





Original article

Easy-BILAG: a new tool for simplified recording of SLE disease activity using BILAG-2004 index

Lucy M. Carter ¹, Caroline Gordon², Chee-Seng Yee³, Ian Bruce ⁴, David Isenberg ⁵, Sarah Skeoch⁶ and Edward M. Vital ¹

Abstract

Objective. BILAG-2004 index is a comprehensive disease activity instrument for SLE but administrative burden and potential frequency of errors limits its use in routine practice. We aimed to develop a tool for more accurate, time-efficient scoring of BILAG-2004 index with full fidelity to the existing instrument.

Methods. Frequency of BILAG-2004 items was collated from a BILAG-biologics registry (BILAG-BR) dataset. Easy-BILAG prototypes were developed to address known issues affecting speed and accuracy. After expert verification, accuracy and usability of the finalized Easy-BILAG was validated against standard format BILAG-2004 in a workbook exercise of 10 case vignettes. Thirty-three professionals ranging in expertise from 14 UK centres completed the validation exercise.

Results. Easy-BILAG incorporates all items present in $\geq 5\%$ BILAG-BR records, plus full constitutional and renal domains into a rapid single page assessment. An embedded glossary and colour-coding assists domain scoring. A second page captures rarer manifestations when needed. In the validation exercise, Easy-BILAG yielded higher median scoring accuracy (96.7%) than standard BILAG-2004 documentation (87.8%, $P = 0.001$), with better inter-rater agreement. Easy-BILAG was completed faster (59.5 min) than the standard format (80.0 min, $P = 0.04$) for 10 cases. An advantage in accuracy was observed with Easy-BILAG use among general hospital rheumatologists (91.3 vs 75.0, $P = 0.02$), leading to equivalent accuracy as tertiary centre rheumatologists. Clinicians rated Easy-BILAG as intuitive, convenient, and well adapted for routine practice.

Conclusion. Easy-BILAG facilitates more rapid and accurate scoring of BILAG-2004 across all clinical settings, which could improve patient care and biologics prescribing. Easy-BILAG should be adopted wherever BILAG-2004 assessment is required.

Key words: systemic lupus erythematosus, BILAG, disease activity, assessment, index

Rheumatology key messages

- Easy-BILAG is a high-accuracy, time-efficient tool for recording BILAG-2004 disease activity in SLE.
- It is the new recommended format for scoring BILAG-2004 index in clinical practice.
- Easy-BILAG and its training material is available free of charge for use in routine care at <https://licensing.leeds.ac.uk/products/healthcare-questionnaires>.

Introduction

Disease activity measurements in SLE are necessary for optimal patient care. They are central to clinical guidelines [1, 2] and treat-to-target approaches, which have been

shown to improve outcomes in SLE [3, 4], and rely on specifically defining and measuring low disease activity and remission [5]. Furthermore, national commissioning policies for biologic agents also increasingly stipulate measured baseline and response disease activity criteria [2, 6].

¹Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, ²Rheumatology Research Group, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, ³Department of Rheumatology, Doncaster and Bassetlaw Teaching Hospitals NHS Foundation Trust, Doncaster, ⁴Centre for Epidemiology Versus Arthritis, Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, University of Manchester, Manchester, ⁵Department of Rheumatology, Division of Medicine, University College London, London and ⁶Department of Rheumatology, Royal

National Hospital for Rheumatic Diseases, Royal United Hospitals Bath NHS Trust, Bath, UK

Submitted 30 July 2021; accepted 16 November 2021

Correspondence to: Edward Vital, Leeds Institute for Rheumatic and Musculoskeletal Medicine, Chapel Allerton Hospital, Leeds LS7 4SA, UK. E-mail: e.m.vital@leeds.ac.uk

Recording complex multisystem manifestations longitudinally is a significant challenge in SLE. As a result, formalized disease measures have been less readily embedded in routine care [7] compared with other rheumatic diseases such as rheumatoid arthritis. Composite disease activity instruments, including the SLEDAI, ECLAM and the British Isles Lupus Assessment Group (BILAG)-2004 index, all have proven validity and reliability [2, 8, 9]. The BILAG-2004 index is the most comprehensive available instrument. It replaced the original 'classic' BILAG index [10] and the current version includes numerical scoring and updates to haematology items [11, 12]. BILAG-2004 index comprises 97 discrete clinical manifestations of SLE across nine organ domains. The activity in each domain is graded separately from A, highly active and likely to necessitate escalation in therapy, to E, no current or previous disease activity [13]. BILAG-2004 captures several important disease features such as haemolysis, lymphadenopathy, polyneuropathy, interstitial pneumonitis, gastrointestinal and significant although rare ophthalmological items that do not feature in SLEDAI. Additionally, all organ domains carry potentially equal weighting. Unlike SLEDAI, BILAG-2004 index differentiates between disease features that have partially, but not completely improved, those that have not changed, and those that are worse. It is thus more sensitive than SLEDAI to changes in disease activity over time and better detects partial response to therapy and the exacerbation of already active disease features [14, 15]. These differences may be particularly important in research studies.

The current BILAG-2004 documentation relies on an index case report form, a detailed glossary of clinical items and a separate scoring algorithm for each of the nine organ domains such that the overall A–E domain scores, notably renal and haematological, are frequently not available at the point of completing the case report form. The requirement to rate items as improving, worse, or the same requires assessment of the status 1 month earlier. Formal training is also recommended. Therefore, despite its advantages, BILAG-2004 index may be difficult or time-consuming to complete during routine clinic visits, particularly for those not familiar with the glossary, the layout of the case report form and scoring document, even though most patients have relatively few abnormal items present at any single visit.

The Easy-BILAG project aimed to develop and validate a simplified tool to record and score the current published version of the BILAG-2004 index [11] more rapidly and accurately for use in routine clinical care.

Methods

Development process

Easy-BILAG was registered as a multicentre quality improvement initiative with the Quality Assurance and Governance department of Leeds Teaching Hospitals NHS Trust and at the relevant governance departments

at individual participating trusts. No real patient data was used in Easy-BILAG validation material and specific research ethics approval was therefore not obtained. BILAG-BR is an ongoing prospective study with research ethics approval from NRES Committee North West–Greater Manchester West (REC: 09/H1014/64) and Health Research Authority approved on 9 November 2009 (IRAS ref. 24407).

Reasons underlying inaccuracy or difficulty completing BILAG-2004 assessment were discussed in meetings of the BILAG group of expert clinicians, experienced in delivering BILAG-2004 training and adjudicating clinical trials. The frequency of BILAG-2004 clinical items was evaluated in an active SLE cohort. Based on these insights, a series of Easy-BILAG prototypes was developed to address the key problems identified. The finalized Easy-BILAG was validated in a workbook exercise. Minor changes to the wording in the constitutional and renal domains were made after validation as directed by feedback from participants. A separate self-adjudication checklist was also added, to assist with use in a clinical trial setting, after the validation.

Determining the frequency of BILAG-2004 SLE manifestations

Pseudonymized BILAG-2004 disease activity scores, from individual SLE patients enrolled in the UK BILAG-Biologics Registry (BILAG-BR) between March 2010 and November 2019, were available for evaluation. The majority of enrolled patients had moderate to severe SLE disease activity and were commencing biologic therapy [16]. The frequency with which each of the 97 BILAG-2004 clinical items was recorded as 'new', 'improving', 'same', or 'worse' was quantified. Quantitative items (i.e. full blood count, creatinine, eGFR, proteinuria and blood pressure) were recorded as raw numerical values in this dataset, independent of attribution to SLE disease activity and these were therefore excluded from this analysis.

Validation exercise

Consultant rheumatologists including members of BILAG as well as those without sub-speciality interests in SLE, rheumatology speciality trainees and experienced lupus specialist nurses from 14 UK centres were invited by BILAG members to complete a timed validation workbook of 10 short case vignettes. Workbooks were distributed to participating centres by mail. Each participant was provided with a workbook that randomly assigned them to Easy-BILAG or standard format. Each clinician thereby scored BILAG-2004 disease activity for the same 10 cases using either Easy-BILAG or standard format BILAG-2004 index form, glossary and scoring algorithm [11]. No prior BILAG-2004 training or experience was prerequisite. The validation workbook was designed to test clinicians in scoring both frequent and uncommon SLE manifestations, and test longitudinal scoring of items in flare and remission. All workbooks contained an introductory overview and detailed instructions on how

to use the BILAG-2004 index. Individuals self-reported the time taken to score the 10 cases. They were also asked to report their level of prior BILAG-2004 experience, job role and whether they worked within a general or tertiary hospital setting. Overall perception and usability of the scoring format was assessed by four Likert-item survey questions. Clinicians returned their workbooks in hard copy to the central study team and could opt to submit anonymously. Item and domain level accuracy was calculated against a model answer scheme verified in advance by original BILAG-2004 authors.

Statistical analysis

After demonstrating non-normality, variables were compared between two groups in SPSS using Wilcoxon rank-sum test and between multiple groups by Kruskal–Wallis followed by relevant pairwise comparisons. Interrater agreement for all domain scores was evaluated according to BILAG-2004 format by Fleiss' kappa coefficient using R Studio *irr* package [17] and the result categorized as previously described [18]. Statistical significance was considered $P < 0.05$ on one-tailed testing.

Results

Key problems identified in BILAG-2004 scoring

Consultation among BILAG-2004 lead authors and expert clinicians identified features of the current published recording format that pose barriers to its accurate and/or time-efficient use in routine practice, many of which are also relevant to clinical trials. Strategies to resolve priority issues were developed. These are collectively summarized in [Supplementary Table S1](#), available at *Rheumatology* online.

Item level frequency of BILAG-2004 SLE manifestations in a UK biologics registry

Item level data from 2395 BILAG-BR disease activity records revealed that the six most frequent items were each present in >20% of records, namely: mild arthritis (72%), mild skin eruption (47%), moderate arthritis (38%), mild mucosal ulceration (34%), mild alopecia (34%) and pleurisy/pericarditis (22%). Twenty-two discrete items were present in 5% or more of cases and no Ophthalmic or Gastrointestinal domain items were among these. Twenty-five items were active in <1% of BILAG-BR records ([Supplementary Fig. S1](#), available at *Rheumatology* online).

The Easy-BILAG template

Easy-BILAG was designed to enhance the visibility of the most frequently scored BILAG-2004 items identified by a BILAG-BR dataset ([Fig. 1](#)). A single page Easy-BILAG now captures full disease activity for 68% of biologic-treated patients in BILAG-BR. The constitutional and renal domains are scored in full on page 1 and all

other clinical items recorded as active in >5% of BILAG-BR records are also captured on the first page ([Fig. 1](#)). Less frequent manifestations are scored, only when necessary, on a second page ([Fig. 2](#)) as prompted by screening questions on page 1 ([Fig. 2A](#)). The design incorporates an abridged glossary definition immediately adjacent to clinical each item to facilitate closer adherence to glossary criteria. A colour-blindness compatible, colour-coding system directs clinicians instantly to the overall A–E score for each domain, so complete scores are readily derived at the time of assessment ([Fig. 1](#)). Constitutional and renal domains follow a similar principle, but, because a combination of features is needed to compute the domain score, we have adopted an arbitrary points system to facilitate scoring with full fidelity to original BILAG-2004 algorithms ([Fig. 1](#)). On completion of Easy-BILAG, an entirely optional self-adjudication checklist, ensuring compliance with core BILAG-2004 scoring rules such as attribution to SLE, improvement criteria and trickle-down rule is available and might assist particularly in a clinical trial setting ([Supplementary Fig. S2](#), available at *Rheumatology* online).

Easy-BILAG facilitates superior BILAG-2004 scoring accuracy and speed

Accuracy of Easy-BILAG scoring was tested against standard format BILAG-2004 index in a validation workbook of 10 case vignettes completed by rheumatology professionals ($n = 33$) in 14 centres around Great Britain and Northern Ireland. Forty-five per cent of participants in the validation exercise were consultant rheumatologists or clinical academics and 45% were speciality trainees. Overall, 42.4% reported their use of BILAG-2004 index in their current practice as infrequent or rare. Further characteristics of professionals participating are shown in [Supplementary Table S2](#), available at *Rheumatology* online.

Accuracy of scoring [% accuracy; median (Quartile 1, Quartile 3)] against expert verified model answers was significantly higher across all domains with use of Easy-BILAG [$n = 16$; 96.7 (94.4, 97.5), Mann–Whitney $U = 53.0$, $P = 0.001$; [Fig. 3A](#)] as compared with the standard BILAG-2004 format [$n = 17$; 87.8 (80.0, 94.4); [Fig. 3A](#)]. Since assessment of active rather than quiescent disease is more prone to error, rating of domains requiring grade A to C scores was examined separately. Rating of active domains retained high levels of accuracy with use of Easy-BILAG [94.6 (89.1, 97.2); [Fig. 3B](#)], which was significantly above that observed using standard format BILAG-2004 [80.4 (65.2, 90.2), $U = 48.0$, $P = 0.001$; [Fig. 3B](#)]. Self-reported completion time for the 10-case workbook [min; median (Q_1 , Q_3)] was significantly shorter among clinicians using Easy-BILAG [$n = 16$; 59.5 min (53.2, 86.3); [Fig. 3C](#)] than standard format [$n = 17$; 80.0 min (61.0, 104.0), $U = 87.0$, $P = 0.04$; [Fig. 3C](#)].

Fig. 1 Annotated example of completed Easy-BILAG main page

		ITEM	ABBREVIATED GLOSSARY	Not Present	Improving	Same	Worse	New	SCORE A-E
Mucocutaneous		Skin eruption - severe	Any lupus rash except panniculitis, bullous lesion and angio-oedema. Must involve >18% BSA	0	1	2	3	4	B
		Skin eruption - mild	As above but <18% BSA. If malar rash: observed by doctor and present at least 1/52	0	1	2	3	4	
		Mucosal ulceration - severe	Disabling (sig interfering with oral intake), extensive & deep. Observed by a physician.	0	1	2	3	4	
		Mucosal ulceration - mild	Localized and/or non-disabling ulceration	0	1	2	3	4	
		Alopecia - severe	Clinically detectable (diffuse or patchy) hair loss with scalp inflammation (redness over scalp)	0	1	2	3	4	
		Alopecia - mild	Diffuse or patchy hair loss without scalp inflammation (clinically detectable /by history)	0	1	2	3	4	
		Digital infarct / nodular vasculitis	Localised single / multiple infarct(s) over digit(s) or tender erythematous nodule(s)	0	1	2	3	4	
		Periungual erythema / chilblains	Chilblains = localised inflammatory lesions (may ulcerate) precipitated by exposure to cold	0	1	2	3	4	
	Other features	Angio-oedema, panniculitis, cutaneous vasculitis, or splinter haemorrhages	No	YES → NEXT PAGE					
MSK		Arthritis - severe	Observed Synovitis ≥2 joints + marked loss of ROM & ADL, on several days (cum) in past 4/52.	0	1	2	3	4	C
		Arthritis (moderate), tendonitis or tenosynovitis	≥1 joint (observed or history), some loss ROM, several days of past 4/52	0	1	2	3	4	
		Arthritis (mild), arthralgia or myalgia	Inflammatory pain over joints / muscle	0	1	2	3	4	
		Other features	Myositis	No	YES → NEXT PAGE				
CardioResp		Pleurisy / Pericarditis	Convincing history and/or physical findings. Do not score if unsure.	0	1	2	3	4	D
		Pleural effusion with dyspnoea	Supportive imaging required	0	1	2	3	4	
		Interstitial alveolitis/ pneumonitis	Supportive imaging required. Corrected Kco <70% normal or fall by > 20%.	0	1	2	3	4	
		Other features	Any other cardiac or respiratory problem due to active SLE.	No	YES → NEXT PAGE				
NeuroPsych		Mononeuropathy (single or multiplex)	Supportive electrophysiology required	0	1	2	3	4	E
		Polyneuropathy	Acute symmetrical distal sensory and/or motor deficit (supportive electrophysiology required)	0	1	2	3	4	
		Seizure disorder	Independent description of seizure by reliable witness	0	1	2	3	4	
		Other features	Any other CNS or PNS feature due to active SLE	No	YES → NEXT PAGE				
Haematology		Full blood count		A	B	C	D	E	C
		Hb WITHOUT evidence of haemolysis		<8.0	8-10.9	>10.9	Never	Never	
		Hb WITH evidence of haemolysis		<8.0	8.0-9.9	>10	Never	Never	
		Total WCC (x10 ⁹ /L)		<4.0	1-3.9	>3.9	Never	Never	
		Neutrophils (x10 ⁹ /L)		<0.5	0.5-1.9	>1.9	Never	Never	
		Lymphocytes (x10 ⁹ /L)		<1.0	>1.0	Never	Never		
	Platelets (x10 ⁹ /L)		<25	25-49	50-149	>149	Never		
	Other features	TTP or isolated Coombs	No	YES → NEXT PAGE					
Gastrointestinal		E		Ophthalmic		E			
If active check next page				If active check next page					
For the following systems, circle all that apply and total the points to determine the domain score									
Constitutional				No	Imp	Same	Worse	New	= 20 C
		Pyrexia	Documented Temp >37.5 (infection excluded)	0	10	1000	1000	1000	
		Weight loss	Unintentional >5% loss in 1 month due to SLE	0	10	100	100	100	
		Lymphadenopathy or splenomegaly	Lymph node > 1cm diameter	0	10	100	100	100	
	Anorexia	Due to active SLE	0	10	100	100	100		
Renal		Select all that apply:	Biopsy nephritis past 3 months? (1000)	Active urinary sediment?	1000	Nephrotic syndrome?	100	= 2000 A	
		Urinary PCR (mg/mmol): or equivalent	25-50 or >25 and improved by ≥25%	1	50-100 and unimproved	10	>100 and unimproved (1000)		
		GFR (ml/min per 1.73m ²):	<80 ml/min & <67% of previous	100	<50 ml/min & previously > 50	100			
		Serum creatinine:	>130 μmol/L & >130% of previous	100	>130 μmol/L & >115% previous	10			
		Blood pressure (mmHg):	Accelerated HTN (↑ to >170/110 within 1/12 with retinal changes)	100	BP >140/90 & ↑ by 30 systolic or 15 diastolic	1			
Update the final domain scores with any items recorded on page two		No other items? BILAG COMPLETE							

Annotated Easy-BILAG page 1 in full demonstrates the structure and scoring mechanism for the most frequently encountered SLE features in addition to the constitutional and renal domains in full. This single page is sufficient to assess the majority of patients. Items are organized within tables by organ domain. Each is anchored to its relevant colour-coded key to the left. Standard BILAG-2004 principles apply throughout. The highest tariff item triggers the final domain score. Scoring requires clinicians to circle each listed item as 'not present' (0), 'improving' (1), 'same' (2), 'worse' (3), or 'new' (4), and colour-coding translates to overall organ domain score from A (blue, highly active), B

(continued)

Easy-BILAG matches general hospital and tertiary centre clinicians for accuracy and speed

The performance of Easy-BILAG outside centralized, subspeciality and research-orientated tertiary centres was evaluated. Accuracy in grading active domains [% accuracy; median (Q_1 , Q_3)] was high among tertiary centre clinicians using standard format BILAG-2004 [$n=9$; 86.9 (70.7, 95.7) Fig. 3D] but still trended to higher accuracy among those using Easy-BILAG [$n=11$; 95.7 (90.2, 96.7), $P=0.06$ Fig. 3D]. In contrast, general hospital clinicians using standard format BILAG-2004 returned significantly lower accuracy [$n=8$; 75.0 (64.1, 83.7)] than those testing Easy-BILAG [$n=5$; 91.3 (90.2, 96.7), $P=0.02$; Fig. 3D] while the latter achieved accuracy comparable to tertiary centre colleagues ($P=0.70$). Among general hospital clinicians, workbook completion was significantly faster [time in min; median (Q_1 , Q_3)] with Easy-BILAG ($n=5$; 59.0 (51.0, 63.0), Fig. 3E) than with standard format [$n=8$; 97.0 (76.0, 113.0), $P=0.01$; Fig. 3E]. Easy-BILAG achieved similarly high and matched levels of accuracy between clinicians with and without prior BILAG-2004 training, and between clinicians who regularly or infrequently use BILAG-2004 in their existing practice (Supplementary Fig. S3, available at *Rheumatology* online).

Easy-BILAG improved inter-rater agreement on disease activity

Overall inter-rater agreement in workbook disease activity grading using standard BILAG-2004 format was classified as good, though levels of agreement in assigning active grade B disease were lower (Table 1). Clinicians testing Easy-BILAG demonstrated higher levels of inter-rater agreement overall and across each level of active disease graded A–C where agreement was classified as very good (Table 1).

Fig. 1 Continued

(pink, moderately active), C (yellow, mild stable disease) to D (white, no disease activity, but domain previously active) or E (white, no current or prior disease activity). Free text space is included to the right of each domain for the clinician to assign the final domain score from A–E. An abbreviated glossary definition is included immediately adjacent to each item. Example shows ‘Skin eruption—severe’ recorded as ‘improving’ (pink) triggering overall mucocutaneous domain grade B; ‘Arthritis (mild)/arthralgia’ recorded as ‘same’ (yellow) triggering overall musculoskeletal domain grade C. Cardiorespiratory domain example is currently in remission prompting grade D when history indicates the domain has previously been active. Grade E is assigned in the neuropsychiatric domain where there is no current or prior disease activity. The haematological domain requires clinicians to circle where each full blood count item is located and indicate whether ‘Other features’ are present. Haemoglobin is scored separately according to whether haemolysis is present or absent. The severity of anaemia/cytopaenia translates to overall domain score based on the same colour-coded key (left) used throughout. Example shows overall grade C triggered by ‘haemoglobin 8–10.9 g/dl without haemolysis’ and ‘lymphocyte count $<1.0 \times 10^9/L$ ’ (both yellow) with no ‘Other features’ present. Gastrointestinal and Ophthalmic domains prompt clinicians to check page 2 of Easy-BILAG for relevant features or assign grade D or E (see Fig. 2). Constitutional and Renal domains derive domain scores using a tally of arbitrary points assigned to each item. A separate key anchored to both domains directs overall scoring according to points tally. Example shows constitutional domain grade C triggered by ‘Anorexia’ and ‘Weight loss’ rated as ‘improving’ (points tally 20, yellow) and renal domain grade A triggered by ‘Biopsy nephritis past 3 months’ and ‘urine protein : creatinine ratio >100 and unimproved’ (points tally 2000, blue). If no items signpost to page 2, as in this example, the assessment is complete.

Easy-BILAG has high usability for routine practice

Clinician feedback was surveyed by Likert scale responses on completion of the validation exercise. All clinicians testing Easy-BILAG agreed that the format was intuitive and sufficiently detailed, while 12/16 (75.0%) agreed Easy-BILAG was a convenient format for disease activity recording (Fig. 3F). By visual analogue scale [VAS 0—highly unlikely; 10—highly likely; median, (Q_1 , Q_3)], clinicians from both tertiary and general hospitals testing Easy-BILAG reported significantly higher likelihood of adopting the tool in regular practice [$n=16$, VAS: 8.5 (7.5, 10.0); Fig. 3G] than those using standard format BILAG-2004 [$n=16$ VAS: 5.0 (3.0, 8.5); $U=68.0$, $P=0.01$].

Discussion

We report a novel tool for recording the BILAG-2004 disease activity index that combines full fidelity to the current instrument [11] with a condensed, data driven, colour-coded design adapted to assist clinicians to assess SLE in routine practice. In this validation exercise it yielded superior accuracy and time efficiency over current standard format BILAG-2004 index using paper forms, with a particularly marked advantage among clinicians based outside tertiary centres. It showed improved inter-rater reliability and perceived usability for routine practice among clinicians of varying levels of prior experience across different hospital settings. Thus Easy-BILAG offers a tool with which clinicians can confidently, accurately and time-efficiently integrate BILAG-2004 index into routine practice.

Treat-to-target principles have demonstrated wide-ranging benefits on patient outcomes across rheumatic diseases and it is clear that this approach reduces both flare rates and damage accrual in SLE [4]. To target

Fig. 2 Easy-BILAG page 2 captures infrequent and rare BILAG-2004 items

A

NeuroPsych	Mononeuropathy (single or multiplex)	Supportive electrophysiology required	0	1	2	3	4	A
	Polyneuropathy	Acute symmetrical distal sensory and/or motor deficit (supportive electrophysiology required)	0	1	2	3	4	
	Seizure disorder	Independent description of seizure by reliable witness	0	1	2	3	4	
	Other features	Any other CNS or PNS feature due to active SLE	No	YES → NEXT PAGE				

B

	ITEM	ABBREVIATED GLOSSARY	Nil	Imp	Same	Worse	New
Gastrointestinal	Lupus peritonitis	Serositis presenting as acute abdomen with rebound/guarding.	0	1	2	3	4
	Abdominal serositis or ascites	Not presenting as acute abdomen.	0	1	2	3	4
	Lupus enteritis / colitis	Vasculitis or inflammation of small or large bowel, with supportive imaging &/or biopsy.	0	1	2	3	4
	Malabsorption	Diarrhoea + abnormal D-xylose absorption / ↑faecal fat losses. Exclude Coeliac & gut vasculitis.	0	1	2	3	4
	Protein-losing enteropathy	See detailed glossary.	0	1	2	3	4
	Intestinal pseudo-obstruction	Subacute intestinal obstruction due to intestinal hypomotility.	0	1	2	3	4
	Lupus hepatitis	Raised transaminases, without AIH specific autoantibodies. Exclude drug- & viral hepatitis.	0	1	2	3	4
	Acute lupus cholecystitis	Exclude gallstones or infection.	0	1	2	3	4
	Acute lupus pancreatitis	Usually associated with multisystem involvement.	0	1	2	3	4
Ophthalmic	Orbital inflam / myositis / proptosis	Orbital inflammation + myositis / extra-ocular muscle swelling / proptosis. Imaging required.	0	1	2	3	4
	Keratitis - severe	Sight-threatening. Includes corneal melt and peripheral ulcerative keratitis.	0	1	2	3	4
	Keratitis - mild	Not sight-threatening.	0	1	2	3	4
	Anterior uveitis		0	1	2	3	4
	Post. uveitis/retinal vasculitis - severe	Sight-threatening and/or retinal vasculitis not due to vaso-occlusive disease.	0	1	2	3	4
	Post. uveitis/retinal vasculitis - mild	Not sight-threatening. Not due to vaso-occlusive disease.	0	1	2	3	4
	Episcleritis		0	1	2	3	4
	Scleritis - severe	Necrotising anterior scleritis. Ant &/or post scleritis requiring systemic therapy.	0	1	2	3	4
	Scleritis - mild	Anterior/posterior scleritis not requiring systemic steroids.	0	1	2	3	4
	Retinal / choroidal vaso-occlusive disease	See detailed glossary table.	0	1	2	3	4
	Isolated cottonwool spots	Also known as cystoid bodies.	0	1	2	3	4
	Optic neuritis	Exclude anterior ischaemic optic neuropathy.	0	1	2	3	4
Anterior ischaemic optic neuropathy	Visual loss with pale swollen optic disc due to occlusion of posterior ciliary arteries.	0	1	2	3	4	

OTHER PAGE 1 FEATURES: Record here and update the score for that domain on page one:

Mucocutaneous	Angio-oedema - severe	Urticaria variant in subcut, submucosal & deep dermal tissues. Potentially life-threatening.	0	1	2	3	4
	Angio-oedema - mild	As above but not life-threatening.	0	1	2	3	4
	Panniculitis/ bullous lupus - severe	See detailed glossary table.	0	1	2	3	4
	Panniculitis/ bullous lupus - mild	Affects <9% BSA & does not fulfil any criteria for severe panniculitis.	0	1	2	3	4
	Major cutaneous vasculitis/thrombosis	Cutaneous vasculitis/thrombosis → extensive gangrene /ulceration / skin infarction.	0	1	2	3	4
Spilinter haemorrhages		0	1	2	3	4	
MSK	Myositis - severe	Significantly ↑ muscle enzymes with significant muscle weakness.	0	1	2	3	4
	Myositis - mild	Significantly ↑ muscle enzymes + myalgia but no significant muscle weakness.	0	1	2	3	4
Cardiorespiratory	Myocarditis - mild	↑ cardiac enzymes &/or ECG changes. No heart failure/arrhythmia/valve dysfunction.	0	1	2	3	4
	Myo /Endocarditis + cardiac failure	See detailed glossary table.	0	1	2	3	4
	Arrhythmias	Due to myocarditis / non-infective inflammation. ECG evidence required.	0	1	2	3	4
	New valvular dysfunction	Due to myocarditis / non-infective inflammation. Supportive imaging required.	0	1	2	3	4
	Cardiac tamponade	Supportive imaging required.	0	1	2	3	4
	Pulmonary haemorrhage/vasculitis	With haemoptysis &/or dyspnoea &/or pulmonary HTN. Imaging &/or histology required.	0	1	2	3	4
	Shrinking lung syndrome	Acute ↓ lung volumes (< 70% predicted) + normal corrected Kco. Diaphragmatic dysfunction.	0	1	2	3	4
Neuropsychiatric	Aortitis	+/- dissection with supportive imaging, claudication, bruits or BP discrepancy >10 mmHg.	0	1	2	3	4
	Coronary vasculitis	Imaging evidence of non-atheromatous coronary narrowing/ obstruction /aneurysm.	0	1	2	3	4
	Aseptic meningitis	See detailed glossary table.	0	1	2	3	4
	Cerebral vasculitis	With features of vasculitis in another system. Supportive imaging &/or biopsy required.	0	1	2	3	4
	Demyelinating syndrome	Discrete white matter lesion + neurological deficit. Ideally ≥1 prior recorded event. Exclude MS.	0	1	2	3	4
	Myelopathy	Acute onset, rapidly evolving paraparesis, quadriparesis and/or sensory level. Exclude SOL.	0	1	2	3	4
	Acute confusional state	See detailed glossary table.	0	1	2	3	4
	Psychosis	Delusions &/or hallucinations. Excludes primary psychotic disorder, drugs or during delirium.	0	1	2	3	4
	Acute inflammatory demyelinating polyradiculoneuropathy	See detailed glossary table.	0	1	2	3	4
	Cranial neuropathy	Exclude optic neuropathy which is classified under ophthalmic system.	0	1	2	3	4
Plexopathy	Supportive electrophysiology study required.	0	1	2	3	4	
Haem	Cognitive dysfunction	Sufficient to impair ADLs. Includes attention, memory, language, visuospatial, psychomotor.	0	1	2	3	4
	Status epilepticus	A seizure or seizures lasting >30 minutes without full recovery to baseline.	0	1	2	3	4
	Cerebrovascular disease	Not vasculitis. See detailed glossary table.	0	1	2	3	4
	Movement disorder	Exclude drug-induced.	0	1	2	3	4
	Autonomic disorder	See detailed glossary table.	0	1	2	3	4
	Cerebellar ataxia	Cerebellar ataxia in isolation of other CNS features. Usually subacute presentation.	0	1	2	3	4
	Severe lupus headache	Unremitting, disabling, unresponsive to narcotics, >3 days. Exclude SOL and CNS infection.	0	1	2	3	4
	Headache from intracranial hypertension	Exclude cerebral sinus thrombosis.	0	1	2	3	4
	TTP	Micro-angiopathic haemolytic anaemia + thrombocytopenia. Other causes excluded.	0	1	2	3	4
	Coombs positive (isolated)	Without evidence of haemolysis.	Negative	Positive			

An annotated example Easy-BILAG page 1 Neuropsychiatric domain (A) shows ‘Mononeuropathy’, ‘Polyneuropathy’ and ‘Seizure disorder’ rated as ‘not present’ but evidence of ‘Other features’ prompts review of the full Neuropsychiatric domain on Easy-BILAG page 2 (B). Easy-BILAG page 2 in full (B) shows the scoring template for uncommon and rare SLE features not captured by page 1. Items are organized within tables by organ domain. Gastrointestinal and Ophthalmic domains are represented in full, followed by less frequently scored spillover items from the Mucocutaneous (Mucocutan), Musculoskeletal (MSK), Cardiorespiratory, Neuropsychiatric and Haematological (Haem) domains, signposted on page 1. Colour-coding translates to overall organ domain scores from A (blue), B (pink), C (yellow) to D or E (white) as directed by the key on Easy-BILAG page 1. Example shows ‘Psychosis’ rated as ‘new’ (grey arrowheads) thereby triggering overall grade A in the Neuropsychiatric domain free text space on the main page (A). The rater need only complete specific sections of page 2 if directed to by prompts on page 1, remaining sections can be left unmarked.

remission and low disease activity states, clinicians must be equipped with tools that allow them to measure disease activity accurately and consistently. BILAG-2004 index is a highly comprehensive disease activity instrument which assures clinicians that they have performed a thorough SLE assessment. It can therefore support and prompt clinical decision making, but its administrative burden is high and inadequately adapted to routine practice. The other major disease activity instrument, SLEDAI, is often felt to be the quicker and easier instrument [9] as it offers a limited item, fixed scoring format. It does, however, have different limitations, for example, some features such as autoimmune haemolysis are not scored at all, and due to the fixed points weighting, features such as lupus rash or thrombocytopaenia, however severe, can never in isolation translate to the high disease activity score. Further, arthritis always scores twice as highly as rash even if a mild arthritis and severe rash are present. Being a binary system it does not distinguish items that are partially improved from those that are unchanged or worse. Therefore SLEDAI does not always align well with the patient experience or physician's intention to treat [15]. By re-ordering items by their frequency and screening questions for rarer manifestations Easy-BILAG offers a balance between the simplicity and speed of the SLEDAI for common features while retaining the sensitivity to change and scope of BILAG-2004.

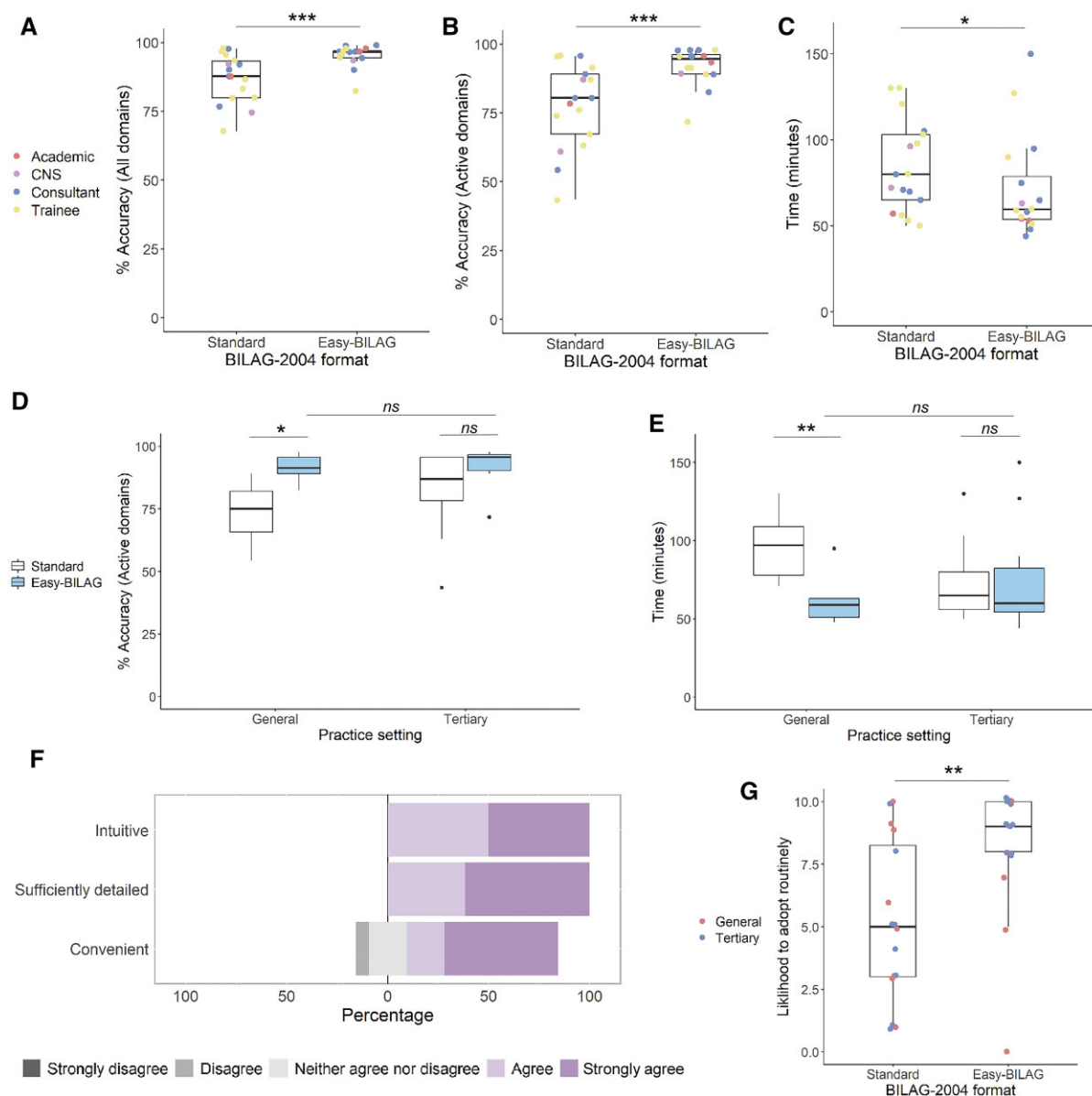
User feedback indicates Easy-BILAG had substantially greater appeal and usability for routine practice than standard format BILAG-2004 index. Thus Easy-BILAG permitted significant time saving for clinicians. Using Easy-BILAG, professionals with a range of prior BILAG-2004 experience completed 10 validation case assessments in an average time of just below 6 min per case. This time accounts for reading and evaluating the case material, and the cases were designed to cover both rare and common manifestations. Although real-world time required to document and derive a score for a patient has not been evaluated here, the present data suggest a potential time saving in clinical settings. Completing BILAG-2004 index accurately in standard published format requires consulting a clinical glossary and a separate scoring algorithm for each organ domain. Easy-BILAG makes use of colour-coding and abridged glossary descriptions to reduce reliance on separate reference documents. Using data-driven design, the same BILAG-2004 assessment would require only a single page assessment for the majority of UK registry patients. Since its design was informed by item frequency in a biologics registry it is likely that far more routine outpatient visits would be less complex and quicker.

Easy-BILAG facilitated excellent scoring accuracy when tested against model case vignettes. Crucially, a similarly high level of accuracy was achieved by clinicians based in general hospital practice compared with those in tertiary centres where more subspecialist and research-focused activities using BILAG-2004 are

typically concentrated. We also found that similar accuracy was maintained irrespective of prior training or current use of BILAG-2004 index. It is important to note that our validation exercise included a detailed overview and scoring instructions for BILAG-2004 index. Therefore Easy-BILAG is not a replacement for appropriate training, but our findings do suggest it is an accessible format for clinicians across all major areas of practice and can be more readily applied by those with less prior experience. Training material used in the current validation exercise is also made available alongside Easy-BILAG through the University of Leeds licensing platform.

Easy BILAG not only showed higher scoring accuracy, but also substantially less variability. We found that in this group of real-world clinicians with a range of prior BILAG-2004 experience, Easy-BILAG achieved better inter-rater agreement across all disease activity grades than standard format scoring. Importantly variation was not simply a function of clinical role, practice setting or prior experience as Easy-BILAG appeared to show an advantage among all professional categories. In practice this has implications for continuity of care where a team of various clinicians may review or manage an individual during follow-up and make treatment decisions such as whether to continue a biologic therapy. In the validation exercises, scoring with standard format appeared particularly vulnerable to inconsistencies assigning scores to grade B disease, which has also been observed in defining moderate flares [19]. Easy-BILAG was particularly beneficial in these cases and among these mixed-experience clinicians. Since grade B disease in two or more organ domains defines moderate disease activity in current EULAR, BSR and NHS England guidelines [1, 2, 6], this is a particularly decisive aspect of scoring where inconsistencies could conceivably introduce unwarranted variation in access to therapies. Improved recording accuracy might improve the identification of flares in clinical practice and the appropriate referral for new therapies.

Although our main goal in development of the Easy-BILAG was for use in routine clinical practice, our findings may have implications for clinical studies. Robust disease activity grading is central to trial outcome measures and the success or failure of new therapies in SLE. While detailed training in BILAG-2004 is essential for clinical investigators, we found that even among experienced lupus clinicians BILAG-2004 accuracy can be enhanced and variability reduced by formatting the assessment in a novel way. Investigators should be mindful of how data collection tools are presented to clinicians and the example of Easy-BILAG may be relevant to disease areas beyond SLE. Specific validation of Easy-BILAG in clinical trials has yet to be undertaken and is beyond the scope of this study but we would suggest that its potential advantage in scoring accuracy could reduce the burden of disease activity adjudication required in clinical trials. We would also anticipate that integrating BILAG-2004 into more

Fig. 3 Accuracy and usability of Easy-BILAG in a case-based validation exercise

Box plots (A–C) show performance of Easy-BILAG against standard format BILAG-2004 in self-timed validation workbook exercise for all professional categories (overlay jitter points). Scoring accuracy (%) against model workbook answers for all organ domains (A) and active organ domains (B) was significantly higher with Easy-BILAG. Time taken to complete workbook exercise (C) was significantly shorter with Easy-BILAG. Box plots (D–E) show Easy-BILAG (blue) versus standard format BILAG-2004 (white) across general and tertiary practice. Scoring accuracy (D) was significantly higher and time taken (E) reduced among general hospital rheumatologists using Easy-BILAG and equivalent to tertiary rheumatologists. Stacked horizontal bar chart (F) shows Easy-BILAG rated favourably for intuitiveness, detail and convenience on 5-point Likert-scale. Box plot (G) shows likelihood to adopt instrument in routine practice was significantly higher by 10-point visual analogue scale (10—highly likely, 0—highly unlikely) for Easy-BILAG among clinicians across general and tertiary practice (overlay jitter). *** $P \leq 0.001$; ** $P \leq 0.01$; $P \leq 0.05$; ns: non-significant, $P > 0.05$

routine practice through use of Easy-BILAG could facilitate more robust data collection in clinical practice and better enable clinicians to take part in clinical trials and identify eligible patients.

The current work has some limitations. First, due to constraints on space not all glossary items could be incorporated for rare items scored on Easy-BILAG page 2 and clinicians would still need to refer to core BILAG-

TABLE 1 Inter-rater agreement in BILAG-2004 domain scoring by validation workbook format

	Easy-BILAG <i>n</i> = 15		Standard BILAG-2004 Index <i>n</i> = 17	
	Fleiss' kappa coefficient, κ	Classification	Fleiss' kappa coefficient, κ	Classification
Overall	0.892	Very good	0.753	Good
BILAG grade				
A	0.877	Very good	0.724	Good
B	0.833	Very good	0.558	Moderate
C	0.882	Very good	0.698	Good
D	0.751	Good	0.627	Good
E	0.970	Very good	0.924	Very good

2004 official glossary if assessing these features. Second, as validation work was conducted on expert standardized training material, Easy-BILAG is yet to be evaluated in real-world practice with real patients. Its ability to maintain accuracy in capturing disease flares and response to therapy requires further validation. Similarly the time advantage in completing training material may not reflect use in the clinical or research setting and this would require confirmation in a further study. Additionally, applying criteria for improvement and worsening criteria as well as the 'trickle-down' rule all still require appropriate understanding and training on BILAG-2004 index. An optional self-adjudication checklist at the end of Easy-BILAG has been devised as a prompt for these areas. Computer-assisted formats with inbuilt quality assurance on all aspects of BILAG-2004 can offer additional error checking, which is particularly valuable in clinical trials [10]. As with the current published BILAG-2004 format, Easy-BILAG has been written in English, which could pose a barrier to use among clinicians and investigators for whom English is not first language. The validation work undertaken did not specifically characterize Easy-BILAG usability in non-first language English speakers.

In conclusion, Easy-BILAG should help to support the use of BILAG-2004 index in routine practice, assisting comprehensive and consistent assessment of patients and detection of disease activity, which in turn could facilitate compliance with biologic and other treatment guidelines. It is our recommended format for recording BILAG-2004 disease activity assessments in future clinical practice.

Acknowledgements

This research was supported by the United Kingdom National Institute for Health Research Leeds Biomedical Research Centre based at the Leeds Teaching Hospitals NHS Trust. The views expressed are those of the

authors and not necessarily those of the NHS, the NIHR or the Department of Health.

We thank rheumatology clinicians from the following centres for their participation in the Easy-BILAG validation exercise: Leeds Teaching Hospitals NHS Trust, Sandwell and West Birmingham NHS Foundation Trust, Doncaster & Bassetlaw Teaching Hospitals NHS Foundation Trust, Royal National Hospital for Rheumatic Diseases, Oxford University Hospitals NHS Foundation Trust, Royal Berkshire NHS Foundation Trust, Frimley Health NHS Foundation Trust—Wexham Park, University College London Hospitals NHS Foundation Trust, University Hospitals Birmingham NHS Foundation Trust, Manchester University NHS Foundation Trust, East Lancashire Hospitals NHS Trust—Royal Blackburn Teaching Hospital, Belfast Health and Social Care Trust—Musgrave Park, Cambridge University Hospitals NHS Foundation Trust and Sheffield Teaching Hospitals NHS Foundation Trust.

Easy-BILAG is copyright of Edward Vital, University of Leeds. It is available for free use for routine clinical practice and academic studies, and for commercial purposes under license at <https://licensing.leeds.ac.uk/products/healthcare-questionnaires>.

Funding: This work was not supported by any dedicated grant funding.

Disclosure statement: Caroline Gordon reports personal fees for honoraria from consultancy work from the Center for Disease Control and Prevention, AbbVie, Amgen, AstraZeneca, MGP, Sanofi and UCB, and personal fees for Speakers Bureau from UCB.

David Isenberg has consulted for AstraZeneca, Merck Serono, Eli Lilly, Servier and Genentech. The honoraria offered are passed onto a local arthritis charity.

Sarah Skeoch reports speaker fees from Pfizer and grant income from Innovative Medicines Initiatives 2 Joint Undertaking under grant agreement number

116106. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

Edward Vital has received research grants paid to his employer from AstraZeneca and Sandoz, and personal fees from Genentech, Aurinia, Lilly, AstraZeneca, ILTOO and Modus Therapeutics.

Lucy Carter and Chee-Seng Yee declare no conflicts of interest.

Data availability statement

The data underlying this article will be shared on reasonable request to the corresponding author.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

- Fanouriakis A, Kostopoulou M, Alunno A *et al.* 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* 2019;78:736–45.
- Gordon C, Amisssah-Arthur M-B, Gayed M *et al.*; British Society for Rheumatology Standards, Audit and Guidelines Working Group. The British Society for Rheumatology guideline for the management of systemic lupus erythematosus in adults: executive summary. *Rheumatology* 2018;57:14–8.
- van Vollenhoven RF, Mosca M, Bertsias G *et al.* Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. *Ann Rheum Dis* 2014;73:958–67.
- Golder V, Kandane-Rathnayake R, Huq M *et al.* Lupus low disease activity state as a treatment endpoint for systemic lupus erythematosus: a prospective validation study. *Lancet Rheumatol* 2019;1:e95–e102.
- van Vollenhoven R, Voskuyl A, Bertsias G *et al.* A framework for remission in SLE: consensus findings from a large international task force on definitions of remission in SLE (DORIS). *Ann Rheum Dis* 2017;76:554–61.
- NHS England. Clinical Commissioning Policy Rituximab for refractory Systemic Lupus Erythematosus (SLE) in adults and post-pubescent children [200402P]. Commissioning Position July 2020 [Internet]. <https://www.england.nhs.uk/wp-content/uploads/2020/07/Rituximab-for-refractory-Systemic-Lupus-Erythematosus-in-adults-and-post-pubescent-children-2.pdf> (21 September 2021, date last accessed).
- Pearce FA, Rutter M, Sandhu R *et al.* BSR guideline on the management of adults with systemic lupus erythematosus (SLE) 2018: baseline multi-centre audit in the UK. *Rheumatology (Oxford)* 2021;60:1480–90.
- 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.
- 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.
- Murphy CL, Yee CS, Gordon C, Isenberg D. From BILAG to BILAG-based combined lupus assessment—30 years on. *Rheumatology (Oxford)* 2016;55:1357–63.
- Yee CS, Cresswell L, Farewell V *et al.* Numerical scoring for the BILAG-2004 index. *Rheumatology (Oxford)* 2010; 49:1665–9.
- Yee CS, Farewell V, Isenberg DA *et al.* The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. *Rheumatology (Oxford)* 2009;48:691–5.
- Isenberg DA, Rahman A, Allen E *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.
- 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.
- Yee CS, Isenberg DA, Prabu A *et al.* BILAG-2004 index captures systemic lupus erythematosus disease activity better than SLEDAI-2000. *Ann Rheum Dis* 2008;67: 873–6.
- McCarthy EM, Sutton E, Nesbit S *et al.*; British Isles Lupus Assessment Group Biologics Register. Short-term efficacy and safety of rituximab therapy in refractory systemic lupus erythematosus: results from the British Isles Lupus Assessment Group Biologics Register. *Rheumatology (Oxford)* 2018;57:470–9.
- Gamer M, Lemon J, Fellows I, Singh P. irr: Various Coefficients of Interrater Reliability and Agreement. 2019 [R package version 0.84.1]; <https://CRAN.R-project.org/package=irr> (19 September 2021, date last accessed).
- Altman D, ed. *Practical statistics for medical research*. 2nd edn. London: Chapman & Hall, 2020.