

## Revisión

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



Isabel Castrejón<sup>a,\*</sup>, Iñigo Rúa-Figueroa<sup>b</sup>, María Piedad Rosario<sup>c</sup> y Loreto Carmona<sup>d</sup>

<sup>a</sup> Departamento de Medicina, División de Reumatología, New York University School of Medicine, NYU Hospital for Joint Diseases, Nueva York, Estados Unidos

<sup>b</sup> Departamento de Reumatología, Hospital Dr. Negrín, Las Palmas de Gran Canaria, España

<sup>c</sup> Unidad de investigación, Sociedad Española de Reumatología, Madrid, España

<sup>d</sup> Universidad Camilo Jose Cela. Instituto de Salud Musculoesquelética, Madrid, España

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 9 de agosto de 2013

Aceptado el 31 de enero de 2014

On-line el 10 de julio de 2014

### Palabras clave:

Validación

Lupus eritematoso sistémico

Actividad de la enfermedad

Daño

Revisión sistemática

## R E S U M E N

**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.

**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  $\alpha$  de Cronbach de 0,35 a 0,87. La fiabilidad intraobservador fue examinada mediante el coeficiente de correlación intraclass 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 medico, 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.

## Clinical composite measures of disease activity and damage used to evaluate patients with systemic lupus erythematosus: A systematic literature review

### A B S T R A C T

**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.

### Keywords:

Validation

Systemic lupus erythematosus

Disease activity

Damage

Systematic review

\* Autor para correspondencia.

Correos electrónicos: [isabelcastrejonf@gmail.com](mailto:isabelcastrejonf@gmail.com) (I. Castrejón), [iruafer@gobiernodecanarias.org](mailto:iruafer@gobiernodecanarias.org) (I. Rúa-Figueroa), [mariapiedad.rosario@ser.es](mailto:mariapiedad.rosario@ser.es) (M.P. Rosario), [loretocarmo@gmail.com](mailto:loretocarmo@gmail.com) (L. Carmona).

**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, LIDIQ and the BILD were self-administered. Feasibility and internal consistency was only studied in 3 indices (BILAG, SLAQ and SDI) with a Cronbach's  $\alpha$  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.

## Introducción

Medir la actividad de la enfermedad y el daño irreversible en pacientes con lupus eritematoso sistémico (LES) es de vital importancia para evaluar las medidas de desenlace del paciente y su pronóstico, las diferencias entre grupos de pacientes y las respuestas a nuevos tratamientos. Existen diferentes índices validados, disponibles desde principios de los 80, aunque ninguno ha demostrado una clara superioridad y, por tanto, no tienen una aceptación universal.

Evaluar a los pacientes con LES supone un reto por diferentes motivos. Por un lado, la compleja afectación multisistémica que produce la enfermedad y, por otro, su curso fluctuante, que hace que los pacientes tengan una evolución y una afectación muy diferentes entre sí. Además, la ausencia de un estándar de oro o variable única que permita evaluar la actividad de la enfermedad hace precisa la utilización de índices o escalas compuestas. La valoración global de la enfermedad por el médico (VGM) puede estar muy influida por la experiencia clínica del médico y, por tanto, producir una amplia variabilidad interobservador que dificulta la comparación entre pacientes<sup>1</sup>.

El objetivo de esta revisión es evaluar los aspectos de la validación de los índices empleados para medir tanto la actividad de la enfermedad como el daño acumulado en pacientes con LES.

## Métodos

Como parte del consenso de la Sociedad Española de Reumatología para el uso de terapias biológicas en pacientes con LES, se realizó una revisión sistemática para examinar la validez de los índices compuestos utilizados para evaluar actividad de la enfermedad y el daño acumulado.

### Estrategia de búsqueda

Se diseñó una estrategia de búsqueda en las siguientes bases de datos bibliográficas: MEDLINE, EMBASE y la biblioteca central Cochrane hasta marzo del 2012. La búsqueda incluyó términos MeSH y texto libre. Las estrategias de búsqueda para MEDLINE, EMBASE y la biblioteca Cochrane han sido incluidas en el [anexo 1 \(disponible en la web\)](#). La búsqueda se limitó a estudios en seres humanos y a estudios publicados en inglés, francés y español. Además, se realizó una búsqueda manual en la bibliografía de los artículos incluidos.

La búsqueda se limitó a estudios en seres humanos y a estudios publicados en inglés, francés.

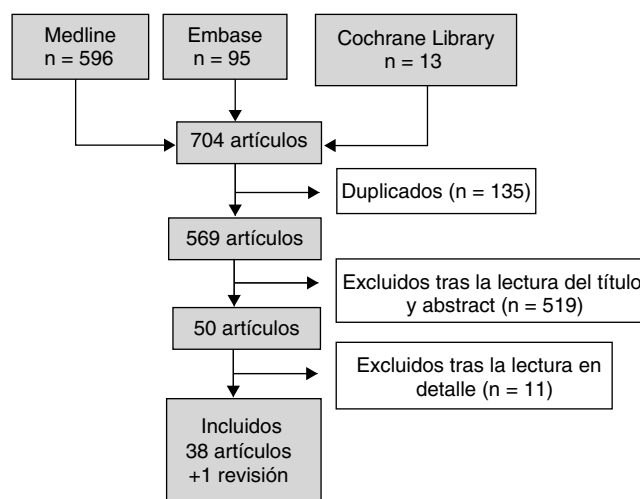
### Selección de estudios

Se incluyeron estudios de validación, cohortes, metaanálisis y revisiones sistemáticas. Respecto al tipo de participantes, se seleccionaron estudios con pacientes con una edad igual o superior a 18 años y con diagnóstico de LES siguiendo los criterios del colegio americano de reumatología (CAR)<sup>2,3</sup>. Se consideró intervención cualquier índice, instrumento o escala empleado para evaluar actividad de la enfermedad o daño estructural. El comparador fue cualquier índice comparado consigo mismo o con otros índices. Por último, las medidas de desenlace evaluadas fueron aspectos de la validación como sensibilidad, especificidad, viabilidad, fiabilidad, validez y sensibilidad al cambio. En la [tabla 1](#) se incluye una información más detallada sobre los aspectos de validación para una fácil interpretación de los resultados.

Un revisor (IC) realizó la selección de los artículos por título y resumen, siguiendo los criterios de inclusión. En la [tabla 2](#) se exponen las características de los estudios incluidos. La información referente a la validación de cada índice compuesto fue recogida usando formularios estandarizados. Se generaron tablas de evidencia para resumir cada aspecto de la validación, así como el método empleado para su evaluación.

## Resultados

En el diagrama de la [figura 1](#) se detallan los resultados de la búsqueda. En la estrategia de búsqueda se identificaron un total



**Figura 1.** Artículos recuperados gracias a la estrategia de búsqueda y resultado de la selección.

**Tabla 1**  
Guía para interpretar los resultados sobre los aspectos de la validación

Término	Aspectos	Técnica de análisis
Viabilidad (feasibility)	Tiempo empleado. Claridad de elementos (sencillo). Aceptado por pacientes y usuarios	Estudio piloto (30 pacientes)
Fiabilidad (reliability). Grado con el que el instrumento mide con precisión: sin error. Fiable, preciso y libre de error (error sistemático/sesgo y error aleatorio)	Consistencia interna: evalúa si los ítems que miden un mismo atributo presentan homogeneidad entre ellos. Depende del número de ítems y su correlación entre ellos F. intraobservador o test-retest; mide la estabilidad de las puntuaciones otorgadas por el mismo evaluador en los mismos sujetos F. interobservador o error de medición: mide el grado de acuerdo entre 2 o más evaluadores	$\alpha$ de Cronbach (0-1): se interpreta como un coeficiente de correlación  CCI (coef. de correlación intraclass): cuantitativas kappa de Cohen: cualitativas
Validez (validity). Capacidad para medir aquello para lo que se ha diseñado	Lógica o aparente (face validity): grado en que parece que un índice mide lo que quiere medir De contenido: si contempla todos los aspectos relacionados con el concepto de estudio. Muestra representativa de los ítems De constructo: refleja el concepto a medir o la capacidad del instrumento para medir adecuadamente V. de criterio: comparando con un método de referencia ya descrito y validado	Error estándar de medida, mínimo cambio detectable límites de acuerdo Opinión de expertos sobre la relevancia y comprensibilidad: redacción de preguntas Panel de expertos  Estructural: análisis factorial. Test de Ho: correlaciones. Validación transcultural
Sensibilidad al cambio (responsiveness)	Capacidad para medir cambios. Intrínseca y extrínseca	Estándar de oro V. continuas: correlaciones o curvas ROC. V. dicotómicas: sensibilidad y especificidad Múltiples métodos

de 704 artículos, de los cuales 50 fueron seleccionados para una revisión en detalle tras la exclusión de 519 referencias por título y resumen. Un total de 38 artículos fueron incluidos y 11 fueron excluidos al no cumplir los criterios de inclusión. En la [tabla 3](#) se detallan los estudios excluidos y los motivos de su exclusión.

Los artículos incluidos abordan la validación de 6 índices compuestos de actividad de la enfermedad: 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) y Systemic Lupus Activity questionnaire (SLAQ), y 3 índices para evaluar daño acumulado: Systemic Lupus International Collaborating Clinics/American College of Rheumatology-Damage Index (SLICC/ACE-DI [SDI]), Lupus Damage Index Questionnaire (LDIQ) y Brief Index of Lupus Damage (BILD).

A continuación, se revisarán en detalle las características de cada índice, así como su validación. En la [tabla 4](#) se resumen los aspectos de la validación de todos estos índices para permitir una sencilla comparación entre ellos.

### Índices para evaluar actividad de la enfermedad

#### British Isles Lupus Assessment group

El grupo de las Islas británicas para la valoración del lupus (BILAG) se empezó a reunir de forma regular en 1984 y en 1988 elaboró, por primera vez, un índice para medir la actividad de la enfermedad en pacientes con LES<sup>4</sup>. Este índice fue desarrollado basándose en la intención de tratar por el médico y evalúa manifestaciones específicas que precisan tratamiento en un total de 8 órganos o aparatos: general, mucocutáneo, neurológico, musculoesquelético, cardiorrespiratorio, vasculitis, renal y hematológico. Al contrario de otros índices que proporcionan una valoración global de la enfermedad, el BILAG proporciona una valoración por órgano o sistema.

Posteriormente, se realizaron ciertas modificaciones menores de este índice y se evaluaron tanto su fiabilidad como su validez. En un estudio de Hay et al. de 1993<sup>5</sup>, se evalúa la versión 3 del BILAG que incluye un glosario con recomendaciones y explicaciones para facilitar su uso. El BILAG mostró una buena fiabilidad intraobservador, con un kappa que osciló entre 0,79 y 0,97, dependiendo del

sistema evaluado. Para evaluar la validez de criterio el BILAG, se comparó con un estándar de oro definido como el inicio o aumento de tratamiento, bien con prednisona u otro tratamiento inmunosupresor. La sensibilidad del BILAG para el grado A en cualquier órgano o sistema fue del 87% y la especificidad del 99%. El coeficiente  $\alpha$  de Cronbach del BILAG, calculado para evaluar la consistencia interna y, por tanto, la asociación entre los componentes del BILAG, es de 0,35, por debajo de lo que se recomienda para comparaciones fiables. El BILAG mostró una correlación moderada con la evaluación global del médico ( $\rho = 0,43$ ) y con la valoración del paciente ( $\rho = 0,50$ )<sup>6</sup>.

El BILAG también se utiliza para evaluar la aparición de brotes en pacientes con lupus. La presencia de un brote severo se define como una puntuación de A de nueva aparición y un brote moderado se define con una puntuación de B, si previamente era de D o E. Un estudio publicado en 2003 por Gordon et al. evaluó el grado de acuerdo entre la definición de brote mediante BILAG y la actitud real en cuanto a tratamiento de los reumatólogos en práctica clínica rutinaria<sup>7</sup>. El 92% de los pacientes con un brote severo (A) recibió una intensificación del tratamiento por su médico y tan solo el 41% en el caso de brote moderado (B). No solo el BILAG es útil para evaluar la presencia de un brote lúpico. En un estudio de Isenberg et al. se compararon 3 métodos para evaluar el brote: BILAG 2004, un índice de brote generado para el Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA), denominado SELENA.SLEDAI Flare Index (SFI) y la VGM<sup>8</sup>. El BILAG 2004 presentó el CCI más alto (0,54 [IC el 95%, 0,32-0,78]) comparado con el SFI (0,21 [IC del 95%, 0,08-0,48]) y la VGM [0,18 (IC 95%: 0,06-0,45)].

La validez del BILAG ha sido evaluada en numerosos estudios. En un estudio de 2010 se evaluó la validez de constructo<sup>9</sup> mediante la hipótesis de que pacientes con puntuaciones más altas en el BILAG presentarían niveles más altos de velocidad de sedimentación globular (VSG), anti-ADN de doble cadena y SLEDAI, y por el contrario, presentarían niveles más bajos de complemento. La asociación entre estas variables fue la esperada con un OR = 2,6 (1,2-4,3) para una VSG > 60 mm/h, OR = 2,5 (1,4-3,6) para un anti-ADN de doble cadena superior a 5 veces el valor normal, OR = 4,8 (1,4-15,1) para un C3 por debajo de la mitad del valor normal, OR = 4,1 (2,3-5,8) para un C4 por debajo de la mitad del valor normal y una OR = 215,6 (99,8-387,6) para un valor del SLEDAI por encima de 6.

**Tabla 2**  
Características de los estudios incluidos

Primer autor y año	Tipo de estudio y participantes	INDICE	Aspectos de la validación evaluados	Comentarios
Symmons 1988 <sup>4</sup>	Estudio multicéntrico (5 centros de Reino Unido e Irlanda) sobre síntomas que precisan tratamiento	BILAG	Validez aparente: se comparó la puntuación con decisiones reales de tratamiento	Descripción y validación preliminar
Hay 1993 <sup>5</sup>	Estudio multicéntrico con 82 pacientes	BILAG versión 3 (glosario explicativo)	Fiabilidad; validez de criterio (EO = decisión de tratamiento); validez de constructo (ESR, dsADN)	La mayoría de los pacientes inactivos o poco activos
Stoll 1996 <sup>6</sup>	Estudio transversal de 141 pacientes	BILAG	Consistencia interna y validez de constructo	BILAG vs. VGM
Isenberg 2000 <sup>11</sup>	Descripción narrativa	Programa informático sobre el BILAG	Utilización del British Lupus Integrated Program	Presentación del programa informático
Gordon 2003 <sup>7</sup>	250 pacientes con LES en consulta de rutina	BILAG para definir «brote lúpico»	Validez de criterio para la definición de «brote» (EO = decisión de tratamiento)	Buena validez para evaluar brote severo
Isenberg 2005 <sup>12</sup>	2 tandas de ejercicios con pacientes reales (8 pacientes/8 médicos)	BILAG 2004, actualización de versiones previas	Fiabilidad: CCI entre médicos por órgano/sistema y grado de acuerdo entre médicos	Todos los evaluadores fueron miembros del BILAG
Yee 2007 <sup>13</sup>	Estudio multicéntrico y transversal de 369 pacientes	BILAG 2004	Validez de criterio y de constructo	EO: cambio en el tto
Cresswell 2009 <sup>15</sup>	Estudio multicéntrico y transversal de 369 pacientes	Escala numérica para el BILAG	Comparación entre 3 escalas mediante curvas ROC	Estudio transversal: no tiene en cuenta actividad previa
Yee 2009 <sup>10</sup>	Estudio longitudinal multicéntrico	BILAG 2004	Sensibilidad al cambio	Difícil diferenciar actividad y daño
Nasiri 2010 <sup>9</sup>	Estudio transversal de 100 pacientes	BILAG	Validez de constructo Hipótesis: si ↑ BILAG entonces ↑ VSG, dsADN y ↑ complemento	Algunos pacientes con LES tienen un perfil serológico normal
Vitali 1992 <sup>18</sup>	Estudio multicéntrico con 704 pacientes	ECLAM	Validez aparente y de contenido	Validación inicial
Bencivelli 1992 <sup>19</sup>	Estudio multicéntrico de 75 pacientes	ECLAM	Validez de constructo y de criterio (EO: VGM)	La VGM mostró poca fiabilidad
Vitali 1999 <sup>21</sup>	Comparación: sistema manual/computarizado	ECLAM	Validez de criterio	Correlaciones entre los 2 sistemas
Mosca 2000 <sup>20</sup>	Revisión de 64 historias clínicas	ECLAM	ECLAM directo vs. calculado revisando historias clínicas	Los médicos sabían que se revisarían
Petri 1992 <sup>27</sup>	150 pacientes con LES. Subgrupo de 6 evaluados por 9 médicos (fiabilidad)	UCSF/JHU LAI	Fiabilidad y validez de constructo (VGM)	Estudio transversal
Bombardier 1992 <sup>24</sup>	574 historias evaluadas por 14 reumatólogos	SLEDAI	Desarrollo inicial del índice, fiabilidad y validez aparente	Validación inicial
Guzmán 1992 <sup>27</sup>	39 pacientes en un centro (3 visitas)	MEX-SLEDAI	Comparación con SLEDAI. Fiabilidad y validez de constructo	Sin variables de laboratorio, menor coste
Chang 2002 <sup>29</sup>	Post hoc análisis de un ECA multicéntrico	SLEDAI y SLAM-R	Sensibilidad al cambio	Cambios relevantes para médico y paciente
Gladman 2002 <sup>26</sup>	Cohorte de 960 pacientes	SLEDAI-2K	Validez de constructo	Referencia: SLEDAI
Uribe 2004 <sup>33</sup>	93 pacientes (3 hospitales)	SLAM-R, MEX-SLEDAI y SLEDAI modificado	Validez de constructo y de criterio (EO = SLEDAI-2K)	Ref. V. constructo: VGM (↑ variabilidad)
Touma 2011 <sup>32</sup>	Estudio transversal de 298 pacientes	Respuesta del 50% del SLEDAI	Validez aparente, de contenido y de constructo	Estudio transversal
Bae 2001 <sup>23</sup>	30 pacientes evaluados por 2 médicos/2 visitas	SLAM-R	Fiabilidad y validez de constructo	Médicos inexpertos con SLAM-R
Karlson 2003 <sup>38</sup>	93 pacientes evaluados en una consulta de LES	SLAQ	Validez de constructo	Pacientes asintomáticos con alteraciones de lab.
Yazdany 2008 <sup>39</sup>	Cohorte observacional	SLAQ	Fiabilidad, validez de constructo y sensibilidad al cambio	Cohorte con alto nivel educativo vs. población general
Gladman 1996 <sup>42</sup>	42 casos por 19 médicos	SDI	Validez aparente y de contenido	Validación inicial
Gladman 1997 <sup>43</sup>	10 pacientes por 6–10 médicos de 5 países	SDI	Fiabilidad interobservador	Variabilidad entre observadores mediante ANOVA
Stoll 1997 <sup>14</sup>	141 pacientes en un solo centro	SDI	Consistencia interna y validez de constructo	Compara SDI con BILAG y puntuación de medicación
Gladman 2000 <sup>45</sup>	1.297 pacientes en estudio multicéntrico	SDI	Validez aparente	Evalúa asociación a mortalidad
Costenbader 2010 <sup>47</sup>	Estudio multicéntrico con 569 pacientes y 14 reumatólogos	LDIQ	Validez aparente, de contenido, de criterio y de constructo	EO: respuesta de los médicos en el SDI
Yazdany 2011 <sup>49</sup>	81 pacientes de 2 hospitales universitarios	BILD	Validez aparente, de criterio y de constructo	Alta aceptación por los pacientes
Gladman 1992 <sup>57</sup>	7 pacientes evaluados por 4 a 7 reumatólogos de distintos países	Comparación entre BILAG, SLAM y SLEDAI	Validez de constructo	Médicos de distintos países evalúan la actividad similar a pesar del índice

Tabla 2 (continuación)

Primer autor y año	Tipo de estudio y participantes	INDICE	Aspectos de la validación evaluados	Comentarios
Gladman 1994 <sup>28</sup>	8 pacientes con 3 visitas evaluados por 8 médicos	Comparación entre BILAG, SLAM y SLEDAI	Validez de constructo y sensibilidad al cambio	Sensibilidad al cambio basada en el cambio medio de cada índice
Liang 1989 <sup>22</sup>	25 pacientes evaluados por 2 reumatólogos	Compara: SLAM BILAG y SLEDAI	Fiabilidad y validez de constructo	Los 2 evaluadores del mismo centro
Fortin 2000 <sup>40</sup>	96 pacientes	Comparación entre SLAM y SLEDAI	Sensibilidad al cambio	Utilizan 5 métodos distintos
Ward 2000 <sup>41</sup>	Estudio prospectivo con 20 pacientes	SLAM, BILAG, SLEDAI, LAI, y ECLAM	Validez de constructo y sensibilidad al cambio	Baja correlación entre los índices y VGP
Wollaston 2004 <sup>1</sup>	80 casos evaluados por 20 expertos en LES (miembros de SLICC)	BILAG y SLEDAI vs. la valoración del médico	Fiabilidad y validez de constructo	Comparación con la VGM, que mostró gran variabilidad
Griffiths 2005 <sup>51</sup>	Revisión narrativa	Descripción de la mayoría de índices	–	–
Isenberg 2011 <sup>8</sup>	16 pacientes evaluados por 16 reumatólogos	Comparación: BILAG2004, SELENA y la VGM en la evaluación de brote	Fiabilidad intraobservador	Peor concordancia en el brote moderado
Romero-Díaz 2011 <sup>17</sup>	Revisión narrativa	Descripción que incluye la mayoría de índices	–	Actualización de la versión de 2003

BILAG: British Isles Lupus Assessment group; C: complemento; CCI: coeficiente de correlación intraclase; ECLAM: European Consensus Lupus Activity Measurement; EO: estándar de oro; LDIQ: 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; tto: tratamiento; VGM: valoración global del médico; VGP: valoración global del paciente.

Tabla 3

Estudios excluidos y razones de exclusión

Estudio	Razones de exclusión
Liang, 1991	Descripción narrativa sobre los índices validados hasta el momento
Stoll, 1996	Evaluación del SLICC como predictor de mortalidad
Gladman, 1999	Descripción narrativa del desarrollo del SLICC/ACR Damage Index
Silvestris, 1999	Revisión en italiano
Ward, 2000	Escala para valorar el estado de salud deseado por el paciente
Mosca, 2006	Revisión narrativa sobre las implicaciones de remisión en pacientes con LES
Mattson, 2008	Validación de una escala de fatiga en LES
Ruperto, 2011	Encuesta tipo Delphi entre expertos para llegar a un consenso para la definición de brote lúpico
Lai, 2011	Validación de una escala de fatiga para pacientes con LES
Jolly, 2012	Instrumento para medir la calidad de vida en pacientes con LES
Julian, 2012	Evaluación de una herramienta para valorar el deterioro cognitivo en pacientes con AR y LES

Solo un estudio ha sido específicamente diseñado para evaluar la sensibilidad al cambio del BILAG<sup>10</sup>. En este estudio multicéntrico y longitudinal, se evaluó la relación entre el cambio en el BILAG 2004 y el cambio en el tratamiento entre 2 visitas consecutivas. El aumento del BILAG se asoció a un aumento en el tratamiento (coeficiente de la regresión logística multinomial: 1,35; IC del 95%, 1,01-1,70). En el sentido opuesto, también se encontró esta asociación.

El BILAG se puede calcular manualmente, aunque existe un software informático conocido como British Lupus Integrated Prospective System (BLIPS) que incorpora variables demográficas y la información clínica necesaria para calcular no solo el BILAG, sino otros índices compuestos como el SLAM, SLEDAI, SDI y el SF-36<sup>11</sup>.

Una versión actualizada del BILAG ha sido publicada en 2005 por Isenberg et al. en un intento de mejorar las características del índice<sup>12</sup>. En este estudio, donde se describe el BILAG 2004, se evaluó su viabilidad o capacidad para ser aplicado mediante 2 ejercicios con pacientes reales que fueron examinados por 8 médicos pertenecientes al grupo BILAG. Se calculó el CCI para el BILAG total (0,48; IC del 95%, 0,23-0,81) y para cada uno de los 9 sistemas. El sistema que mostró un CCI más alto fue el renal (0,98; IC del 95%, 0,96-0,99) y el

que mostró un CCI más bajo fue el musculoesquelético (0,17; IC del 95%, 0,01-0,56). Otra medida empleada en este estudio para evaluar la viabilidad fue la ratio entre la desviación estándar atribuible a los médicos y la misma medida atribuible al paciente consigo mismo, mostrando un acuerdo alto en el sistema mucocutáneo, nervioso, renal, oftalmológico y hematológico. En ambos ejercicios, el sistema musculoesquelético mostró una baja fiabilidad en términos de CCI y nivel de acuerdo entre los médicos.

El BILAG 2004 ha sido desarrollado a partir del índice original e incluye 9 sistemas: general, mucocutáneo, sistema nervioso, musculoesquelético, cardiovascular/respiratorio, gastrointestinal, renal, oftalmológico y hematológico. Se compone de varios ítems o preguntas que son evaluadas en una escala de 0 a 4, donde 0 corresponde a «no presente», 1 equivale a «mejorando», 2 es «similar», 3 es «peor» y 4 es «acontecimiento nuevo». Las respuestas se combinan para formar 5 posibles estados de actividad por sistema: A, el paciente está muy activo y, por tanto, precisaría tratamiento con inmunosupresores, corticoides a dosis media o alta (> 20 mg de prednisolona o equivalente) o anticoagulación a dosis alta; B, cierta actividad y, por tanto, plantea la necesidad de tratamiento con dosis moderada de corticoides (< 20 mg de prednisolona), antimaláricos, antidepresivos, o AINE; C, poca actividad con necesidad solo de tratamiento sintomático; D, no actividad en este momento, aunque sí previamente, y E, nunca ha presentado actividad en este sistema.

Otros intentos de validación del BILAG 2004 fueron llevados a cabo por Yee et al. en 2007. Se evaluó la validez de criterio utilizando como estándar de oro el cambio de tratamiento y la validez de constructo intentando corroborar la hipótesis de que pacientes con valores más altos en el BILAG 2004 tendrían valores más altos de VSG, anticuerpos anti-ADN de doble cadena y valores más bajos de complemento<sup>13</sup>. Valores del BILAG 2004 que indican actividad alta o moderada de la enfermedad (A y B) se asociaron de forma significativa con un aumento en la medicación (OR = 19,3,  $p < 0,01$ ). Además, el BILAG 2004 tuvo una sensibilidad del 81%, una especificidad del 91,9%, un valor predictivo positivo del 56,8% y un valor predictivo negativo del 93,6% respecto al cambio de tratamiento. Las puntuaciones más altas del BILAG 2004 se asociaron a una VSG mayor de 60 mm/h (OR = 2,9), una elevación en el anti-ADN de doble cadena (OR = 2,7), una elevación en el SLEDAI por encima de 4 (OR = 20) y una disminución a la mitad de C3 (OR = 5) y C4 (OR = 4,2).

**Tabla 4**  
Aspectos de validación para índice de actividad de la enfermedad y daño

Nombre del instrumento	Viabilidad Tiempo empleado, claridad y aceptación	Fiabilidad			Validez			Sensibilidad al cambio Capacidad para detectar cambios: múltiples métodos	Aplicabilidad
		Consistencia interna: $\alpha$ Cronbach	Intrao test-retest: CCI/Cohen kappa	Intero. error de medición: ESM/MDC/LA	Validez aparente: relevante y comprensible	Validez de constructo: 1. Estructural 2. Test de Ho 3. Cross-cultural	Validez de criterio: Estándar de oro (correlacio- nes/AUC: variables continuas y S/E: dicotómicas)		
<i>Índices para evaluar actividad de la enfermedad</i>									
British Isles Lupus Assessment Group (BILAG) BILAG 2004 <sup>12</sup>	Precisa una HC y EF completa. 5-20 min	$\alpha = 0,35^5$	kappa = 0,79-0,97 <sup>5</sup> CCI = 0,48 (IC del 95%, 0,23-0,81) <sup>12</sup> Kappa evaluación brote lúpico <sup>8</sup> BILAG = 0,54 SELENA = 0,21 VGM = 0,18	NE	Se comparó la puntuación con decisiones reales de tratamiento en un grupo de pacientes <sup>4</sup>	Asociación con $\uparrow$ VSG, dsADN, SLEDAI y $\downarrow$ C3/C4 <sup>9</sup> VGP rho = 0,50 y VGM rho = 0,43 <sup>6</sup> $\uparrow$ VSG: OR = 2,9, $\uparrow$ dsADN: OR = 2,7 SLEDAI > 4: OR = 20 $\downarrow$ C3: OR = 5 $\downarrow$ C4: OR = 4,2 <sup>13</sup>	BILAG «A» score: S = 87% y E = 99% VPP = 80% (> dosis de esteroides e IS) <sup>5</sup> EO/cambio medicación: OR = 19,3, p < 0,01 S = 81%, E = 91,9%, VPP = 56,8% y VPP = 93,6% <sup>13</sup>	RME (VGM) = 0,68 <sup>41</sup> Un aumento en el BILAG se asocia a un aumento del tratamiento <sup>10</sup>	Entrenamiento Recomendaciones específicas para ECA y multicéntricos
European Consensus Lupus Activity Measurement (ECLAM) <sup>18,20,41</sup>	Precisa una HC y EF 5-10 min Sencillo de calcular	NE	NE	Coef. de correlación = 0,90- 0,93 <sup>20</sup>	Variables que mejor se correlación con la VGM	Correlación con SLAM, SLEDAI y BILAG = 0,72 a 0,78 <sup>41</sup>	EO/VGM: r = 0,69 Similar a: BILAG = 0,63 SLAM = 0,61 SLEDAI = 0,66 NE	RME (VGM) = 0,75 <sup>41</sup>	Precisa cierto entrenamiento
Systemic Lupus Activity Measure (SLAM) <sup>22,23,28,40</sup>	Precisa una HC y EF 10-15 min	NE	NE	= 0,86 <sup>22</sup>	NE	SLE índices = 0,81-0,97 VGM = 0,76-0,96 <sup>22</sup> BILAG = 0,79 <sup>30</sup>	NE	RME (VGM) = 0,62 <sup>41</sup> RME (LS) Mejor: -0,88 Peor: 0,61 <sup>40</sup>	Precisa entrenamiento
SLAM-R <sup>23</sup>	NE	NE	= 0,78 mediante ANOVA	= 0,78 mediante ANOVA	NE	VGM r = 0,87 dsADN r = 0,51 C3 r = -0,60 C4 r = -0,29 VGM r = 0,64-0,79 <sup>24</sup> VGM rho = 0,55 <sup>37</sup> BILAG rho = 0,76 SLAM rho = 0,73 <sup>57</sup>	EO: SLEDAI-2K <sup>33</sup> : S = 73% E = 33%	RME (VGM) M: -0,47; p = 0,65 RME (VGP): M: -0,31; p = 0,48 <sup>29</sup>	Incluye manifestaciones subjetivas: fatiga, artralgia y mialgias
Systemic Lupus Erythematosus Disease Activity Measure (SLEDAI) <sup>24,28</sup> SLEDAI-2K <sup>26</sup>	Requiere HC, EF y análisis. 10 a 20 min Ídem	NE	Análisis de variancia, coef. correlación = 0,33 <sup>37</sup>	Intrao = 0,66-0,99 Intero = 0,60-0,80 <sup>24</sup> Intero = 0,47 <sup>37</sup>	NE	r = 0,64-0,79 <sup>24</sup> VGM rho = 0,55 <sup>37</sup> BILAG rho = 0,76 SLAM rho = 0,73 <sup>57</sup>	NE	RME (VGM) = 0,48 <sup>41</sup> RME (VGM) = 0,66 RME (VGP) = 0,05 <sup>29</sup> RME (LS) = 0,57 <sup>40</sup>	
MEX-SLEDAI <sup>27</sup>	Se completa en 16,9 min	NE	NE	NE	NE	SLEDAI rho = 0,97 <sup>26</sup> SLAM-R rho = 0,59 VGM rho = 0,68 <sup>33</sup> VGM rho = 0,68 SLEDAI rho = 0,77 <sup>27</sup> SLAM-R rho = 0,75 VGM rho = 0,54 <sup>33</sup> LAI/VGM = 0,64 <sup>37</sup>	EO: SLEDAI-2K <sup>33</sup> S = 58% y E = 93%	NE	Método de Guyatt (gráfica)
UCSF/JHU Lupus Activity Index (LAI) <sup>37</sup>	1 min	NE	Análisis de variancia, coef. correlación = Intrao = 0,81 <sup>37</sup>	Intero = 0,89 <sup>37</sup>	NE	LAI/VGM = 0,64 <sup>37</sup>	NE	NE	

Tabla 4 (continuación)

Nombre del instrumento	Viabilidad Tiempo empleado, claridad y aceptación	Fiabilidad			Validez			Sensibilidad al cambio Capacidad para detectar cambios: múltiples métodos	Aplicabilidad
		Consistencia interna: $\alpha$ Cronbach	Intrao test-retest: CCI/Cohen kappa	Intero. error de medición: ESM/MDC/LA	Validez aparente: relevante y comprensible	Validez de constructo: 1. Estructural 2. Test de Ho 3. Cross-cultural	Validez de criterio: Estándar de oro (correlacio- nes/AUC: variables continuas y S/E: dicotómicas)		
Systemic Lupus Activity questionnaire (SLAQ) <sup>a</sup> 38,39	Lo completa el propio paciente	$\alpha = 0,87$	NE	NE	NE	SLAM (no lab) rho = 0,62, p < 0,001 VGM rho = 0,73 SF-36 rho = 0,66	NE	RME = 0,12	Desarrollado para estudios epidemiológicos
Índices para evaluar daño acumulado Systemic Lupus International Collaborating Clinics Damage Index (SLICC-DI; SDI) <sup>42,43,45,58</sup>	Se necesita una HC y EF. Aprox. 15 min (según complejidad)	$\alpha = 0,41^{14}$	CCI = 0,553	10 pacientes y 6-10 médicos de 5 países <sup>43</sup>	NE	BILAG rho = 0,19 Tratamiento rho = 0,33 <sup>14</sup>	NE	NE	Seguir recomendaciones de algún miembro del SLICC
Lupus Damage Index Questionnaire (LDIQ) <sup>47</sup>	Lo completa el paciente	NE	NE	NE	Variables incluidas en SDI	LDIQ/SDI rho = 0,48	EO: SDI: S = 53,3% y E = 94,6%	NE	
Brief Index of Lupus Damage (BILD) <sup>49</sup>	Lo completa el paciente	NE	NE	NE	Variables incluidas en SDI	Características demográficas y clínicas por cuartiles	BILD/SDI rho = 0,64 LDIQ/SDI rho = 0,54	NE	Alta aceptación por los pacientes

AUC: área debajo la curva; CCI: coeficiente de correlación intraclase; ECA: ensayo clínico aleatorizado E: especificidad; EF: exploración física; EO: estándar de oro; ESM: error estándar de medida; HC: historia clínica; Intero: interobservadores; Intrao: intraobservadores; LA: límites de acuerdo; LHR+: razón de verosimilitud positiva; LHR-: razón de verosimilitud negativa; LS: escala Likert; MDC: mínimo cambio detectable; NE: no evaluado; r: coeficiente de correlación Pearson; rho: coeficiente de correlación Spearman; RME: respuesta media estandarizada; S: sensibilidad; VGM: valoración global por el médico; VGP: valoración global por el paciente; VPN: valor predictivo negativo; VPP: valor predictivo positivo.

<sup>a</sup> Lo completa el paciente.

Aunque el BILAG es una escala ordinal que permite dar una evaluación global de la enfermedad, algunos autores han propuesto su transformación en una escala numérica para facilitar los análisis estadísticos. Inicialmente, se propuso el siguiente esquema: A=9 puntos, B=3 puntos, C=1 punto y tanto D como E equivaldrían a 0 puntos, con un rango de 0 a 72<sup>14</sup>. En 2009, Cresswell et al. presentaron un estudio que tenía como objetivo validar una puntuación numérica para el BILAG<sup>15</sup>. Describieron 3 métodos distintos de puntuación mediante modelos de regresión logística y, posteriormente, los comparan mediante curvas ROC. La transformación finalmente seleccionada fue A=12, B=5, C=1 y D/E=0.

El BILAG es el único instrumento validado que da una idea de la actividad del lupus en cada órgano de un simple vistazo en lugar de combinar la información en una puntuación global<sup>16</sup>. Se recomienda familiarizarse con el glosario donde se define cada ítem y seguir las pautas y recomendaciones establecidas por el grupo BILAG, especialmente las recomendaciones específicas para su uso en ECA y estudios multicéntricos<sup>5,17</sup>.

#### *European Consensus Lupus Activity Measurement*

El ECLAM fue descrito por el grupo de trabajo europeo para el consenso de medición de actividad en LES en 1992<sup>18</sup>. Es un índice diseñado para medir la actividad de la enfermedad en el último mes en pacientes con lupus.

Este índice fue desarrollado a partir de una cohorte que incluía a un importante número de pacientes reales. En un análisis inicial, se desarrolló el índice y se evaluaron su viabilidad y validez, tanto aparente como de contenido<sup>18</sup>. Para ello se seleccionaron aquellas variables clínicas y de laboratorio que mejor se correlacionaran con la VGM como estándar de oro, tanto en una escala cuantitativa (0-10) como en una escala cualitativa (de inactivo a muy activo). Mediante un análisis univariable, se seleccionaron un total de 15 variables capaces de predecir la actividad de la enfermedad. Posteriormente, mediante un análisis multivariable se definió el peso de cada una de estas variables, a las que se les asignó una puntuación específica. El ECLAM no solo permite clasificar correctamente a los pacientes según su nivel de actividad de la enfermedad, es, además, un índice fácil de calcular. En un segundo estudio, se evaluó la validez de constructo, siendo el estándar de oro la VGM y la validez de criterio, comparando el ECLAM con otros índices compuestos<sup>19</sup>. La correlación del ECLAM con el SLAM, BILAG y SLEDAI fue mayor de 0,72. Estos 4 índices se correlacionaron de forma similar con la VGM.

El ECLAM incluye la evaluación de 10 órganos y/o aparatos y 2 valores de laboratorios que son la VSG y los niveles de complemento. Se compone de un total de 33 ítems que se evalúan de 0,5 a 2, según el tipo de afectación, y suman una puntuación global que va de 0 a 17,5.

El ECLAM es un índice sencillo que se puede usar en estudios retrospectivos, ya que existe una buena correlación entre el ECLAM calculado de forma inmediata y aquel calculado con los datos recogidos en la historia clínica ( $\rho=0,871$ )<sup>20</sup>. Además, se puede calcular mediante un sistema informatizado con resultados muy similares a su cálculo manual ( $r=0,90-0,92$ )<sup>21</sup>.

#### *Systemic Lupus Activity Index*

Este índice mide el grado de actividad de forma global en el último mes. La primera publicación en la que se describe fue en 1986<sup>22</sup>. Dos años después, fue revisado por los residentes de la Universidad de Harvard, que modificaron el apartado de manifestaciones cardiovasculares y «otras» para mejorar la claridad y la reproductibilidad, dando como resultado una nueva versión, el SLAM-R. En 2001, Bae et al. presentan un estudio de 30 pacientes para evaluar la viabilidad y la validez de constructo de esta nueva versión<sup>23</sup>. Se estimó la fiabilidad mediante análisis de variancia, siendo la fiabilidad interobservador de 0,78 y la intraobservador de

0,61. Respecto a la validez de constructo, se calcularon correlaciones de Pearson del SLAM-R con la VGM (0,87), niveles de anti-ADN de doble cadena (0,51), C3 (-0,60) y C4 (-0,29)

El SLAM evalúa manifestaciones específicas en 9 órganos e incluye 7 medidas de laboratorio. Algunos ítems se puntúan de 0 a 3, según el grado de severidad, y otros solo se evalúan de 0 a 1. La máxima puntuación es 84, siendo la parte de laboratorio un máximo de 21 puntos. Una puntuación de 7 o más se considera relevante, ya que el paciente va a precisar un cambio en el tratamiento en el 50% de los casos.

El SLAM es considerado por algunos expertos el índice menos adecuado, ya que incluye medidas subjetivas, como la fatiga y las artralgias; sin embargo, estas variables deben ser puntuadas por el médico siempre y cuando considere que son debidas a la actividad del lupus<sup>8</sup>. Se necesita entrenamiento, ya que es necesario alcanzar un consenso para evaluar los aspectos subjetivos, especialmente en estudios multicéntricos<sup>17</sup>.

#### *Systemic Lupus Erythematosus Disease Activity Measure*

El SLEDAI es un índice global que fue desarrollado por un grupo de expertos de Toronto en 1986 y descrito con detalle por Bombardier et al. en 1992<sup>24</sup>. Fue modificado por el grupo SELENA durante un estudio para evaluar el uso de estrógenos y progesterona en mujeres con LES<sup>25</sup> y, posteriormente, fue actualizado por Gladman et al. en el 2000<sup>26</sup>. Además, existe una versión desarrollada por investigadores mexicanos que, para disminuir los costes del uso de este índice, excluyeron algunos valores de laboratorio<sup>27</sup>.

Por lo tanto, actualmente existen 4 versiones de este índice: SLEDAI, SELENA-SLEDAI, SLEDAI 2000 y MEX-SLEDAI.

Para desarrollar la primera versión del SLEDAI se identificaron 24 variables que podrían ser factores importantes para evaluar la actividad de la enfermedad en pacientes con LES. Con estas variables, se generaron 574 perfiles de posibles pacientes y se presentó esta información a 14 reumatólogos expertos en LES para que evaluaran el grado de enfermedad en una escala de 0 a 10. Se utilizaron modelos de regresión múltiple para estimar la importancia relativa de cada una de estas 24 variables clínicas según la valoración de los expertos y de esta forma se generó este índice global<sup>24</sup>.

El SLEDAI, por tanto, es un índice global que evalúa la actividad de la enfermedad en los últimos 10 días y se compone de 24 ítems que recoge manifestaciones específicas en 9 órganos o sistemas con una puntuación máxima de 105.

En un estudio de 1992 de Guzmán et al. se compararon varios aspectos de la validación del MEX-SLEDAI con el SLEDAI original<sup>27</sup>. La correlación entre 2 evaluadores calculando estos índices osciló entre 0,87 y 0,89. El acuerdo para la valoración de la enfermedad entre los médicos fue moderado, con un kappa de 0,43 ( $p=0,17$ ). La validez de constructo se evaluó usando como comparador la VGM (en una escala de 0 a 10), con una correlación prácticamente idéntica para ambas versiones de 0,68; la correlación entre ambas versiones fue de 0,77.

El SLEDAI se puede utilizar tanto en investigación como en práctica clínica. Y ha demostrado ser una herramienta sensible al cambio<sup>28</sup>. Para evaluar la sensibilidad al cambio del SLEDAI comparándolo con el SLAM-R, Chang et al. realizaron un análisis secundario de los datos obtenidos de un ECA multicéntrico canadiense donde se evaluó la eficacia de metotrexato en pacientes con LES<sup>29</sup>. Se compararon los cambios en estos 2 índices respecto a los cambios en la valoración del médico y del paciente. Respecto a la evaluación del médico, la respuesta estandarizada media de control (C-SRM) para el SLAM-R y el SLEDAI fue de -0,47 vs. -0,42 para mejoría y 0,65 vs. 0,66 para empeoramiento. Respecto a la evaluación del paciente, la C-SRM para el SLAM-R y el SLEDAI fue de -0,31 vs. -0,18 para mejoría y 0,48 vs. 0,05 para empeoramiento. Solo el SLAM-R mantuvo el 0 fuera del IC del 95% para detectar mejoría o deterioro. Ambos índices exhiben una sensibilidad al cambio



relevante para los médicos pero solo el SLAM-R exhibe una sensibilidad al cambio relevante para los pacientes, posiblemente por la inclusión de variables más subjetivas evaluadas por el propio paciente.

En 2002 se introdujo una versión revisada y actualizada del SLEDAI, el SLEDAI 2000 (SLEDAI-2K), en el cual se puntúan el rash, la alopecia, las úlceras o la proteinuria persistente, y no solo de nueva aparición, como ocurría en la versión previa<sup>26</sup>. El SLEDAI-2K fue validado frente al SLEDAI original con una evaluación en los 10 días previos a la visita y además fue validado con un espacio temporal de 30 días para su utilización en ECA<sup>30,31</sup>. El SLEDAI-2K y la versión previa tuvieron una alta correlación ( $r = 0,97$ ) y ambos índices predicen mortalidad de forma similar<sup>26</sup>.

El SLEDAI-2K recoge los aspectos de la enfermedad como presentes o ausentes y no puede reflejar mejoría parcial, lo cual limita su utilización en ECA. Por este motivo, en 2011 Touma et al. desarrollaron un índice de respuesta del 50% (SRI-50) para documentar un 50% de mejoría en SLEDAI<sup>32</sup>. En este estudio, evaluaron la validez de constructo utilizando como comparador externo una valoración de la respuesta al tratamiento por el médico en una escala tipo Likert (LS) que va de 7 (mejoría importante) a 1 (mucho peor). Un 50% de mejoría en esta escala sería un cambio mayor o igual a 6. En pacientes con una mejoría  $\geq 6$  en la LS se produjo una disminución del SRI-50 de  $4,15 \pm 3,01$  ( $p < 0,0001$ ).

En 2004, Uribe et al., en un intento de profundizar más en la validación de las versiones disponibles del SLEDAI, presentaron un estudio en el que se analizó la validez de constructo del SLAM-R, el MEX-SLEDAI y el SLEDAI-2K usando como criterio externo la VGM en una escala visual analógica (EVA) de 0 a 10<sup>33</sup>. Las correlaciones de Spearman de estos índices con la VGM oscilaron entre 0,54 para el MEX-SLEDAI y 0,67 para el SLEDAI-2K. En un intento de evaluar la validez de criterio usando como estándar de oro el SLEDAI-2K, se constató que la sensibilidad del SLAM-R fue del 73%, mientras que la del MEX-SLEDAI fue del 58%. En cuanto a la especificidad, fue del 63% para el SLAM-R y el 93% para el MEX-SLEDAI.

#### UCSF/JHU Lupus Activity Index

El índice LAI se utilizó inicialmente en estudios de infecciones graves y afectación renal en pacientes con LES<sup>34-36</sup>. Se trata de una escala con 5 dominios que reflejan la actividad de la enfermedad en la 2 semanas previas y se completa por el médico en aproximadamente un minuto. Como en el resto de los índices, solo se evalúan aquellas manifestaciones atribuibles al LES. Incluye una VGM en una EVA de 0 a 3, e incluye 4 EVA para fatiga, rash, articulaciones y serositis, y una parte que cuantifica la afectación en 4 órganos: neurológico, renal, pulmonar y hematológico, en EVA de 0 a 3. Además, esta última parte asigna distintas puntuaciones según la necesidad de medicación y valores de laboratorio.

En un estudio de Petri et al.<sup>37</sup>, se evaluaron la fiabilidad y la validez de constructo de este índice comparado con el SLEDAI utilizando como referencia la VGM. La VGM tuvo una correlación más alta con el LAI ( $r = 0,64$ ) que con el SLEDAI ( $r = 0,55$ ). En cuanto a la fiabilidad, el LAI mostró mayor fiabilidad, tanto intra como inter-observador.

#### Systemic Lupus Activity questionnaire

El SLAQ se desarrolló como herramienta de medida por el propio paciente para ser utilizada en estudios epidemiológicos y grandes cohortes de pacientes. Fue descrito y validado por Karlson et al. en 2003<sup>38</sup>. Fue desarrollado en una cohorte clínica de 93 pacientes y se comparó con el SLAM (sin incluir datos de laboratorio), mostrando una buena correlación ( $r = 0,62$ ,  $p < 0,001$ ). La correlación individual de cada ítem entre paciente-médico osciló entre 0,06 para la evaluación de linfadenopatías y vasculitis y 0,7 para la evaluación de síndrome de Raynaud.

Posteriormente, Yazdany et al. evaluaron otros aspectos de la validación en una cohorte de 982 pacientes<sup>39</sup>. El SLAQ mostró una buena consistencia interna, con un  $\alpha$  de Cronbach de 0,87. Para evaluar la validez de constructo, los autores examinaron la correlación del SLAQ con la VGP ( $r = 0,73$ ) y con el SF-36 ( $r = 0,66$ ). El índice mostró una baja sensibilidad al cambio con una SRM de 0,12, aunque cuando fue comparado con el cambio en otras medidas la puntuación del SLAQ cambió en la dirección adecuada.

#### Estudios que comparan los índices

El primer estudio en el que se comparan varios índices compuestos se presenta en 1989 por Liang et al.<sup>22</sup>. En este estudio se analiza la fiabilidad del SLAM, SLEDAI y BILAG, y su validez de constructor; por un lado, se calculó la correlación entre cada índice y los demás (que osciló entre 0,81 y 0,97) y, por otro, calculando la correlación de cada uno de ellos con la VGM en una EVA de 0 a 10 ( $r = 0,76-0,96$ ). En un intento de evaluar no solo la validez de constructo, sino también la sensibilidad al cambio, Gladman et al. diseñan un estudio en el que 8 pacientes, con 3 visitas por paciente, son evaluados por 8 reumatólogos<sup>28</sup>. La correlación entre estos 3 índices osciló entre 0,35 y 0,61. Respecto a la sensibilidad al cambio, solo el SLEDAI fue capaz de diferenciar entre visitas mediante el análisis de variancia. Estos resultados están basados en el cambio medio de cada índice y, por tanto, es difícil evaluar la relevancia clínica de este cambio observado en cada paciente. Unos años más tarde, Fortin et al. realizan un estudio para valorar la sensibilidad al cambio mediante distintos métodos y utilizando como criterio externo la valoración del cambio realizada por el médico en una LS de 5 puntos, que va de «mucho mejor» a «mucho peor»<sup>40</sup>. Los 3 índices mostraron ser sensibles al cambio, el SLAM con mejores resultados. Los autores interpretan que esta pequeña diferencia puede ser debida a que el SLAM permite dar valores crecientes según la gravedad mientras que el SLEDAI tiene una puntuación fija para cada ítem.

En otro estudio del mismo año, Ward et al. evaluaron la validez de constructo y la sensibilidad al cambio del SLAM, SLEDAI, LAI, BILAG y el ECLAM<sup>41</sup>. Todos estos índices resultaron válidos para medir la actividad de la enfermedad en pacientes con LES. El SLAM es el que capturaba mejor la percepción del paciente ( $r = 0,22$ ;  $p < 0,0001$ ). La correlación entre el cambio en cada uno de los índices y el cambio en la valoración del médico fueron en orden decreciente: LAI  $r = 0,75$ ; ECLAM  $r = 0,65$ ; BILAG  $r = 0,61$ ; SLAM  $r = 0,54$  y el SLEDAI  $r = 0,52$ , todos con una  $p < 0,0001$ . El índice que mostró mayor sensibilidad al cambio medido por la SRM y usando la VGM como referencia fue el LAI (SRM = 0,74) y el más pequeño, el SLEDAI (SRM = 0,48).

En un intento por desarrollar un índice que evaluara la respuesta al tratamiento, en 2004 Wollaston et al. compararon el cambio en 2 índices validados, el BILAG y el SLEDAI, con el cambio en la actividad de la enfermedad evaluado por un médico experto en lupus usando una LS de 7 puntos (1 = mejoría importante y 7 = empeoramiento respecto a actividad basal)<sup>1</sup>. El CCI de la valoración del médico en los 4 grupos de evaluadores osciló entre 0,25 y 0,46. El cambio de basal a 3 meses del SLEDAI tuvo una buena correlación con el cambio del BILAG ( $r = 0,75$ , IC del 95%, 0,63–0,83). La conclusión de este estudio fue que, mientras los índices compuestos son comparables entre ellos, existe una gran variabilidad en la valoración realizada por los médicos, incluso entre médicos con gran experiencia con pacientes con lupus.

#### Índices para evaluar daño en el lupus eritematoso sistémico

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

Desarrollado en 1996 por un grupo de colaboración internacional (grupo SLICC) y adoptado por el CAR<sup>42</sup>. En este primer estudio, se

desarrolló el índice y se evaluó su validez aparente y de contenido. Se generó una lista de variables que podrían reflejar daño en pacientes con LES y se llegó a un consenso sobre qué variables debían ser incluidas en un índice para evaluar daño irreversible. En otro estudio del mismo año y por los mismos autores, se evaluó la fiabilidad interobservador mediante 10 pacientes a los que reumatólogos de 5 países distintos aplicaron este índice<sup>43</sup>. Los autores concluyeron en este estudio que médicos de distintos países evalúan el daño en pacientes con LES de forma muy similar.

El SDI evalúa daño irreversible en pacientes con lupus sin tener en cuenta la causa que lo produce. Incluye 42 ítems que miden la afectación de 12 dominios con una puntuación máxima de 46 puntos. Cada ítem es evaluado como presente o ausente con la posibilidad de puntuar un 2 o 3 en caso de eventos recurrentes, como por ejemplo puede ser el caso de un accidente cerebrovascular.

La definición de daño en pacientes con LES es un cambio irreversible en un órgano o sistema que ha ocurrido desde el inicio del LES y está presente al menos durante los últimos 6 meses. En el momento del diagnóstico, la puntuación debería ser 0 por definición.

El SDI lo completa el médico en la visita basal al entrar en un ECA. Tiene una consistencia interna moderada ( $\alpha$  de Cronbach = 0,41). En un intento por evaluar la validez de constructo, Stoll et al. compararon el SDI con el BILAG y con una puntuación de la medicación del paciente, en ambos casos con muy baja correlación (0,19 para el BILAG y 0,33 para el índice de medicación)<sup>14</sup>.

Cuando el SDI es completado por otro médico al revisar de forma retrospectiva la historia clínica del paciente, también muestra buena fiabilidad interobservador<sup>44</sup>. Los valores de SDI aumentan con la evolución de la enfermedad de forma similar en pacientes de distintos países y, además, es una herramienta que predice mortalidad en pacientes con LES<sup>45,46</sup>.

#### *Lupus Damage Index Questionnaire*

El LDIQ fue descrito por Costenbader et al. en 2010, en un intento de desarrollar una medida de daño irreversible basada en el SDI pero completada por el propio paciente para ser utilizada en práctica clínica o en estudios epidemiológicos<sup>47</sup>. En este estudio hay una primera etapa de desarrollo del índice donde se evalúa la validez aparente y de contenido con un grupo de 37 pacientes y 7 reumatólogos, y posteriormente, una validación más elaborada a nivel multicéntrico. Para evaluar la validez de criterio se utiliza como estándar de oro el SDI, dando como resultado una sensibilidad del 53,3%, una especificidad del 94,6% y un grado de acuerdo entre las 2 medidas del 93,2%. Para evaluar la validez de constructo, se calculó la correlación entre LDIQ y SDI (coeficiente de correlación de Spearman = 0,48,  $p < 0,001$ ).

El LDIQ consta de 56 preguntas para evaluar cada dominio incluido en el SDI y está diseñado para ser administrado como una encuesta. Sus versiones en español, portugués y francés han sido validadas<sup>48</sup>.

#### *Brief Index of Lupus Damage*

El BILD fue descrito en 2011 por Yazdany et al. como una medida de daño en pacientes con LES para ser usada en estudios poblacionales<sup>49</sup>. Al igual que el LDIQ, lo completa el propio paciente pero es una versión más corta que incluye solo 28 preguntas. La validez de criterio del BILD se examinó mediante su correlación con otro instrumento que también mide daño, como es el SDI, y que se puede considerar como el estándar de oro ( $r = 0,64$ ,  $p < 0,001$ ). La validez de constructo se evaluó comparando las características demográficas y clínicas de los pacientes divididos en los 4 cuartiles del BILD. Los pacientes con valores más altos para el BILD fueron mayores, con mayor duración de la enfermedad y con mayor actividad de la enfermedad.

## Discusión

Para medir la actividad de la enfermedad y el daño irreversible en pacientes con LES necesitamos una medida compuesta cuantitativa que tenga una aceptable validación. Diseñar estas medidas compuestas supone un reto, dada la afectación multisistémica en el LES y su extensa variabilidad. En el momento actual, existen múltiples herramientas con mayor o menor grado de validación, aunque ninguna ha sido aceptada como la única medida recomendada a nivel internacional. La elección de la medida más apropiada en cada caso dependerá fundamentalmente del contexto en el que se vaya a utilizar y la pregunta que se quiera responder en términos de evaluación de la enfermedad.

A grandes rasgos, existen 2 tipos de índices de actividad, los que funcionan como medidas globales (como el ECLAM, SLAM, SLEDAI y LAI) y los que dan una puntuación específica para cada órgano o sistema (BILAG).

Los índices globales se consideran útiles para comparar cohortes de pacientes con LES, ya que son más sencillos de aplicar y, por tanto, más útiles en estudios poblacionales. También se pueden utilizar para definir criterios de inclusión en ECA o criterios de respuesta a un tratamiento. Pero estos índices no dan información sobre el grado de actividad en un órgano específico, para lo que serían más útiles los índices por órgano/sistema.

En cuanto a su viabilidad, la mayoría de estos índices son complejos y requieren información de la historia clínica y la exploración física. El índice que requiere más tiempo es el BILAG y los más sencillos de aplicar, el LAI y el SLAQ. La consistencia interna fue solo evaluada en el BILAG y el SLAQ, siendo mayor en este último. Respecto a la fiabilidad interobservador, el índice que mostró un mayor CCI fue el SLEDAI (ICC = 0,79). La fiabilidad interobservador fue evaluada de una forma muy heterogénea entre los distintos estudios, haciendo muy difícil la comparación entre los índices. Este aspecto de la validación no es necesario en los índices realizados por el paciente.

La validez de constructo fue la parte de la validación más estudiada para la mayoría de los índices con similares resultados. Es difícil establecer un estándar de oro que sirva de referencia para evaluar la actividad de la enfermedad. Para evaluar la validez de criterio, en algunos estudios se describe como estándar de oro el cambio de tratamiento y en otros la VGM. Utilizar la VGM como estándar de oro tiene ciertos problemas, ya que diversos estudios han demostrado que existe un grado de acuerdo entre médicos muy bajo a la hora de valorar la actividad de la enfermedad en pacientes con LES<sup>1,27</sup>. Otros estudios emplean como estándar de oro otro índice ya descrito en la literatura, con el que se compara el nuevo índice evaluado. La mayoría de los índices se correlacionan de forma aceptable entre sí<sup>22</sup>.

Por último, el índice que muestra una mayor sensibilidad al cambio es el ECLAM.

Algunos aspectos importantes que se deben tener en cuenta es que en los índices de actividad solo se deben puntuar aspectos que están directamente relacionados con la enfermedad. El SLAM y el BILAG recogen manifestaciones que aparezcan en el último mes, mientras los demás recogen manifestaciones ocurridas en los últimos 10 días. Algunos índices incluyen variables de laboratorio y es importante tener en cuenta que hay pacientes con LES activo que no presentan alteraciones analíticas. Por ejemplo, solo el 60% de los pacientes tendrán anticuerpos anti-dsADN positivos durante la evolución de su enfermedad<sup>50</sup>.

Una crítica sobre estos índices podría ser que la mayoría de ellos han sido validados en el contexto de cohortes de larga duración en lugar de en ECA<sup>51</sup>. Aun así, en 1998, uno de los objetivos de trabajo de OMERACT fue definir de forma preliminar las variables mínimas que debían incluirse en la evaluación de los pacientes con LES, tanto en ECA como en cohortes longitudinales de larga duración,

y estas fueron actividad de la enfermedad, daño, calidad de vida y toxicidad/efectos adversos<sup>52</sup>.

El concepto de daño en el LES se ha convertido en una medida de desenlace importante, ya que no solo predice mortalidad, también capacidad funcional y utilización de recursos sanitarios<sup>53-55</sup>. En esta revisión, 3 medidas de daño han sido identificadas, una que completa el médico, el SDI, y 2 que completa el paciente, LDIQ y BILD. Las 2 medidas de daño autoadministradas mostraron una buena correlación con la medida del médico considerada como estándar de oro.

En conclusión, la mayoría de los índices compuestos para valorar la actividad de la enfermedad en pacientes con LES han sido validados y son comparables entre sí. Aunque son complejos a la hora de ser aplicados en la práctica clínica, estos índices facilitan la recogida de información clínica relevante de forma cuantitativa y EULAR recomienda su uso para monitorizar a los pacientes<sup>56</sup>. El uso de índices compuestos para evaluar tanto la actividad, como el daño estructural, ayudaría en práctica clínica a guiar las decisiones terapéuticas de la forma más objetiva posible<sup>51</sup>. Por otro lado, la VGM exhibe una gran variabilidad, incluso entre médicos con una larga experiencia tratando a pacientes con LES, con lo cual no es una herramienta ideal para valorar la actividad.

Tanto el BILAG como el SLEDAI son los índices que tienen una validación más completa y los que más se usan en ECA y estudios de cohortes. El ECLAM tiene la ventaja de ser el más sencillo de calcular y el SLAM es el más sensible al cambio, ya que es el único que permite dar una puntuación mayor a mayor gravedad. El SLAQ es un índice que completa el paciente y exhibe una buena correlación con el SLAM, del que deriva, por lo que puede suponer una buena alternativa para consultas en las que se puede dedicar menos tiempo por paciente o para estudios poblacionales.

## Responsabilidades éticas

**Protección de personas y animales.** Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

**Confidencialidad de los datos.** Los autores declaran que en este artículo no aparecen datos de pacientes.

**Derecho a la privacidad y consentimiento informado.** Los autores han obtenido el consentimiento informado de los pacientes y/o sujetos referidos en el artículo. Este documento obra en poder del autor de correspondencia.

## Financiación

Este estudio ha sido financiado por la Fundación Española de Reumatología.

## Conflicto de intereses

Iñigo Rúa-Figueroa ha tenido contratos con GSK y ha asistido a cursos/congresos patrocinado por GSK y MSD.

Los demás autores declaran no tener ningún conflicto de interés.

## Anexo. Material adicional

Se puede consultar material adicional como las estrategias de búsqueda de este artículo en su versión electrónica disponible en <http://dx.doi.org/10.1016/j.reuma.2014.01.012>.

## Bibliografía

- 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 SLENA. *Ann Rheum Dis*. 2011;70:54–9.
- Nasiri S, Karimifar 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*. 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.
- Bencivelli W, Vitali C, 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.

23. 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.
24. 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.
25. Petri M, Buyon J, Skovron ML, Kim M. Disease activity and health status (SF-36) in postmenopausal systemic lupus erythematosus: The SELINA trial [abstract]. *Arthritis Rheum*. 1987;40.
26. Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol*. 2002;29:288–91.
27. 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.
28. 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.
29. 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.
30. 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.
31. Touma Z, Urowitz MB, Gladman DD. SLEDAI-2K for a 30-day window. *Lupus*. 2010;19:49–51.
32. 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.