Assessment of lupus: where are we now?

Management of patients with systemic lupus erythematosus (SLE) presents the doctor with particular problems. The heterogeneity of clinical manifestations and the fluctuating course of the disease produce difficulty in assessing the need for, and the response to, treatment. Central to this problem is the measurement of disease activity and, in particular, the differentiation between reversible organ dysfunction (activity) and irreversible organ damage. The potential for recovery, a prerequisite for treatment with potentially toxic drugs, may only be apparent in retrospect. There are various approaches to the measurement of disease activity in SLE; these include the monitoring of certain laboratory tests, the assessment of clinical features, or various combinations of the two. A major problem is the discrepancy between the needs of the clinician and the desires of the epidemiologist. The epidemiologist needs a reproducible method of assessing disease activity to ensure comparability between groups of patients – for example, for clinical trials, whereas the clinician needs a system which will indicate when a change of treatment is required in individual subjects. Satisfying these disparate requirements provides the major challenge for any monitoring system.

**Laboratory markers of disease activity**
Of the laboratory markers of disease activity in SLE, no single test has proved sufficiently reliable, sensitive, and specific to be used in isolation. The erythrocyte sedimentation rate is probably the most useful non-specific marker for distinguishing between active and inactive SLE, but may be normal even in the presence of major organ involvement and, conversely, may remain raised in remission, being maintained by changes in immunoglobulins and lipids. C reactive protein, the classical acute phase protein, is seldom raised in active disease unless the exacerbation is accompanied by serositis or infection. Of the more specific tests for SLE, raised titres of antibodies to dsDNA and low complement usually indicate active disease. In particular, if significant changes in these parameters are seen, they often predict the development of renal disease, and in this subset of lupus patients are useful markers of disease activity. Some patients without renal disease may have persistently raised levels of anti-dsDNA, however, and others may have low C4 owing to null alleles; whereas other patients with active renal disease may have normal results, thus limiting the applicability of these tests. Consequently, other measures have been developed, such as those indicating complement turnover, for example C3d and C4d, which have proved more sensitive and specific. Indicators of lymphocyte activation, such as soluble interleukin 2 receptors, have also had their advocates. In summary, although changes in laboratory tests are extremely helpful indicators of disease activity, most doctors do not base therapeutic decisions solely upon them because their sensitivity is less than absolute. Disease activity assessment in SLE must use the additional information provided by the clinical features, especially when non-renal disease is present.

**Clinical measures of disease activity**
In everyday clinical practice, disease activity is usually judged subjectively; the doctor decides whether the patient is getting better or worse using qualitative criteria, largely based on prior experience. However, 'disease activity' may mean different things to different clinicians. A dermatologist may consider a widespread, potentially scarring maculopapular eruption affecting the face to represent active disease, whereas a renal physician, using a different yardstick, might argue that a patient with skin disease should never be considered to have significant activity in comparison with patients with life threatening renal disease. This difficulty in measuring activity arises partly because it is difficult to separate the concepts of activity and severity. Severity connotes the seriousness or gravity of a disease manifestation with respect to the threat posed to the overall health or potential survival of a patient. It is clearly an important modifier of disease activity which has important implications for treatment, which itself can pose threats to health. To allow meaningful comparisons between centres managing lupus patients the need was recognised for standardised quantitative measures of disease activity to enable clinicians to communicate using a common language.

**Potential applications of clinical activity scales**
Clinical scales for measuring disease activity in SLE have a number of potential applications. The course of SLE is characterised by exacerbations (flares) and remissions, usually associated with treatment. A clinical activity index would enable standardisation of the definition of 'flare' and measurement of the reduction in disease activity in...
response to treatment. Such scales could then be used to dictate the need for, or response to, treatment in clinical trials, or to monitor the course of the disease in longitudinal studies of outcome. Unfortunately, at present this is still some way off and the major current use of indices is to ensure comparability between studies. Ultimately, the measures must move on to be clinically relevant and useful, or else they will be doomed to fall into disrepute. A further application of clinical activity indices is the validation of immunological tests or other potential markers of disease activity.

**Selected activity scales**

To be certain about comparability of data from different centres all clinicians should, ideally, use the same scale for measuring disease activity. In reality, however, the differing purposes for which scales are required mean that no single instrument has proved to be universally acceptable. This was illustrated by Liang et al., who reported more than 60 scales which have been developed in recent years. It is not within the scope of this article to provide a comprehensive review of all the currently available activity scales. Those described below represent recent instruments which are currently used at a number of centres, and which have been subjected to testing of reliability and validity.

**BILAG**

In 1984 a group of British rheumatologists convened with the aim of developing and testing a disease activity index which would score disease activity in different organ systems separately, and which would be suitable for use in studies designed to follow up patients longitudinally. The scale was developed according to the principle of the 'physician's intention to treat', on the premise that there is broad agreement amongst clinicians about when patients with SLE require disease modifying treatment. A nominal consensus approach was used to develop the scale, which has recently undergone modification and testing for reliability and validity in a number of centres in the United Kingdom. The BILAG index contains 86 items, each of which is rated by taking into account change in activity over time. It uses complex formulae to calculate separate scores for eight organ based systems, each of which is scored on a five point ordinal scale; an overall score is not produced. Immunological tests do not contribute to the score. The BILAG index is a complicated tool, but the development of a computer program has facilitated quick and easy recording, and analysis of data.

**SIS**

The SLE Activity Scoring System (SIS) consists of 21 clinical and 11 laboratory criteria, which are weighted to give a maximum possible score of 52. Patients are classified as either inactive (scoring 0–4), mildly active (5–8), moderately active (9–12), active (13–15), or very active (>15). It is one of the simplest instruments currently available.

**SLEDAI**

The SLE Disease Activity Index (SLEDAI) was developed in 1985 at the conference in prognosis studies in lupus, convened in Toronto. SLEDAI consists of 24 variables covering nine organ systems (including some immunological tests) scored according to weights derived using multiple regression techniques. A final weighted total SLEDAI score is then calculated. The activity index was validated on a 'test set' of patient profiles, which had been compiled using the 24 variables which are included in SLEDAI, by correlating the SLEDAI score with the doctor's global assessment.

**SLAM**

Liang et al in Boston developed the Systemic Lupus Activity Measure (SLAM), which contains disease manifestations originally derived from a review of publications and refined by clinical judgment using a consensus technique. It covers symptoms that occurred over the previous month and includes 24 clinical manifestations and eight laboratory parameters to evaluate organ systems that could not otherwise be assessed. Parameters of immune function are not included. This scale makes an attempt to include the dimensions of both activity and severity. Once a symptom or sign has been scored as either active or inactive the element of severity is assessed by a grading scale based on the need to treat with immunosuppressive agents, the need to follow up the patient more closely, or the functional or prognostic consequences of the manifestation. In this way a single weighted score is devised.

**LAI**

The Lupus Activity Index is a five part scale, reflecting disease activity over the previous two weeks. It includes the doctor's global assessment of disease activity, assessment of four symptoms (fatigue, rash, joint involvement, serositis), assessment of activity in renal, neurological, pulmonary, and haematological systems, scores attributed on the basis of drug treatment, and a variety of laboratory (including immunological) tests.

**Testing of disease activity indexes**

A prerequisite of clinical activity scales is that they should be reliable (give the same answer when used by different raters) and valid (measure what they purport to measure). Only recently have scales been systematically tested for these properties.

Liang et al tested between-rater and test-reset (within-rater) reliability for six systems including BILAG, SLEDAI, and SLAM. Twenty-five patients selected for a wide range of SLE activity were assessed separately by two doctors on two occasions about one month apart. Three components of variability were estimated: between patients, between two visits for the same patient, and between two raters at each visit. The largest part of the variance on every instrument was the 'patient's' component, which is reassuring because this is what the instruments are supposed to be measuring. The two instruments with the best overall metric properties were SLAM and BILAG. Three instruments (SLAM, SLEDAI, and BILAG) had high between-rater and between-visit reliability. Convergent validity was tested comparing the scales with each other, both cross sectionally and between visits. Strong convergent validity among the instruments was shown for cross sectional use, but the similarity between instruments was less when considering change between visits. Sensitivity to change was tested by the ability of an instrument to detect change in disease activity after the start of specific treatment. BILAG seemed to be the most sensitive, followed by SLAM; SLEDAI proved to be the least sensitive. One apparent drawback of BILAG was that it appeared difficult to complete, but the BILAG scores were calculated manually; use of the computer program would have made it much easier to use.
High correlations between BILAG, SLAM, and SLEDAI have also been demonstrated cross-sectionally by Gladman et al., suggesting that they are all measuring the same thing. There was good agreement between observers regardless of the instrument used. Worryingly, there was a lack of correlation between a 10-point clinical rating scale and any of the instruments. Furthermore, the clinical rating scale failed to distinguish between patients, emphasising the need for superior quantitative measures of disease activity.

In one of the few studies to test clinical activity indices longitudinally Petri et al. reported changes in the LAI and SLEDAI scores at times of lupus ‘flare’ (defined as an increase of one point on the doctor’s global assessment, a 1–3 point visual analogue scale) showing sensitivity to change. There were, however, important differences between the two scales in the reporting of the presence of certain clinical features at the time of a ‘flare’ – for example, the prevalence of active cutaneous lupus was 46–9% using LAI, but only 12–2% using SLEDAI. This difference was attributed to the fact that SLEDAI only scores new manifestations and cannot be used to score worsening of pre-existing manifestation, limiting its usefulness as a clinical disease index.

As part of an initiative to reach a consensus for measuring clinical disease activity within Europe, 25 participating centres contributed data on 25 consecutive patients with SLE to the coordinating centre in Pisa. Demographic and historical details, current clinical and laboratory data, and an overall doctor’s assessment (qualitative and semiquantitative) and management decision were collected on a chart and then transferred to a computer database. These data have been used to compare a number of systems including SLAM, SLEDAI, BILAG, SIS, and the French system. The five indices showed very high intercorrelation for the whole sample (which largely had patients with mild to moderate degrees of disease activity). When only those patients with very active disease were considered the correlation was less satisfactory. This is an important point which clearly needs further study; one of the applications of an activity index may be to monitor response to treatment in trials of treatment options in patients who certainly will have active disease. It is clearly unsatisfactory if the close correlations obtained between scales are just for patients in a stable state. At the high end of the activity scale the best correlations were between SLAM, SIS, and BILAG.

Outcome in SLE has been shown to be determined by, among other things, the number of exacerbations and the presence of renal or neurological disease. It would seem, therefore, important to measure disease activity in designated organ systems, which most of the indices fail to do.

The inclusion of immunological tests in some scales (for example, SIS, SLEDAI, LAI) makes them unsuitable for use as instruments to validate immunological or other scatological markers. Furthermore, given the heterogeneity of disease manifestations in SLE, and evidence linking scatological abnormalities with specific clinical manifestations, it is perhaps naïve to expect a new scatological test to correlate strongly with overall disease activity.

Three of the currently available activity scales have been shown to be reliable, both between and within raters (BILAG, SLAM, SLEDAI). The lack of a ‘gold standard’ for measuring disease activity in SLE makes it difficult to be sure that these scales are actually measuring what they are supposed to be measuring (criterion validity), but they do correlate strongly with each other in cross sectional studies, suggesting that they are, at least, all measuring the same thing (convergent validity). Convergent validity for these instruments used longitudinally remains to be established, particularly for patients with very active disease.

In conclusion, measurement of disease activity in SLE is central to patient care, and a number of instruments are available for this purpose. Although none is perfect, most are reliable and valid, and are suitable for classifying and monitoring groups of patients in the research setting. In reality, the indices are used with the additional benefit of laboratory markers and, as yet, no one has found the instruments sufficiently sensitive and specific to rely entirely upon them. The exact choice of instrument should be governed by the purpose for which it is required in clinical practice. Disease activity scales are unlikely to be appropriate for dictating treatment decisions in individual cases. An instrument which would be comprehensive and flexible enough for this purpose would necessarily prove too complicated and cumbersome for widespread use.

Summary
Decisions about treatment for patients with SLE are based on numerous criteria, including the rate of change of clinical features and disease markers (especially antibodies to dsDNA and markers of complement turnover), which organ systems are affected, the severity of manifestations, and the presence of pre-existing damage (which may reduce the reserve capacity of the organ system). Most of the currently available organ systems calculate a single overall score, at a single point in time, and take few of these considerations into account. SLAM is the only index to consider directly the scoring of severity as well as activity, though this concept is probably inherent in most of the other indices because of various methods for weighting the scores. Preliminary studies have indicated that four of the scales (BILAG, SLAM, LAI, and SLEDAI) are sensitive to change. Few of the indices have been tested longitudinally, hence their role in clinical trials remains to be established. None of the indices considers the impact of damage, indeed this is not their remit, but this concept is being considered by an international working group.

Epidemiology Unit, University of Manchester, Oxford Road, Manchester M13 9WL, United Kingdom
Department of Rheumatology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

Correspondence to: Dr Emery.


Manifestations of inflammation described by Celsus (42 BC to 37 AD) included colour, calor, rubor, and tumour. Galen (130–203 AD) added loss of function. He also systemised medicine up to this period. He publicised his 'galenic system', which became articles of faith for 1500 years. He was one of the earliest to use dissection, but mainly on animals. He dissected two gladiators who were killed. He coined the aphorism that gout is the daughter of Bacchus and Venus. He suspected the relation of renal calculi to gout. He described muscles, bones, and joints with great accuracy and used the term 'rheumatism' for involvement of these tissues. He turned practical hygiene into physiology, stressing the eclectic application of rest, exercise, diet, and sleep.

Yale Anzelman
Assessment of lupus: where are we now?

E Hay, C Gordon and P Emery

Ann Rheum Dis 1993 52: 169-172
doi: 10.1136/ard.52.3.169

Updated information and services can be found at:
http://ard.bmj.com/content/52/3/169

These include:

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/