

Review Article

Clinical Composite Measures of Disease Activity and Damage Used to Evaluate Patients With Systemic Lupus Erythematosus: A Systematic Literature Review[☆]



Isabel Castrejón,^{a,*} Iñigo Rúa-Figueroa,^b María Piedad Rosario,^c Loreto Carmona^d

^a Departamento de Medicina, División de Reumatología, New York University School of Medicine, NYU Hospital for Joint Diseases, Nueva York, United States

^b Departamento de Reumatología, Hospital Dr. Negrín, Las Palmas de Gran Canaria, Spain

^c Unidad de Investigación, Sociedad Española de Reumatología, Madrid, Spain

^d Universidad Camilo Jose Cela, Instituto de Salud Musculoesquelética, Madrid, Spain

ARTICLE INFO

Article history:

Received 9 August 2013

Accepted 31 January 2014

Available online 10 August 2014

Keywords:

Validation

Systemic lupus erythematosus

Disease activity

Damage

Systematic review

ABSTRACT

Objectives: To determine the most appropriate indices to evaluate the disease activity and damage in patients with systemic lupus erythematosus (SLE).

Methods: A systematic literature search was performed to identify validation studies of indices used to evaluate disease activity and damage. We collected information for each instrument on every aspect of validation including feasibility, reliability, validity and sensitivity to change using ad hoc forms.

Results: A total of 38 articles were included addressing the validation of 6 composite indices to evaluate disease activity (BILAG, ECLAM, SLAM, SLEDAI, LAI and SLAQ); and 3 indices to evaluate damage (SLICC/ACE-DI, LDIQ and BILD). Only the SLAQ, LDIQ and the BILD were self-administered. Feasibility and internal consistency was only studied in 3 indices (BILAG, SLAQ and SDI) with a Cronbach α ranging from 0.35 to 0.87. The intra-observer reliability was examined by the intraclass correlation coefficient for BILAG with a result of 0.48 (95% CI: 0.23–0.81) and using analysis of variance for SLAM-R (0.78), SLEDAI (0.33) and the LAI (0.81). The inter-observer feasibility was evaluated using the correlation coefficient for ECLAM (0.90–0.93), the SLAM (0.86) and MEX-SLEDAI (0.97–0.89). The construct validity was examined by means of convergence with other instruments, specifically with global assessment by the physician, with similar results between indices (0.48–0.75). Lastly, responsiveness was tested in all indices except LAI, SDI and LDIQ, with a standardized response mean ranging from 0.12 to 0.75.

Conclusions: Although multiple instruments have been validated for use in SLE it was not possible to find direct evidence of which is the most appropriate. BILAG and SLEDAI, with moderate feasibility and low responsiveness, are the 2 indices with a most complete validation and more extensively used.

© 2013 Elsevier España, S.L.U. All rights reserved.

Índices compuestos para evaluar la actividad de la enfermedad y el daño estructural en pacientes con lupus eritematoso: revisión sistemática de la literatura

RESUMEN

Objetivos: Determinar que índices compuestos serían los más apropiados para la evaluación de la actividad o el daño estructural en pacientes con lupus eritematoso sistémico (LES).

Métodos: Se realizó una revisión sistemática para identificar estudios de validación de índices de actividad y daño en pacientes con LES. De cada instrumento se recogió información sobre cada aspecto de la validación, como la viabilidad, la fiabilidad, la validez y la sensibilidad al cambio utilizando un formulario ad hoc.

Palabras clave:

Validación

Lupus eritematoso sistémico

Actividad de la enfermedad

Daño

Revisión sistemática

[☆] Please cite this article as: Castrejón I, Rúa-Figueroa I, Rosario MP, Carmona L. Índices compuestos para evaluar la actividad de la enfermedad y el daño estructural en pacientes con lupus eritematoso: revisión sistemática de la literatura. Reumatol Clin. 2014;10:309–320.

* Corresponding author.

E-mail addresses: isabelcastrejonf@gmail.com (I. Castrejón), iruafer@gobiernodecanarias.org (I. Rúa-Figueroa), mariapiedad.rosario@ser.es (M.P. Rosario), loretocarmo@gmail.com (L. Carmona).

Resultados: Se incluyeron 38 estudios de validación de 6 índices compuestos para la evaluación de actividad (BILAG, ECLAM, SLAM, SLEDAI, LAI y SLAQ) y 3 índices para la evaluación de daño (SLICC/ACR-DI, LDIQ y BILD). De estos instrumentos, solo el SLAQ, el LDIQ y el BILD son autoadministrados. En relación con los parámetros de fiabilidad, solo se evaluó la consistencia interna de 3 índices (BILAG, SLAQ y SDI); con resultados para el α de Cronbach de 0,35 a 0,87. La fiabilidad intraobservador fue examinada mediante el coeficiente de correlación intraclase para el BILAG, con un resultado de 0,48 (IC del 95%, 0,23–0,81) y mediante análisis de variancia para el SLAM-R (0,78), SLEDAI (0,33) y el LAI (0,81). La fiabilidad interobservador fue evaluada mediante coeficientes de correlación para el ECLAM (0,90–0,93), el SLAM (0,86) y el MEX-SLEDAI (0,87–0,89). Respecto a las dimensiones de validez, la validez de constructo se evaluó mediante medidas de convergencia con otros instrumentos, en concreto con la valoración global del médico, mostrando resultados similares entre todos los índices (0,48–0,75). Por último, la sensibilidad al cambio se evaluó en todos los índices, excepto el LAI, SDI y LDIQ, obteniendo resultados de respuesta media estandarizada de 0,12 a 0,75.

Conclusiones: A pesar de los múltiples índices validados para la evaluación de pacientes con LES, no se ha encontrado suficiente evidencia para determinar cuál es el más apropiado. Los índices BILAG y SLEDAI, con moderada fiabilidad y poca sensibilidad al cambio, son los 2 índices con una validación más completa y los más empleados.

© 2013 Elsevier España, S.L.U. Todos los derechos reservados.

Introduction

Measuring disease activity and irreversible damage in patients with systemic lupus erythematosus (SLE) is of vital importance to evaluate patient outcomes and prognosis, differences between patient groups and responses to new treatments. There are several validated indices, available since the early 1980s, although none have shown a clear superiority and therefore do not have universal acceptance.

Assessing patients with SLE is challenging for several reasons. On the one hand, the complex multisystem involvement caused by the disease and, on the other, its fluctuating course, which makes patients have very different evolutions and affections. Furthermore, the absence of a gold standard or a single variable to assess the disease activity makes the use of composite indices or scales necessary. The physicians global assessment (PGA) can be strongly influenced by the clinical experience of the physician and therefore produce a wide interobserver variability complicating comparison between patients.¹

The objective of this review is to evaluate aspects of the validation of the indices used to measure both disease activity and accumulated damage in patients with SLE.

Methods

As part of the consensus of the Spanish Society of Rheumatology for the use of biological therapies in patients with SLE, a systematic review was conducted to examine the validity of the composite indexes used to assess disease activity and cumulative damage.

Search Strategy

A search strategy was designed using the following bibliographic databases: MEDLINE, EMBASE and Cochrane Central Library up until March 2012. The search included MeSH terms and free text. The search strategies for MEDLINE, EMBASE and the Cochrane library were included in the Appendix (available on the web). The search was limited to human studies and studies published in English, French and Spanish. In addition, a manual search for the included articles was conducted in the bibliography.

The search was limited to human studies and studies published in English, French.

Selection of Studies

Validation studies, cohort studies, meta-analyses and systematic reviews were included. Regarding the type of participants,

studies were selected with patient age ≥ 18 and diagnosed with SLE according to the criteria of the American College of Rheumatology (ACR).^{2,3} An intervention was considered as any index, instrument or scale used to assess disease activity and structural damage. The comparator was any index compared with itself or with other indices. Finally, the outcome measures were evaluated aspects of validation and sensitivity, specificity, feasibility, reliability, validity and sensitivity to change. More detailed information on aspects of validation, for an easy interpretation of the results, is given in Table 1.

One reviewer (IC) selected the articles by title and abstract, following the inclusion criteria. Table 2 shows the characteristics of the included studies. Information concerning the validation of each composite index was collected using standardized forms. Evidence tables were generated to summarize every aspect of validation as well as the method used for evaluation.

Results

Fig. 1 details the search results. In the search strategy, of a total of 704 articles, 50 were selected for a detailed review after exclusion of 519 references by title and abstract. A total of 38 articles were included and 11 were excluded because they did not meet the inclusion criteria. Table 3 details the reasons for the exclusion of studies.

The included articles address the validation of 6 composite indices of disease activity: British Isles Lupus Assessment Group (BILAG), European Consensus Lupus Activity Measurement (ECLAM), Systemic Lupus Activity Index (SLAM), Systemic Lupus Erythematosus Disease Activity Measure (SLEDAI), UCSF/JHU Lupus Activity Index (LAI) and Systemic Lupus Activity Questionnaire (SLAQ), and 3 indices to assess cumulative damage: Systemic Lupus International Collaborating Clinics/American College of Rheumatology-Damage Index (SLICC/ACE-DI [SDI]), Lupus Damage Index Questionnaire (LDIQ) and Brief Index of Lupus Damage (BILD).

The characteristics and validation of each index were then reviewed in detail. Table 4 shows aspects of validation of these indices to allow for easy comparison.

Indices to Assess Disease Activity

British Isles Lupus Assessment Group

The British Isles for the Assessment of Lupus Group (BILAG) began to meet regularly in 1984 and in 1988 developed for the first time an index to measure disease activity in patients with SLE.⁴ This index was developed based on the physicians intention to

Table 1
Guide to Interpret the Results on Aspects of Validation.

Term	Aspects	Analysis technique
Viability (feasibility)	Time spent. Clarity of elements (simple). Accepted by patients and users	Pilot study (30 patients)
Degree (reliability) to which the instrument accurately measures: without error. Reliable, accurate and free of error (systematic error/bias and random error)	Internal consistency evaluates whether items measuring the same attribute have homogeneity among them. Depends on the number of items and correlation between them	Cronbach α (0–1): is interpreted as a correlation coefficient
Ability (validity) to measure that for which it is designed	Intraobserver or test–retest reliability: measures the stability of the scores given by the same evaluator in the same subjects Interobserver reliability or measurement error: the extent of agreement between 2 or more reviewers Logic or apparent (face validity): degree to which an index appears to measure what it wants to measure Content: if it contemplates all aspects related to the concept studied. Representative sample of the items Construct: reflects the concept to be measured or the instrument's ability to adequately measure it Criteria variables: comparison with a reference method already described and validated	ICC (intraclass correlation coef.): quantitative Cohen's kappa: qualitative Standard error of measurement, minimum detectable change agreement limits Expert opinion on the relevance and understandability: reduction of questions Panel of experts Structural: factor analysis. Test of Ho: correlations. Cross-cultural validation Gold standard Continuous variable: correlations and ROC curves. Dichotomous variables: sensitivity and specificity
Sensitivity to change (responsiveness)	Ability to measure changes. Intrinsic and extrinsic	Multiple methods

treat and evaluates specific manifestations requiring treatment in a total of 8 organs or systems: general, mucocutaneous, neurological, musculoskeletal, cardiorespiratory, vasculitis, renal and hematology. Unlike other indices which provide a global assessment of the disease, BILAG provides an assessment by organ system.

Subsequently, certain minor modifications of this index were conducted and assessed for both their reliability and validity. A study of Hay et al.,⁵ evaluated the BILAG version 3, which includes a glossary with explanations and recommendations for ease of use. BILAG showed good intraobserver reliability, with kappa ranging from 0.79 to 0.97, depending on the target system. To assess the criterion validity of the BILAG, it was compared to a gold standard defined as starting or increasing treatment with either prednisone or other immunosuppressive therapy. BILAG sensitivity for grade A in any organ or system was 87% and specificity of 99%. BILAG's Cronbach coefficient α calculated to assess the internal consistency and, therefore, the association between its components, is 0.35, lower than what is recommended for reliable comparisons. The BILAG showed moderate correlation with the physician global assessment ($\rho = 0.43$) and patient assessment ($\rho = 0.50$).⁶

The BILAG was also used to evaluate the occurrence of relapses in patients with lupus. The presence of a severe flare is defined as a score of A, new appearance, and a moderate flare is defined with a rating of B, if it was previously in D or E. A study published in 2003 by Gordon et al. assessed the degree of agreement between the definition of flare by BILAG and the real attitude about treatment by rheumatologists in routine clinical practice.⁷ 92% of patients with a severe flare (A) received intensive treatment by their doctor and only 41% received it in case of a moderate flare (B). Not only BILAG is useful for evaluating the presence of a lupus flare. In a study by Isenberg et al., 3 methods to assess flares were compared; BILAG 2004 index flair index generated for the Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) called the SELENA. SLEDAI, the Flare Index (SFI) and the VGM⁸: The BILAG 2004 had the highest ICC (0.54 [95% CI: 0.32–0.78]) compared to the SFI (0.21 [95% CI: 0.08–0.48]) and VGM (0.18 [95% CI: 0.06–0.45]).

The validity of the BILAG has been evaluated in numerous studies. In a 2010 study⁹ construct validity was assessed with the hypothesis that patients with higher scores on the BILAG presented higher levels of erythrocyte sedimentation rate (ESR), anti-dsDNA titers and SLEDAI scores, and conversely, presented lower levels of complement. The association between these variables was expected, with an OR=2.6 (1.2–4.3) for ESR>60 mm/h, OR=2.5 (1.4–3, 6) for anti-dsDNA greater than 5 times the normal value, OR=4.8 (1.4–15.1) for C3 below half the normal value, OR=4.1 (2.3–5.8) for a C4 below half the normal value and OR=215.6 (99.8–387.6) for a value of SLEDAI above 6.

Only one study has been specifically designed to evaluate the sensitivity to change of BILAG.¹⁰ It was a longitudinal multicenter study where the relationship between the change in BILAG 2004 and the change in treatment between 2 consecutive visits was evaluated. BILAG increase was associated with an increase in treatment (coefficient multinomial logistic regression: 1.35, 95% CI: 1.01–1.70). In the opposite direction, this association was also found.

The BILAG can be manually calculated, but there is a computer software known as the British Lupus Integrated Prospective System (Blips) incorporating demographic variables and clinical information necessary to calculate not only the BILAG, but other composite indices such as SLAM, SLEDAI, and SDI the SF-36.¹¹

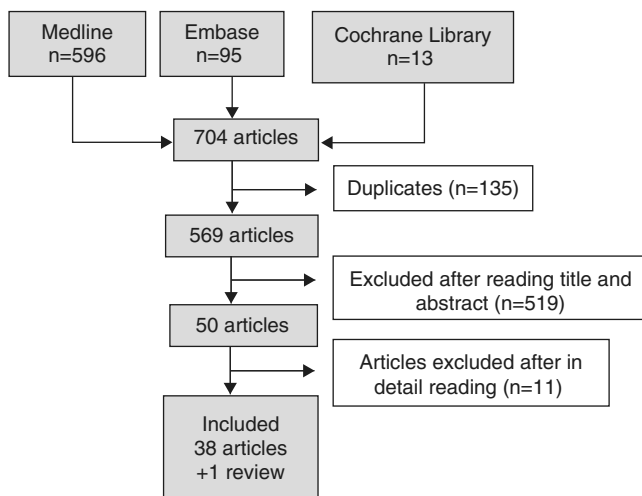
**Fig. 1.** Articles retrieved by the search strategy and results of the selection.

Table 2
Characteristics of Studies Included.

First author and year	Type of study and participants	Index	Evaluated aspects of validation	Comments
Symmons 1988 ⁴	Multicenter study (5 centers in the UK and Ireland) on symptoms that require treatment	BILAG	Face validity: the score was compared with actual treatment decisions	Description and preliminary validation
Hay 1993 ⁵	Multicenter study with 82 patients	BILAG version 3 (explanatory glossary)	Reliability; criterion validity (GS = treatment decision); construct validity (ESR, dsDNA)	Most patients inactive or less active
Stoll 1996 ⁶	Cross-sectional study of 141 patients	BILAG	Internal consistency and construct validity	BILAG vs PGA
Isenberg 2000 ¹¹	Narrative description	BILAG Software	Usage of the British Lupus Integrated Program	Presentation of software
Gordon 2003 ⁷	250 SLE patients in routine office care	BILAG to define "lupus flare"	Validity of criteria for the definition of 'flare' (GS = treatment decision)	Good validity for assessing severe flare
Isenberg 2005 ¹²	2 batches of exercises with real patients (8 patients/8 doctors)	BILAG 2004 update from previous versions	Reliability: ICC among physicians for organ/system and degree of agreement between doctors	All evaluators were members of BILAG
Yee 2007 ¹³	Multicenter cross-sectional study of 369 patients	BILAG 2004	Criterion validity and construct	GS: Treatment change
Cresswell 2009 ¹⁵	Multicenter cross-sectional study of 369 patients	Numerical scale for the BILAG	Comparison between 3 indexes using ROC curves	Cross-sectional study: ignores previous activity
Yee 2009 ¹⁰	Multicenter longitudinal study	BILAG 2004	Sensitivity to change	Difficult to differentiate activity and damage
Nasiri 2010 ⁹	Cross-sectional study of 100 patients	BILAG	Construct validity Hypothesis: if ↑↑ BILAG then ESR ↑ dsDNA and complement	Some patients with SLE have a normal serological profile
Vitali 1992 ¹⁸	Multicenter study with 704 patients	ECLAM	Face and content validity	Initial Validation
Bencivelli 1992 ¹⁹	Multicenter study of 75 patients	ECLAM	Construct validity and criterion (GS: PGA)	The PGA showed unreliability
Vitali 1999 ²¹	Comparison: manual/computerized system	ECLAM	Criterion validity	Correlations between the 2 systems
Fly 2000 ²⁰	Review of 64 medical records	ECLAM	Direct vs calculated ECLAM by reviewing medical records	Doctors knew they were to be reviewed
Petri 1992 ³⁷	150 patients with SLE. Subgroup of 6 evaluated by 9 doctors (reliability)	UCSF/JHU LAI	Reliability and construct validity (PGA)	Cross-sectional study
1992 Bombardier ²⁴	574 cases evaluated by 14 rheumatologists	SLEDAI	Initial development of the index, reliability and face validity	Initial Validation
Guzman 1992 ²⁷	39 patients in a center (3 views)	MEX-SLEDAI	Comparison with SLEDAI. Reliability and construct validity	No laboratory variables, lower cost
Chang 2002 ²⁹	Post hoc analysis of a multicenter RCT	SLEDAI and SLAM-R	Sensitivity to change	Relevant changes for doctor and patient
Gladman 2002 ²⁶	Cohort of 960 patients	SLEDAI-2K	Construct validity	Reference: SLEDAI
Uribe 2004 ³³	93 patients (3 hospitals)	SLAM-R, MEX-SLEDAI and modified SLEDAI	Construct validity and criterion (GS = SLEDAI-2K)	V. construct Ref: PGA (↑ variability)
Touma 2011 ³²	Cross-sectional study of 298 patients	Answer 50% of the SLEDAI	Face validity, content and construct	Cross-sectional study
Bae 2001 ²³	30 patients evaluated by 2 physicians/2 visits	SLAM-R	Reliability and construct validity	Physicians inexperienced with SLAM-R
2003 Karlson ³⁸	93 patients evaluated in a SLE clinic	SLAQ	Construct validity	Asymptomatic patients with abnormal lab.
Yazdany 2008 ³⁹	Observational cohort	SLAQ	Reliability, construct validity and sensitivity to change	Highly educated cohort vs general population
Gladman 1996 ⁴²	42 cases of 19 physicians	SDI	Face and content validity	Initial Validation
Gladman 1997 ⁴³	10 patients for 6–10 physicians from 5 countries	SDI	Interobserver reliability	Variability between observers through ANOVA
Stoll 1997 ¹⁴	141 patients in a single center	SDI	Internal consistency and construct validity	Compare SDI BILAG score and medication
Gladman 2000 ⁴⁵	Multicenter study in 1297 patients	SDI	Apparent validity	Evaluates association with mortality
Costenbader 2010 ⁴⁷	Multicenter study with 569 patients and 14 rheumatologists	LDIQ	Face, content, criterion and construct validity	GS: Response of doctors in the SDI

Table 2 (Continued)

First author and year	Type of study and participants	Index	Evaluated aspects of validation	Comments
Yazdany 2011 ⁴⁹ Gladman 1992 ⁵⁷	81 patients 2 university hospitals 7 patients evaluated by 4–7 rheumatologists from different countries	BILD Comparison between BILAG, SLAM and SLEDAI	Face, criterion and construct validity Construct validity	High acceptance by patients Doctors from different countries evaluate similar activity despite index
Gladman 1994 ²⁸	8 patients with 3 visits evaluated by 8 doctors	Comparison between BILAG, SLAM and SLEDAI	Construct validity and sensitivity to change	Sensitivity to change based on the average exchange rate for each center
Liang 1989 ²²	25 patients evaluated by two rheumatologists	Compare: SLAM BILAG and SLEDAI	Reliability and construct validity	The 2 reviewers from the same center
Fortin 2000 ⁴⁰ Ward 2000 ⁴¹	96 patients Prospective study of 20 patients	Comparison between SLAM and SLEDAI SLAM, BILAG, SLEDAI, LAI, and ECLAM	Sensitivity to change Construct validity and sensitivity to change	Using five different methods Low correlation between indices and PaGA
Wollaston 2004 ¹	80 cases evaluated by 20 experts in SLE (SLICC members)	SLEDAI and BILAG vs doctor's assessment	Reliability and construct validity	Comparison with PaGA, which showed great variability
Griffiths 2005 ⁵¹ Isenberg 2011 ⁸	Narrative review 16 patients evaluated by 16 rheumatologists	Most indexes described Comparison: BILAG2004, SELENA and PGA in evaluating flare	– Intraobserver reliability	– Worst agreement in moderate flare
Romero-Diaz 2011 ¹⁷	Narrative review	Description that includes most indexes	–	Updated version of 2003

BILAG: British Isles Lupus Assessment group; C: complement; ICC: intraclass correlation coefficient; ECLAM: European Consensus Lupus Activity Measurement; GS: gold standard; LDIOQ: Lupus Damage Index Questionnaire; SDI: Systemic Damage Index; SLAM: Systemic Lupus Activity Index; SLAQ: Systemic Lupus Activity Questionnaire; SLEDAI: Systemic Lupus Erythematosus Disease Activity Measure; PaGA: physician global assessment; PaGA: patient global assessment.

An updated version was published in BILAG 2005 by Isenberg et al. in an attempt to improve the characteristics of the index.¹² In this study, where the BILAG 2004 is described, its viability or ability to be applied by 2 exercises with real patients who were examined by 8 doctors belonging to the BILAG group was evaluated. The ICC was calculated for the total BILAG (0.48, 95% CI: 0.23–0.81) and for each of the 9 systems. The system that showed a higher ICC was the kidney (0.98, 95% CI: 0.96–0.99) and the one that showed a lower ICC was musculoskeletal (0.17, 95% CI: 0.01–0.56). Another measure used in this study to assess feasibility was the ratio between the standard deviation attributable to the physician and the same measure attributable to the patient, showing a high agreement in the mucocutaneous, nervous, renal, ophthalmic and hematological systems. In both exercises, the musculoskeletal system showed low reliability in terms of ICC and level of agreement among physicians.

The BILAG 2004 has been developed from the original index, and includes 9 systems: general, mucocutaneous, nervous, musculoskeletal, cardiovascular/respiratory, gastrointestinal, renal, ophthalmic and hematological system. It consists of several items or questions that are assessed on a scale of 0–4, where 0 represents “not present”, 1 means “improving”, 2 is “similar”, 3 is “worse” and 4 is “new event.” Responses were combined into 5 states of system activity: A, the patient is very active and, therefore, would require treatment with immunosuppressants, corticosteroids at medium or high doses (>20 mg prednisolone or equivalent) or anticoagulation at a high dose; B, certain activity and, therefore, need for treatment with moderate dose corticosteroid (<20 mg prednisolone), antimalarial, anti-depressants, or NSAIDs; C, little activity and need only for symptomatic treatment; D, no activity at the time, although present previously, and E, has never presented activity in this system.

Other attempts to validate the BILAG 2004 were conducted by Yee et al. in 2007. Criterion validity using as a gold standard changes in treatment and the construct validity of trying to corroborate the hypothesis that patients with higher values in the BILAG 2004 would have higher ESR, anti-dsDNA titers and lower values of complement was evaluated.¹³ BILAG 2004 values indicating high or moderate disease activity (A and B) were significantly associated with an increase in medication (OR=19.3, $p<0.01$). In addition, the BILAG 2004 had a sensitivity of 81%, a specificity of 91.9%, a positive predictive value of 56.8% and a negative predictive value of 93.6% compared to the change in treatment. Top BILAG 2004 scores were associated with increased ESR of 60 mm/h (OR=2.9), a rise in anti-dsDNA titers (OR=2.7), an elevation in the SLEDAI above 4 (OR=20) and halving of C3 (OR=5) and C4 (OR=4.2).

Although the BILAG is an ordinal scale that provides a general assessment of the disease, some authors have proposed its transformation into a numerical scale to facilitate statistical analysis. Initially, the following scheme was proposed: A=9 points, B=3 points, C=1 point and both D and E were equivalent to 0 points, with a range of 0–72.¹⁴ In 2009, Cresswell et al. presented a study aimed to validate a numerical score for the BILAG.¹⁵ They described three different scoring methods using logistic regression models and later the ROC curves were compared. The transformation finally selected was A=12, B=5, C=1 and D/E=0.

The BILAG is the only validated instrument that gives an idea of lupus activity in each organ at a glance instead of merging the information into an overall score.¹⁶ The clinician should be familiar with the glossary where each item is defined, and follow the guidelines and recommendations established by the BILAG group, especially the specific recommendations for use in RCTs and multicenter studies.^{5,17}

European Consensus Lupus Activity Measurement

The ECLAM was described by the European working group consensus for measuring activity in SLE in 1992.¹⁸ It is an index

Table 3
Excluded Studies and Reasons for Exclusion.

Study	Reasons for exclusion
Liang, 1991	Narrative description on indices validated so far
Stoll, 1996	Assessment of SLICC as a predictor of mortality
Gladman, 1999	Narrative description of the development of the SLICC/ACR Damage Index
Silvestris, 1999	Review in Italian
Ward, 2000	Scale to assess the health status desired by the patient
Mosca, 2006	Narrative review of the implications of remission in patients with SLE
Mattson, 2008	Validation of a fatigue scale in SLE
Ruperto, 2011	Delphi type survey among experts to reach a consensus on the definition of lupus flare
Lai, 2011	Validation of a fatigue scale for patients with SLE
Jolly, 2012	Instrument to measure quality of life in patients with SLE
Julian, 2012	Evaluation of a tool to assess cognitive impairment in patients with RA and SLE

designed to measure disease activity in the last month in patients with lupus.

This index was developed from a cohort that included a significant number of real patients. In an initial analysis, the index was developed and its feasibility and validity were assessed, both apparent and in the content.¹⁸ To do so, those clinical and laboratory variables that best correlate with PGA as the gold standard, both on a quantitative scale (0–10) and a qualitative scale (inactive to active) were selected. By univariate analysis, a total of 15 variables able to predict disease activity were selected. Subsequently, a multivariate analysis the weight of each of these variables, which were assigned a specific score, was defined. The ECLAM not only allows to properly classify patients according to their level of disease activity, it is also easy to calculate. In a second study, construct validity was assessed, with the gold standard being the PGA and criterion validity, comparing ECLAM with other composite indices.¹⁹ The ECLAM correlation with SLAM, BILAG and SLEDAI was greater than 0.72. These four levels are correlated similarly to the PGA.

The ECLAM includes evaluation of 10 organs and/or systems and 2 laboratory values that are ESR and complement levels. It consists of a total of 33 items that are evaluated from 0.5 to 2, depending on the type of involvement, and a combined global score ranging from 0 to 17.5.

The ECLAM is a simple index that can be used in retrospective studies, since there is a good correlation between the ECLAM calculated immediately and that calculated with data collected on clinical history ($\rho=0.871$).²⁰ Furthermore, it can be calculated by a computerized system with very similar results to the manual calculation ($r=0.90-0.92$).²¹

Systemic Lupus Activity Index

This index measures the degree of activity globally in the last month. The first publication in which it was described appeared in 1986.²² Two years later, it was revised by the residents of Harvard University, who modified the section on cardiovascular events and “others” to improve clarity and reproducibility, resulting in a new version, the SLAM-R. In 2001, Bae et al. presented a study of 30 patients to evaluate the feasibility and construct validity of this new version.²³ Reliability was estimated by analysis of variance, with 0.78 interobserver reliability and 0.61 intraobserver reliability. Regarding construct validity, Pearson correlations were calculated SLAM-R with PGA (0.87), levels of anti-dsDNA (0.51), C3 (–0.60) and C4 (–0.29).

The SLAM evaluates specific manifestations in 9 organs and includes 7 laboratory measurements. Some items are scored from 0 to 3, depending on the degree of severity, and others are evaluated only 0–1. The maximum score is 84, being part of the laboratory parameters a maximum 21 points. A score of 7 or more is considered relevant, since the patient will require a change in treatment in 50% of cases.

The SLAM is considered by some experts the less appropriate index because it includes subjective measures, such as fatigue and joint pain; however, these variables must be scored by the doctor if considered to be due to lupus activity.⁸ Training is required, since it is necessary to reach a consensus to assess the subjective aspects, especially in multicenter studies.¹⁷

Systemic Lupus Erythematosus Disease Activity Measure

The SLEDAI is a global index that was developed by an expert group of Toronto in 1986 and described in detail by Bombardier et al.²⁴ It was amended by the SELINA group for a study to evaluate the use of estrogen and progesterone in women with SLE²⁵ and later was updated by Gladman et al.²⁶ In addition, a version developed by Mexican researchers, to reduce the costs of using this index, excluded some laboratory values.²⁷

Therefore, there are currently 4 versions of this index: SLEDAI, SELINA-SLEDAI, SLEDAI 2000 and MEX-SLEDAI.

To develop the first version of SLEDAI, 24 variables that might be important factors for evaluating disease activity in SLE patients were identified. With these variables, 574 profiles were generated for potential patients and this information was provided to 14 rheumatologists considered to be SLE experts to assess the extent of disease on a scale of 0–10. Multiple regression models were used to estimate the relative importance of each of these 24 clinical variables as assessed by the experts and thus the overall index²⁴ was generated.

The SLEDAI, therefore, is a global index that evaluates the activity of the disease in the last 10 days and consists of 24 items that include specific manifestations in 9 organ systems with a maximum score of 105.

In a 1992 study by Guzman et al. various aspects of the validation of MEX-SLEDAI with the original SLEDAI²⁷ were compared. The correlation between 2 reviewers calculating these indices ranged between 0.87 and 0.89. The agreement on the assessment of the disease among physicians was moderate, with a kappa of 0.43 ($p=0.17$). Construct validity was assessed using the PGA comparator (on a scale of 0–10), with an almost identical correlation of 0.68 for both versions; the correlation between the two versions was 0.77.

SLEDAI can be used in research and clinical practice. It has proven to be a tool sensitive to change.²⁸ To assess the sensitivity to change of SLEDAI compared with SLAM-R, Chang et al. conducted a secondary analysis of data from a Canadian multicentric RCT where the effectiveness of methotrexate in patients with SLE were evaluated.²⁹ Changes were compared in these 2 indices with respect to changes in the physician and patient global evaluations. Regarding the assessment of physician, the average standardized control response (C-SRM) for the SLAM-R and SLEDAI was –0.47 vs –0.42 for improvement and 0.65 vs 0.66 for worsening. Regarding the assessment by the patient, the C-SRM for the SLAM-R and SLEDAI was –0.31 vs –0.18 for improvement and 0.48

Table 4
Aspects of validation for a Disease Activity and Damage Index.

Name of instrument	Viability Time employed, clarity and acceptance	Reliability			Validity			Sensitivity to change	Applicability
		Internal consistency: <i>Cronbach α</i>	Intrao. test-retest: ICC/Cohen kappa	Intrao. measurement error: ESM/MDC/LA	Face validity: relevant and understandable	Construct validity: 1. Structural 2. Ho Test 3. Cross-cultural	Criterion validity: gold Standard (correlations/AUC: continuous variables and S/E: dichotomous)		
<i>Indices to assess disease activity</i>									
British Isles Lupus Assessment Group (BILAG) 2004 ¹²	Accurate and complete history, 5–20 min	α=0.35 ⁵	kappa=0.79–0.97 ⁵ CCI=0.48 (95% CI: 0.23–0.81) ¹² Kappa evaluation lupus flare ⁸ BILAG=0.54 SELENA=0.21 PGA=0.18	NE	Score with actual treatment decisions was compared in a group of four patients	↑ Association ESR, dsDNA, SLEDAI and C3/C4 ↓ ⁹ PaGA rho=0.50 and rho = PGA 0.43 ⁵ ↑ ESR: OR =2.9, ↑ dsDNA: OR =2.7 SLEDAI>4: OR =20 ↓ C3: OR 5 ↓C4: OR =4.2 ¹³ Correlation with SLAM, SLEDAI and BILAG=0.72–0.78 ⁴¹	BILAG 'A' score: S=87% and E=99% PPV=80% (>dose steroids and IS) ⁵ EO/change medication: OR =19.3, P<0.01 S=81%, E=91.9%, PPV=56.8% and PPV=93.6% ¹³ GS/PGA: r=0.69 Similar to: BILAG=0.63 SLAM=0.61 SLEDAI=0.66	RME (PGA)=0.68 ⁴¹ An increase in the BILAG is associated with an increase of treatment ¹⁰	Training Multicenter RCT specific recommendations
European Consensus Lupus Activity Measurement (ECLAM) ^{18,20,41}	Needs clinical examination 5–10 minutes Simple to calculate	NE	NE	Coef. correlation =0.90–0.93 ²⁰	Variables that best correlated with the PGA			RME (PGA)=0.75 ⁴¹	Requires some training
Systemic Lupus Activity Measure (SLAM) ^{22,23,28,40} SLAM-R ²³	Accurate HC and EF 10–15 min	NE	NE	=0.86 ²²	NE	SLE indices =0.81–0.97 PGA=0.76–0.96 ²² BILAG=0.79 ⁵⁰	NE	RME (PGA)=0.62 ⁴¹ RME (LS) Best: –0.88 Worst: 0.61 ⁴⁰	Requires training
Systemic Lupus Erythematosus Disease Activity Measure (SLEDAI) ^{24,28} SLEDAI-2K ²⁶	Requires examination and analysis. 10–20 min	NE	Analysis of variance coef. correlation =0.33 ³⁷	Intrao. =0.66–0.99 Intero. =0.60–0.80 ²⁴ Intero. =0.47 ³⁷	NE	R PGA =.64–.79 ²⁴ PGA rho =0.55 ³⁷ BILAG rho =0.76 SLAM rho =0.73 ⁵⁷	GS: SLEDAI-2K ³³ : S = 73% E = 33%	RME (PGA)=0.48 ⁴¹ RME (PGA)=0.66 RME (PaGA)=0.05 ²⁹ RME (LS)=0.57 ⁴⁰	Subjective manifestations include fatigue, joint pain and myalgia
MEX-SLEDAI ²⁷	The same	NE	NE	NE	NE	rho SLEDAI=0.97 ²⁶ SLAM-R rho=0.59 PGA rho=0.68 ³³	GS: SLEDAI-2K ³³ S=58% and E=93%	NE	Method Guyatt (Video)
	Is completed in 16.9 min	NE	NE	NE	NE	PaGA rho=0.68 SLEDAI rho =0.77 ²⁷ SLAM-R rho =0.75 PGA rho=0.54 ³³			

Table 4 (Continued)

Name of instrument	Viability Time employed, clarity and acceptance	Reliability			Validity			Sensitivity to change Ability to detect changes: multiple methods	Applicability
		Internal consistency: <i>Cronbach α</i>	Intrao. test–retest: ICC/Cohen kappa	Intrao. measurement error: ESM/MDC/LA	Face validity: relevant and understandable	Construct validity: 1. Structural 2. Ho Test 3. Cross-cultural	Criterion validity: gold Standard (correlations/AUC: continuous variables and S/E: dichotomous)		
UCSF/JHU Lupus Activity Index (LAI) ³⁷	1 min	NE	Analysis of variance coef. intrao. correlation=0.81 ³⁷	Intero.=0.89 ³⁷	NE	LAI/PGA=0.64 ³⁷	NE	NE	
Systemic Lupus Activity Questionnaire ^a (SLAQ) ^{38,39}	Patient completed	$\alpha=0.87$	NE	NE	NE	SLAM (no lab) rho=0.62, $p<0.001$ PaGA rho=0.73 SF-36 rho=0.66	NE	SMR=0.12	Developed for epidemiological studies
<i>Indices to assess cumulative damage</i> Systemic Lupus International Collaborating Clinics Damage Index (SLICC-DI, SDI) ^{42,43,45,58}	Requires examination Approx. 15 min (depends on complexity)	$\alpha=0.41$ ¹⁴	ICC=0.553	10 patients and 6–10 physicians from 5 countries ⁴³	NE	BILAG rho=0.19 Treatment rho=0.33 ¹⁴	NE	NE	Follow recommendations of any member of the SLICC
Lupus Damage Index Questionnaire ^a (LDIQ) ⁴⁷	Patient completed	NE	NE	NE	Variables included in SDI	LDIQ/SDI rho=0.48	GS: SDI: S=53.3% and E=94.6%	NE	
Brief Index of Lupus Damage ^a (BILD) ⁴⁹	Patient completed	NE	NE	NE	Variables included in SDI	Demographic and clinical characteristics by quartiles	BILD/SDI rho=0.64 LDIQ/SDI rho=0.54	NE	High acceptance by patients

AUC: area under the curve; ICC: intraclass correlation coefficient; RCT: randomized clinical trial S: specificity; PE: physical examination; GS: gold standard; SEM: standard error of measurement; MH: medical history; Intero.: interobserver; Intrao.: intraobserver; LA: limits of agreement; PLR+: positive likelihood ratio; NLR–: negative likelihood ratio; LS: Likert scale; MDC: minimum detectable change; NE: not evaluated; r: Pearson correlation coefficient; rho: Spearman correlation coefficient; SRM: standardized response mean; S: sensitivity; PGA: physician global assessment; PaGA: global assessment by the patient; NPV: negative predictive value; PPV: positive predictive value.

^a Patient completed.

vs 0.05 for worsening. Only SLAM-R maintained 0 out of the 95% CI to detect improvement or deterioration. Both indices exhibit significant sensitivity to change for the physicians but only SLAM-R exhibits a sensitivity to change relevant for patients, possibly by the inclusion of more subjective variables assessed by the patient.

In 2002 a revised and updated version of the SLEDAI, SLEDAI 2000 (SLEDAI-2K), which scored rash, alopecia, ulcers or persistent proteinuria, not only recently introduced but also of new appearance, as occurred in the previous version.²⁶ The SLEDAI-2K was validated against the original SLEDAI with an evaluation within 10 days prior to the visit and it was validated with a temporal space of 30 days for use in RCTs.^{30,31} The SLEDAI-2K and the previous version had a high correlation ($r=0.97$) and both indices similarly predict mortality.²⁶

The SLEDAI-2K reflects aspects of the disease as present or absent and may not reflect partial improvement, which limits their use in RCT. For this reason, in 2011 Touma et al. developed a response rate of 50% (SRI-50) to document a 50% improvement in SLEDAI.³² This study evaluated the construct validity using assessment of response to treatment by a physician as an external comparator on a Likert scale (LS) that ranged from 7 (significant improvement) to 1 (much worse). A 50% improvement on this scale would be greater than or equal to a change of 6. In patients with an improvement ≥ 6 in the LS, there was a decrease in the SRI-50 of 4.15 ± 3.01 ($P < 0.0001$).

In 2004, Uribe et al., in an attempt to dig deeper into the validation of available versions of the SLEDAI, presented a study in which the construct validity of the SLAM-R, the MEX-SLEDAI and SLEDAI-2K was analyzed, using the PGA as the external criterion on a visual analog scale (VAS) from 0 to 10.³³ Spearman correlations of these indices with the PGA ranged from 0.54 to 0.67 for the MEX-SLEDAI and SLEDAI-2K. In an attempt to assess the criterion validity using as a gold standard the SLEDAI-2K, it was found that the sensitivity of the SLAM-R was 73%, while that of the MEX-SLEDAI was 58%. The specificity was 63% for the SLAM-R and 93% for MEX-SLEDAI.

UCSF/JHU Lupus Activity Index

The LAI was initially used in studies of serious infections and renal involvement in patients with SLE.^{34–36} This is an index with 5 domains that reflect the disease activity in the previous 2 weeks and is completed by the physician in about a minute. As in the rest of the indices, only those manifestations attributable to SLE were evaluated. It includes PGA VAS from 0 to 3, and includes 4 VAS regarding fatigue, rash, joint and serositis, and a part that quantifies the involvement in 4 organs: neurological, renal, pulmonary, and hematologic, with a VAS ranging from 0 to 3. Moreover, this last part assigns different scores according to the need for medication and laboratory values.

In a study by Petri et al.,³⁷ the reliability and construct validity of this index were evaluated using SLEDAI compared to the PGA reference. The PGA had a higher correlation with LAI ($r=0.64$) than with SLEDAI ($r=0.55$). As for reliability, LAI showed greater reliability, both intra- and interobserver.

Systemic Lupus Activity Questionnaire

The SLAQ was developed as a measurement tool for the patient to be used in epidemiological studies and large patient cohorts. It was described and validated by Karlson et al.³⁸ It was developed in a clinical cohort of 93 patients and compared with SLAM (excluding laboratory data), showing a good correlation ($r=0.62$, $P < 0.001$). Each individual item correlation between patients–physicians ranged from 0.06 for the evaluation of lymphadenopathy and vasculitis and 0.7 for evaluation of Raynaud's syndrome.

Subsequently Yazdany et al. evaluated other aspects of validation in a cohort of 982 patients³⁹. The SLAQ showed good internal

consistency, with a Cronbach α 0.87. To assess construct validity, we examined the correlation of SLAQ with PGA ($r=0.73$) and the SF-36 ($r=0.66$). The index showed a low sensitivity to change with an SRM of 0.12, but when it was compared with the change in other measures, the SLAQ score changed in the right direction.

Studies Comparing Rates

The first study that compares several composite indexes was presented in 1989 by Liang et al.²² In this study the reliability of the SLAM, SLEDAI and BILAG, and their construct validity is discussed; on the one hand, the correlation between each index (which ranged between 0.81 and 0.97) was calculated and, secondly, calculating the correlation of each of them with a PGA VAS 0–10 ($r=0.76$ – 0.96). In an attempt to assess not only the construct validity, but also the sensitivity to change, Gladman et al. designed a study in 8 patients with 3 visits per patient and evaluated by 8 rheumatologists.²⁸ The correlation between these three indices ranged from 0.35 to 0.61. Regarding the sensitivity to change, only the SLEDAI was able to differentiate between visits by analysis of variance. These results are based on the average rate change for each index and therefore the clinical relevance of this change observed in each patient is difficult to assess. A few years later, Fortin et al. conducted a study to assess the sensitivity to change using different methods and using as an external criterion the change assessment made by the physician in a LS of 5 points, ranging from “much better” to “much worse”.⁴⁰ The 3 indices were shown to be sensitive to change, SLAM performing best. The authors interpret that this small difference may be due to the SLAM allowing increasing values according to severity while the SLEDAI has a fixed score for each item.

In another study of the same year, Ward et al. evaluated the construct validity and sensitivity to change of the SLAM, SLEDAI, LAI, BILAG and ECLAM.⁴¹ All these indices were valid for measuring disease activity in patients with SLE. The SLAM best captured the patient's perception ($r=0.22$, $P < 0.0001$). The correlation between the change in each of the indices and the change in the health evaluation were in decreasing order: LAI $r=0.75$; R ECLAM $=0.65$; R BILAG $=0.61$; SLAM $r=0.54$ and r SLEDAI $=0.52$, all with $P < 0.0001$. The index that showed greater sensitivity to change measured by the SRM using and using the PGA as reference was the LAI (SRM $=0.74$) and to a lesser extent, the SLEDAI (SRM $=0.48$).

In an attempt to develop an index to assess the response to treatment, in 2004, Wollaston et al. compared the change in two validated indices, the BILAG and SLEDAI, with the change in disease activity evaluated by a skilled practitioner in lupus using LS of 7 points (1 = important improvement and 7 = worsening with respect to baseline).¹ The ICC of the physicians evaluation in the 4 groups of evaluators ranged between 0.25 and 0.46. The change in baseline to 3 months of SLEDAI correlated well with the change in BILAG ($r=0.75$, 95% CI: 0.63–0.83). The conclusion of this study was that, while the composite indices are comparable between themselves, there exists a large variability in the assessment made by physicians, even among experienced physicians with lupus patients.

Indices to Assess Damage in Systemic Lupus Erythematosus

Systemic Lupus International Collaborating Clinics/American College of Rheumatology-Damage Index

It was developed in 1996 by an international collaboration (SLICC group) and adopted by the CAR.⁴² In this first study, the index was developed and its content validity was assessed. A list of variables that could reflect damage in patients with SLE was generated and a consensus was reached on which variables should be included in an index generated to assess irreversible damage. In another study of the same year and by the same authors, interobserver

reliability was assessed with 10 patients who rheumatologists from 5 different countries applied this index to.⁴³ The authors concluded in this study that doctors in different countries assess damage in patients with SLE very similarly.

The SDI evaluates irreversible damage in lupus patients regardless of the cause that produces it. It includes 42 items measuring involvement in 12 domains, with a maximum score of 46 points. Each item is scored as present or absent with the possibility of scoring a 2 or 3 in case of recurring events, such as may be the case of a stroke.

The definition of damage in patients with SLE is an irreversible change in an organ or system that has occurred since the onset of SLE and is present for at least the last 6 months. At the time of diagnosis, the score should be 0 by definition.

The SDI is completed by the physician at baseline in order to enter an RCT. It has moderate internal consistency (Cronbach $\alpha=0.41$). In an attempt to assess construct validity, Stoll et al. compared SDI with BILAG and a patient's medication score, both with very low correlation (0.19 BILAG and 0.33 for the medication index).¹⁴

When SDI is completed by another physician to retrospectively review the patient's history, it also shows good interobserver reliability.⁴⁴ SDI values increase with disease progression similarly in patients from different countries and is also a tool that predicts mortality in patients with SLE.^{45,46}

Lupus Damage Index Questionnaire

The LDIQ was described by Costenbader et al. in an attempt to develop a measure of irreversible damage based on the SDI but completed by the patient to be used in clinical practice or epidemiological studies.⁴⁷ This study is a first step in the development of an index, evaluating the apparent content validity with a group of 37 patients and 7 rheumatologists, with a more elaborate multicentric level validation later. To assess criterion validity, SDI is used as the gold standard, resulting in a sensitivity of 53.3%, a specificity of 94.6% and a degree of agreement between the 2 measures of 93.2%. To assess construct validity, the correlation between LDIQ and SDI (Spearman correlation coefficient =0.48, $P<0.001$) was calculated.

The LDIQ consists of 56 questions to assess each domain included in the SDI and is designed to be administered as a survey. Versions in Spanish, Portuguese and French have been validated.⁴⁸

Brief Index of Lupus Damage

The BILD was described in 2011 by Yazdany et al. as a measure of damage in SLE patients for use in population studies.⁴⁹ Like LDIQ, it is completed by the patient but is a shorter version that includes only 28 questions. BILD criterion validity was examined through its correlation with other instrument that also measured damage, such as SDI, and can be considered as the gold standard ($r=0.64$, $P<0.001$). Construct validity was assessed by comparing the demographic and clinical characteristics of the patients divided into the 4 quartiles of BILD. Patients with higher values for the BILD were older, with longer duration of illness and greater disease activity.

Discussion

To measure disease activity and irreversible damage in patients with SLE, there is a need for quantitative composite measures that have an acceptable validation. Designing these composite measures is a challenge, given the multisystem involvement in SLE and its extensive variability. At present, there are multiple tools with varying degrees of validation, although none has been accepted as the only measure recommended internationally. Choosing the most appropriate measure in each case mainly depends on the context in which it will be used and the question one wants to answer in terms of disease assessment.

Broadly speaking, there are 2 types of activity indices, those which function as global measures (such as ECLAM, SLAM, SLEDAI and LAI) and those that give a specific score for each organ system (BILAG).

The global indices are considered useful to compare cohorts of patients with SLE, as they are simpler to implement and therefore more useful in population studies. They can also be used to define inclusion criteria for RCTs or response criteria for treatment. But these indices do not give information on the degree of activity in a specific organ, for which indices of organ/system would be most useful.

In terms of implementation, most of these indices are complex and require information from the clinical history and physical examination. The index that takes longer to implement is the BILAG, and the LAI and SLAQ are the simplest to apply. Internal consistency was assessed only in the BILAG and SLAQ, being higher in the latter. Regarding the interobserver reliability, the index showed a higher ICC was the SLEDAI (ICC =0.79). The interobserver reliability was assessed in a very heterogeneous way across studies, making it difficult to compare the indices. This aspect of validation is not required in the indexes answered by the patient.

Construct validity was the most studied part of the validation for most indices, with similar results. It is difficult to establish a gold standard as a reference to assess disease activity. To assess criterion validity, in some studies the gold standard was change of treatment and, in others, the PGA. Use of the PGA as a gold standard has certain problems, since several studies have shown that there is a very low degree of agreement between doctors when assessing the disease activity in patients with SLE.^{1,27} Other studies used as gold standard another index already described in the literature, with which the new index was compared and evaluated. Most indices correlate acceptably with each other.²²

Finally, the index that shows a greater sensitivity to change is the ECLAM.

Some important aspects to consider is that the activity indexes should only rate aspects that are directly related to the disease. The SLAM and BILAG collected manifestations that appear in the last month, while others collected manifestations in the last 10 days. Some indexes include laboratory variables and it is important to note that there are patients with active SLE who have no laboratory abnormalities. For example, only 60% of patients have positive anti-dsDNA antibodies during the course of the illness.⁵⁰

A critique of these indices could be that most of them have been validated in the context of long-term cohorts instead of RCT.⁵¹ Even so, in 1998, one of the OMERACT objectives was to preliminarily define minimum variables to be included in the evaluation of patients with SLE, both for RCTs and longitudinal cohort of long-term activity, and these were illness, injury, quality of life and toxicity/adverse effects.⁵²

The concept of damage in SLE has become an important outcome measure, as it not only predicts mortality also functional capacity and health resource utilization.^{53–55} In this review, three measures of injury have been identified, one that is completed by the doctor, the SDI, and 2 to be completed by the patient LDIQ and BILD. The 2 self-administered measures of damage showed a good correlation with the extent to which the physician measured was considered the gold standard.

In conclusion, most composite indices to assess disease activity in SLE patients have been validated and are comparable. Although complex, when applied in clinical practice, these indices facilitate the collection of relevant clinical information quantitatively and EULAR recommends their use to monitor patients.⁵⁶ The use of composite indices to assess both the activity and structural damage would assist in clinical practice to guide therapeutic decisions as objectively as possible.⁵¹ Furthermore, the PGA exhibits great variability, even among physicians with extensive experience

treating patients with SLE, not making it an ideal tool to assess activity.

Both the BILAG and SLEDAI are the indexes that have a more complete validation and are most commonly used in RCTs and cohort studies. The ECLAM has the advantage of being the easiest to calculate and SLAM is the most sensitive to change, since it is the only one that can give a greater score to high severity. The SLAQ is a patient answered index that exhibits good correlation with SLAM, from which it derives, so it may be a good alternative for patient visit in which there is less time to spend per patient or in population studies.

Ethical Responsibilities

Protection of people and animals. The authors declare that no experiments were performed on humans or animals for this investigation.

Data confidentiality. The authors declare that no patient data appears in this article.

Right to privacy and informed consent. The authors have obtained the informed consent of the patients and/or subjects mentioned in the article. The author for correspondence is in possession of this document.

Financing

This study was funded by the Spanish Foundation of Rheumatology.

Conflict of Interest

Iñigo Rúa-Figueroa has had contracts with GSK and has attended courses/conferences sponsored by GSK and MSD.

The rest of the authors have no disclosures to make.

Appendix. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.reumae.2014.01.011>.

References

- Wollaston SJ, Farewell VT, Isenberg DA, Gordon C, Merrill JT, Petri MA, et al. Defining response in systemic lupus erythematosus: a study by the Systemic Lupus International Collaborating Clinics group. *J Rheumatol.* 2004;31:2390–4.
- Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982;25:1271–7.
- Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;40:1725.
- Symmons DP, Coppock JS, Bacon PA, Bresnihan B, Isenberg DA, Maddison P, et al. Development and assessment of a computerized index of clinical disease activity in systemic lupus erythematosus. Members of the British Isles Lupus Assessment Group (BILAG). *Q J Med.* 1988;69:927–37.
- Hay EM, Bacon PA, Gordon C, Isenberg DA, Maddison P, Snaith ML, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. *Q J Med.* 1993;86:447–58.
- Stoll T, Stucki G, Malik J, Pyke S, Isenberg DA. Further validation of the BILAG disease activity index in patients with systemic lupus erythematosus. *Ann Rheum Dis.* 1996;55:756–60.
- Gordon C, Sutcliffe N, Skan J, Stoll T, Isenberg DA. Definition and treatment of lupus flares measured by the BILAG index. *Rheumatology (Oxford).* 2003;42:1372–9.
- Isenberg DA, Allen E, Farewell V, D'Cruz D, Alarcon GS, Aranow C, et al. An assessment of disease flare in patients with systemic lupus erythematosus: a comparison of BILAG 2004 and the flare version of SELENA. *Ann Rheum Dis.* 2011;70:54–9.
- Nasiri S, Karimifard M, Bonakdar ZS, Salehi M. Correlation of ESR, C3, C4, anti-DNA and lupus activity based on British Isles Lupus Assessment Group Index in patients of rheumatology clinic. *Rheumatol Int.* 2010;30:1605–9.
- Yee CS, Farewell V, Isenberg DA, Griffiths B, Teh LS, Bruce IN, et al. The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. *Rheumatology (Oxford).* 2009;48:691–5.
- Isenberg DA, Gordon C. From BILAG to BLIPS—disease activity assessment in lupus past, present and future. *Lupus.* 2000;9:651–4.
- Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology (Oxford).* 2005;44:902–6.
- Yee CS, Farewell V, Isenberg DA, Rahman A, Teh LS, Griffiths B, et al. British Isles Lupus Assessment Group 2004 index is valid for assessment of disease activity in systemic lupus erythematosus. *Arthritis Rheum.* 2007;56:4113–9.
- Stoll T, Stucki G, Malik J, Pyke S, Isenberg DA. Association of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index with measures of disease activity and health status in patients with systemic lupus erythematosus. *J Rheumatol.* 1997;24:309–13.
- Cresswell L, Yee CS, Farewell V, Rahman A, Teh LS, Griffiths B, et al. Numerical scoring for the Classic BILAG index. *Rheumatology (Oxford).* 2009;48:1548–52.
- Ramsey-Goldman R, Isenberg DA. Systemic lupus erythematosus measures: British Isles Lupus Assessment Group (BILAG), European Consensus Lupus Activity Measurement (ECLAM), Systemic Lupus Activity Measure (SLAM), Systemic Lupus Erythematosus Disease Activity Measure (SLEDAI), and Systemic Lupus International Collaborating Clinics/American College of Rheumatology-Damage Index (SLICC/ACR-DI; SDI). *Arthritis Care Res.* 2003;49:S225–33.
- Romero-Diaz J, Isenberg D, Ramsey-Goldman R. Measures of adult systemic lupus erythematosus: updated version of British Isles Lupus Assessment Group (BILAG 2004), European Consensus Lupus Activity Measurements (ECLAM), Systemic Lupus Activity Measure, Revised (SLAM-R), Systemic Lupus Activity Questionnaire for Population Studies (SLAQ), Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K), and Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SDI). *Arthritis Care Res (Hoboken).* 2011;63 Suppl. 11:S37–46.
- Vitali C, Bencivelli W, Isenberg DA, Smolen JS, Snaith ML, Sciuto M, et al. Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research. II. Identification of the variables indicative of disease activity and their use in the development of an activity score. *The European Consensus Study Group for Disease Activity in SLE. Clin Exp Rheumatol.* 1992;10:541–7.
- Vitali C, Bencivelli W, Isenberg DA, Smolen JS, Snaith ML, Sciuto M, et al. Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research. III. Development of a computerised clinical chart and its application to the comparison of different indices of disease activity. *The European Consensus Study Group for Disease Activity in SLE. Clin Exp Rheumatol.* 1992;10:549–54.
- Mosca M, Bencivelli W, Vitali C, Carrai P, Neri R, Bombardieri S. The validity of the ECLAM index for the retrospective evaluation of disease activity in systemic lupus erythematosus. *Lupus.* 2000;9:445–50.
- Vitali C, Bencivelli W, Mosca M, Carrai P, Sereni M, Bombardieri S. Development of a clinical chart to compute different disease activity indices for systemic lupus erythematosus. *J Rheumatol.* 1999;26:498–501.
- Liang MH, Socher SA, Larson MG, Schur PH. Reliability and validity of six systems for the clinical assessment of disease activity in systemic lupus erythematosus. *Arthritis Rheum.* 1989;32:1107–18.
- Bae SC, Koh HK, Chang DK, Kim MH, Park JK, Kim SY. Reliability and validity of systemic lupus activity measure-revised (SLAM-R) for measuring clinical disease activity in systemic lupus erythematosus. *Lupus.* 2001;10:405–9.
- Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. *Arthritis Rheum.* 1992;35:630–40.
- Petri M, Buyon J, Skovron ML, Kim M. Disease activity and health status (SF-36) in postmenopausal systemic lupus erythematosus: the SELENA trial. *Arthritis Rheum.* 1987;40 [abstract].
- Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol.* 2002;29:288–91.
- Guzman J, Cardiel MH, Arce-Salinas A, Sanchez-Guerrero J, Alarcon-Segovia D. Measurement of disease activity in systemic lupus erythematosus. Prospective validation of 3 clinical indices. *J Rheumatol.* 1992;19:1551–8.
- Gladman DD, Goldsmith CH, Urowitz MB, Bacon P, Bombardier C, Isenberg D, et al. Sensitivity to change of 3 systemic lupus erythematosus disease activity indices: international validation. *J Rheumatol.* 1994;21:1468–71.
- Chang E, Abrahamowicz M, Ferland D, Fortin PR. Comparison of the responsiveness of lupus disease activity measures to changes in systemic lupus erythematosus activity relevant to patients and physicians. *J Clin Epidemiol.* 2002;55:488–97.
- Touma Z, Urowitz MB, Ibanez D, Gladman DD. SLEDAI-2K 10 days versus SLEDAI-2K 30 days in a longitudinal evaluation. *Lupus.* 2011;20:67–70.
- Touma Z, Urowitz MB, Gladman DD. SLEDAI-2K for a 30-day window. *Lupus.* 2010;19:49–51.
- Touma Z, Urowitz MB, Fortin PR, Landolt C, Toloza SM, Riddell C, et al. Systemic lupus erythematosus disease activity index 2000 responder index-50: a reliable index for measuring improvement in disease activity. *J Rheumatol.* 2011;38:868–73.

33. Uribe AG, Vila LM, McGwin Jr G, Sanchez ML, Reveille JD, Alarcon GS. The Systemic Lupus Activity Measure-revised, the Mexican Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), and a modified SLEDAI-2K are adequate instruments to measure disease activity in systemic lupus erythematosus. *J Rheumatol.* 2004;31:1934–40.
34. Ratain JS, Petri M, Hochberg MC, Hellmann DB. Accuracy of creatinine clearance in measuring glomerular filtration rate in patients with systemic lupus erythematosus without clinical evidence of renal disease. *Arthritis Rheum.* 1990;33:277–80.
35. Petri M, Bockenstedt L, Colman J, Whiting-O'Keefe Q, Fitz G, Sebastian A, et al. Serial assessment of glomerular filtration rate in lupus nephropathy. *Kidney Int.* 1988;34:832–9.
36. Hellmann DB, Petri M, Whiting-O'Keefe Q. Fatal infections in systemic lupus erythematosus: the role of opportunistic organisms. *Medicine (Baltimore).* 1987;66:341–8.
37. Petri M, Hellmann D, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. *J Rheumatol.* 1992;19:53–9.
38. Karlson EW, Daltroy LH, Rivest C, Ramsey-Goldman R, Wright EA, Partridge AJ, et al. Validation of a Systemic Lupus Activity Questionnaire (SLAQ) for population studies. *Lupus.* 2003;12:280–6.
39. Yazdany J, Yelin EH, Panopalis P, Trupin L, Julian L, Katz PP. Validation of the systemic lupus erythematosus activity questionnaire in a large observational cohort. *Arthritis Rheum.* 2008;59:136–43.
40. Fortin PR, Abrahamowicz M, Clarke AE, Neville C, Du Berger R, Fraenkel L, et al. Do lupus disease activity measures detect clinically important change? *J Rheumatol.* 2000;27:1421–8.
41. Ward MM, Marx AS, Barry NN. Comparison of the validity and sensitivity to change of 5 activity indices in systemic lupus erythematosus. *J Rheumatol.* 2000;27:664–70.
42. Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Urowitz M, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum.* 1996;39:363–9.
43. Gladman DD, Urowitz MB, Goldsmith CH, Fortin P, Ginzler E, Gordon C, et al. The reliability of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index in patients with systemic lupus erythematosus. *Arthritis Rheum.* 1997;40:809–13.
44. Thumboo J, Lee HY, Fong KY, Chan SP, Chapman CA, Leong KH, et al. Accuracy of medical record scoring of the SLICC/ACR damage index for systemic lupus erythematosus. *Lupus.* 2000;9:358–62.
45. Gladman DD, Goldsmith CH, Urowitz MB, Bacon P, Fortin P, Ginzler E, et al. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) damage index for systemic lupus erythematosus international comparison. *J Rheumatol.* 2000;27:373–6.
46. Rahman P, Gladman DD, Urowitz MB, Hallett D, Tam LS. Early damage as measured by the SLICC/ACR damage index is a predictor of mortality in systemic lupus erythematosus. *Lupus.* 2001;10:93–6.
47. Costenbader KH, Khamashta M, Ruiz-Garcia S, Perez-Rodriguez MT, Petri M, Elliott J, et al. Development and initial validation of a self-assessed lupus organ damage instrument. *Arthritis Care Res (Hoboken).* 2010;62:559–68.
48. Pons-Estel BA, Sanchez-Guerrero J, Romero-Diaz J, Iglesias-Gamarra A, Bonfa E, Borba EF, et al. Validation of the Spanish, Portuguese and French versions of the Lupus Damage Index questionnaire: data from North and South America, Spain and Portugal. *Lupus.* 2009;18:1033–52.
49. Yazdany J, Trupin L, Gansky SA, Dall'era M, Yelin EH, Criswell LA, et al. Brief index of lupus damage: a patient-reported measure of damage in systemic lupus erythematosus. *Arthritis Care Res (Hoboken).* 2011;63:1170–7.
50. Ho A, Magder LS, Barr SG, Petri M. Decreases in anti-double-stranded DNA levels are associated with concurrent flares in patients with systemic lupus erythematosus. *Arthritis Rheum.* 2001;44:2342–9.
51. Griffiths B, Mosca M, Gordon C. Assessment of patients with systemic lupus erythematosus and the use of lupus disease activity indices. *Best Pract Res Clin Rheumatol.* 2005;19:685–708.
52. Strand V, Gladman D, Isenberg D, Petri M, Smolen J, Tugwell P. Endpoints: consensus recommendations from OMERACT IV. Outcome measures in rheumatology. *Lupus.* 2000;9:322–7.
53. Alarcon GS, McGwin Jr G, Bastian HM, Roseman J, Lisse J, Fessler BJ, et al. Systemic lupus erythematosus in three ethnic groups. VII [correction of VIII]. Predictors of early mortality in the LUMINA cohort. LUMINA Study Group. *Arthritis Rheum.* 2001;45:191–202.
54. Clarke AE, Petri M, Manzi S, Isenberg DA, Gordon C, Senecal JL, et al. The systemic lupus erythematosus Tri-nation Study: absence of a link between health resource use and health outcome. *Rheumatology (Oxford).* 2004;43:1016–24.
55. Bertoli AM, Fernandez M, Alarcon GS, Vila LM, Reveille JD. Systemic lupus erythematosus in a multiethnic US cohort LUMINA (XLI): factors predictive of self-reported work disability. *Ann Rheum Dis.* 2007;66:12–7.
56. Mosca M, Tani C, Aringer M, Bombardieri S, Boumpas D, Brey R, et al. European League Against Rheumatism recommendations for monitoring patients with systemic lupus erythematosus in clinical practice and in observational studies. *Ann Rheum Dis.* 2010;69:1269–74.
57. Gladman DD, Goldsmith CH, Urowitz MB, Bacon P, Bombardier C, Isenberg D, et al. Crosscultural validation and reliability of 3 disease activity indices in systemic lupus erythematosus. *J Rheumatol.* 1992;19:608–11.
58. Stoll T, Seifert B, Isenberg DA. SLICC/ACR Damage Index is valid, and renal and pulmonary organ scores are predictors of severe outcome in patients with systemic lupus erythematosus. *Br J Rheumatol.* 1996;35:248–54.