity in one or more organ systems involving new or worse clinical signs and symptoms and/or laboratory measurements. It must be considered clinically significant by the assessor, usually there would be at least consideration of initiation or increase in treatment [19]. Physician’s global assessment uses the treating physician’s overall assessment to assign a numerical value to the disease activity on a visual analogue scale of 0 to 3. This assessment has been used to define flares (> 1 point rise) in clinical trials. Physician visual analogue scale rating has been successfully incorporated as part of the SLE Responder Index (SRI). The SLE Responder Index is a composite index developed to incorporate the strengths of different disease activity indices. This index provides a comprehensive definition of meaningful clinical response and has been used to define the primary end point in clinical trials. The SLE Responder Index utilizes the SELENA-SLEDAI score to determine global improvement, the older, less satisfactory version of the BILAG domain scores to ensure no significant worsening in previously unaffected organ systems, and physician’s global assessment to ensure that improvements in disease activity are not achieved at the expense of the patient’s overall condition. The SLE Responder Index has been used successfully in the belimumab phase 3 trials and has the potential to serve as an outcome measure in future SLE therapeutic trials [20].

The SLEDAI, SELENA–SLEDAI, SLEDAI 2000, and BILAG have been successfully used in observational trials and case studies, although baseline Disease Activity Index (DAI) scores were not always predictors of subsequent damage or other outcomes. These DAI were validated in the context of long-term observational trials studies and not in randomized clinical trials (RCTs). The few RCTs conducted have shown that improvement in DAI scores correlates with response rates, disease remission, and flare prevention; however, a threshold of clinically meaningful change has not been established [21].

Assessing damage

It may be difficult to distinguish symptoms due to active current inflammation from those due to damage, which implies permanent change. Permanent organ damage in SLE may be due to disease, treatment of disease, or concomitant diseases. Thus, a pain in the hip might be the consequence of synovitis or aseptic necrosis. In the former case anti-inflammatory drugs, including steroids, may be required. In the latter case, steroids need to be reduced or stopped and surgical intervention sought.

Currently, there is only one internationally agreed measure of damage which is “The Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SLICC/SDI)” score [22]. The SLICC/SDI assesses the cumulative effect of the disease since diagnosis. It has shown good inter- and intra-observer reliability. The index records damage in 12 organs or systems. The change must have been present for at least six months and is ascertained clinically or by simple investigations. As opposed to the activity indices, the SLICC index score does not distinguish whether damage results from the disease, the medications used or other unrelated concomitant causes, e.g., diabetes. It has useful prognostic value in clinical and research studies. It has been shown by several groups that the early acquisition of damage is a sign of a poor prognosis [23,24]. In 2001, Rahman et al. published their remarkable results regarding SLE patients from the University of Toronto Lupus Clinic presenting within one year of their diagnosis prior to 1988. They followed two-hundred and sixty-three patients for 10 years. Twenty-five percent of these patients, who exhibited damage at their first SLICC/ACR Damage Index assessment after only one year of disease, died within 10 years of their illness as compared to only 7.3% who had no early damage [24].

The key characterisations of the main activity and damage indices are summarized in table 1 [21].

Assessing patient health status

Throughout the course of their disease, individuals with SLE face considerable physical, psychological and social challenges. The disease has profound effects on health-related quality-of-life (HRQoL), which have been documented extensively in the literature. The disease activity and damage scores mentioned do not take into account the patient’s health related quality of life neither degree of disability nor impact of disease. The HRQoL incorporates mental, social, and physical health and is assessed by questionnaire. Capturing decrements and improvements in HRQoL has therefore become important in clinical research in SLE.

The Health Assessment Questionnaire (HAQ) is simple to use and is widely used in rheumatoid arthritis. It has been validated in patients with SLE who have arthritis. However, it assesses only the joint disease and does not take into account other
### Table I

**Key characterisations of the main activity and damage indices**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>Disease activity in the past month</td>
<td>Disease activity in the past month</td>
<td>Permanent damage regardless of attribution present for at least six months</td>
<td>Disease activity in the past month, based upon the “intent-to-treat” premise</td>
<td>Disease activity in patients in the last 10 days</td>
<td>Disease activity in large cohorts of patients</td>
</tr>
<tr>
<td><strong>Strengths</strong></td>
<td>This index includes both dimensions: disease activity and disease severity</td>
<td>Derived from large cohorts of patients and from analysis of data from a large multicenter study</td>
<td>The only tool available for disease damage Excellent for prognostic studies</td>
<td>Reflects disease change with time Sensitive to small changes and distinguishes between disease activity and severity Aims to show activity in individual systems “at a glance” rather than combining them into a global score</td>
<td>Simple to use</td>
<td>Useful for large epidemiologic studies</td>
</tr>
<tr>
<td><strong>Caveats and cautions</strong></td>
<td>Many items are subjective not objective Difficulty in distinguishing the severity of the disease Scores damage, not only activity</td>
<td>May miss changes in severity over time</td>
<td>In patients with a long duration of SLE, the accuracy depends on the information available</td>
<td>More complicated and time consuming</td>
<td>Does not record improving or worsening, and does not include severity within an organ system</td>
<td>Not useful for clinical assessment, only for research purposes Depends on the level of education of patients</td>
</tr>
</tbody>
</table>
aspects of the disease; therefore, no correlation with the SLEDAI activity index has been seen. Currently, the Short Form-36 (SF-36) index is preferred for use in clinical practice. It is easy to complete and assesses health status over the preceding month. The SF-36 has been validated by Stoll et al. The methods of assessment mentioned above are most useful in clinical trials and in the longitudinal follow up of patients with SLE [25,26]. Recently three specific lupus assessment tools have been designed to ascertain HRQoL in SLE: the Lupus Quality of Life (LupusQol) [27], SLE-specific Quality of Life questionnaire (SLE-QoL) [28] and SLE Quality of Life Questionnaire (L-Qol) [29]. Administration of these instruments to a clinical cohort wherein physician-assessed measures of both disease activity and damage are available will yield further insight into both construct validity and also discriminate validity, or the independence of these instruments from other disease assessments in SLE. In addition, information on responsiveness is not available and it will be needed to assess whether these measures might be applied to treatment studies of SLE. Finally, validation of these instruments in other populations, including patients with more severe disease phenotypes, will be useful [30].

Measurement of disease activity and damage in the paediatric population

All global disease activity indices have been found to be reliable and valid for use in children and adolescents with SLE, and none of them has shown clearly superior metrological properties. The choice of a specific tool may largely depend on the purposes of the study, the investigational setting (standard clinical practice or research) and the personal preference of the investigator [26,31].

Although the SLICC\SDI has proved suitable to assess damage in paediatric SLE patients, it has some important limitations for use in the paediatric age group; the chief of which is the inability to capture some forms of damage that are unique to children and adolescents, namely growth failure and delayed puberty. The capacity for "repair" in children clearly exceeds that in adults. For example, in the original damage index a GFR < 50 mL/min for more than six months gets a damage item point on the grounds that adults in this situation are unlikely to improve. However, this is not necessarily the case in children, for whom improvement after six months of reduced GFR is possible. Therefore, the use of a modified paediatric version of the SDI in paediatric SLE patients is advised [26,31]. Further work needs to be done in this area.

Conclusions

Comprehensive assessment of a patient with lupus requires the assessment of activity, damage and quality of life. Several, though mostly, global score indices of lupus activity have been developed. They are relatively easy to use and have simple scoring systems. Their ability to distinguish symptoms which are improving from those which are deteriorating is an "Achilles heel" however. The BILAG activity index overcomes this problem and is more comprehensive though it requires more time to complete. The one damage index SLICC/SDI has "stood the test of time" though some adjustments are needed in the paediatric population.

Disclosure of interest: the authors have not supplied their declaration of conflict of interest.

References