

Use of Physician Global Assessment (PGA) in Systemic lupus erythematosus: a systematic review of its psychometric properties

Elisabetta Chessa¹, Matteo Piga¹, Alberto Floris¹, Hervé Devilliers², Alberto Cauli¹, Laurent Arnaud^{3,4}

¹ Rheumatology Unit, AOU University Clinic and University of Cagliari, Cagliari, Italy

² Centre Hospitalier Universitaire de Dijon, Hôpital François-Mitterrand, service de médecine interne et maladies systémiques (médecine interne 2) et Centre d'Investigation Clinique, Inserm CIC-EC 1432, Dijon, France.

³ Service de rhumatologie, Hôpitaux Universitaires de Strasbourg, Université de Strasbourg, Strasbourg, France.

⁴ Centre National de Référence des Maladies Systémiques et Autoimmunes Rares Est Sud-Ouest (RESO), Strasbourg, France.

Corresponding author:

Ms Elisabetta Chessa, Rheumatology Unit, AOU University Clinic and University of Cagliari, Cagliari, Italy. Email: elis.chessa@gmail.com

Acknowledgment: The authors wish to thank Ms. Sylvie Thuong for her invaluable assistance in the preparation of the manuscript.

Funding: The study was funded through the training Bursary Programme 2019 of the SLEuro European Lupus Society (recipient of the bursary: Ms Chessa).

ABSTRACT

Background

Physician Global Assessment (PGA) is a visual analogue score (VAS) that reflects the clinician's judgment of overall Systemic Lupus Erythematosus (SLE) disease activity. The aim of this systematic literature review (SLR) is to describe and analyse the psychometric properties of PGA.

Methods

This SLR was conducted by two independent reviewers in accordance with the PRISMA statement. All articles published until the 1st of July 2019 in Pubmed were screened with no limitation about years of publication, language or patients' age. Psychometric properties data were analysed according to the OMERACT Filter methodology version 2.1.

Results

The literature search identified 91 studies. Face validity was reported in all the articles retrieved, in which the PGA was used alone or as part of composite indices (SRI, SFI, LLDAS, DORIS remission criteria). Content validity was reported in 89 studies. Construct validity was demonstrated by a good correlation ($r \geq 0.50$) between the PGA with the SLEDAI (12 studies), SLAM (4 studies), LAI, BILAG and ECLAM (2 studies each). Criterion validity was assessed exploring the PGA correlation with quality of life measurements, biomarkers levels and treatment changes in 28 studies, while no study has evaluated correlation with damage. A good responsiveness for PGA was shown in 8 studies. A high variability in scales was found, causing a wide range of reliability (ICC=0.67-0.98).

Conclusion

PGA is a valid, responsive and feasible instrument, while its reliability was impacted by the scale adopted, suggesting the major need for a standardization of its scoring.

1. Introduction

Systemic lupus erythematosus (SLE) is an immune-mediated multi-systemic disease characterised by a wide variability of clinical manifestations and a course frequently subject to unpredictable flares (1-2). Because of SLE complexity, the assessment of disease activity is particularly challenging (3). The judgment of whether a patient with SLE has active disease is a central question, both in routine patient management as well as in clinical research (4). A higher disease activity is an important predictor of both organ damage and mortality (2) and the attainment of at least a low disease activity is associated with a reduction of early damage (5-6).

In the last 30 years more than a dozen of scores have been derived to assess disease activity in SLE, but not all of these have proven valid and reliable tools. Nevertheless, the development of a comprehensive index for assessing disease activity still represents one of the most important challenges in SLE (7).

In 1988, Liang et al. (8) suggested the need for a disease activity instrument which would be valid, reliable and sensitive to change. They proposed a physician global scoring on a 10-cm visual analog scale (VAS) to be used as a gold standard to compare six disease activity instruments (4). The term “Physician Global Assessment” (PGA) was coined in 1991 by Petri et al (9) to address a disease activity index scored on a VAS ranging from 0 to 3, with an increase of ≥ 1.0 since the last visit indicating a flare. Subsequently, the PGA was incorporated in the SELENA flare index in 1999 (10), in the Systemic Responder Index in 2009 (3, 11-12) as well as in the definitions of low disease activity (LLDAS) (13) and various definitions of remission (14-15).

The last EULAR/ACR recommendations recommend the use of the PGA in the routine monitoring of SLE (16). However, no precise guidelines exist regarding the optimal use of the PGA in SLE, such as the adequate length of the VAS, the presence of anchored values, the incorporation of laboratory data and the time frame of assessment.

The aim of this systematic review is to describe and analyse the measurement properties of the PGA, including the validity, reliability, responsiveness and feasibility.

2. Material & Methods

2.1. Literature search

A systematic review of the literature was performed according to the PRISMA guidelines (17), searching for articles reporting on the use of PGA in SLE. The following search strategy was used

through MEDLINE via PubMed: (((("lupus erythematosus, systemic"[MeSH Terms] OR ("lupus"[TIAB] AND "erythematosus"[TIAB] AND "systemic"[TIAB]) OR "systemic lupus erythematosus"[TIAB] OR ("systemic"[TIAB] AND "lupus"[TIAB] AND "erythematosus"[TIAB]))) OR "SLE"[TIAB]) AND ("physician global assessment"[TIAB] OR "PGA"[TIAB])). Additional papers were obtained by checking the references from the selected studies. Reviews and case series with less than 5 patients were excluded. Retrieved papers were selected with no limitation about years of publication, language or patients' age. The last Medline search was performed on the 1st of July 2019. Once 2 investigators (E.C., M.P.) have independently selected the articles, initially on the basis of titles and abstracts then, if necessary, on the full texts, eligibility assessment was performed independently in a blinded standardized manner. Disagreements between investigators were solved by consensus. Whenever papers reported duplicate data, the most recent article was selected.

2.2 Extraction of Psychometric Properties of the PGA

Each study was examined in order to extract psychometric property data on PGA according to the OMERACT (Outcome Measures in Rheumatology) Filter methodology version 2.1 (18).

Measurement properties of the PGA were analysed according to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) terminology (19).

3. Results

3.1. Literature search

The literature search identified 93 articles, and 12 additional articles were retrieved from the reference list of those publications. A total of 91 articles were included in the study (Fig. 1), accounting for 49 longitudinal cohort studies, 25 cross-sectional studies, 7 Randomized Controlled Trials (RCT), 3 consensus conferences, 4 post-hoc analyses, 2 retrospective studies and 1 case-series.

3.2 Psychometric Properties of the PGA

The OMERACT defines an instrument fit as outcome measurement if it passes the three pillars of evidence: truth (that refers to validity), discrimination (that includes reliability and responsiveness) and feasibility.

3.2.1 Truth (validity)

Truth refers to whether the measure provided by the scores is able to measure what was intended for (18). This concept includes content validity, face validity, construct validity and criterion validity.

Content validity

Content validity pertains to the degree to which the instrument measures all facets of a construct of interest (20): this property is satisfied if the PGA is considered able to measure all aspects of disease activity in SLE, in a comprehensive way. In 89 studies (2, 3, 9-13, 21-102) the PGA was used to measure disease activity as a whole, therefore satisfying the content validity criteria.

The globality of measurement was intended in the form of a scale from 0 to 3 in 54 studies (2, 3, 9-10, 12-13, 21, 24-69, 103) in a 0-10 scale in 12 studies (4, 60, 70-79), 0-100 in 9 studies (27, 40, 78, 80-85), 0-7 Likert scale (11, 78, 80), 0-2 scale (53), 0-4 (86) and 0-5 (87). Different definitions of PGA retrieved through literature search are reported in table 1.

In 2 studies, the PGA-VAS was used to assess concepts other than disease activity: disease severity (4) and patients' compliance to treatments (103).

Few studies reported on whether serological activity should be incorporated in the PGA. Liang et al. (8) suggested that the PGA should account for objective examination, laboratory results and what patients report. Barr et al. (2, 21, 24) recommended PGA assessment prior to reviewing serological data, based only on clinical visit. This scoring modality was used for the SRI (3). In one study (25), PGA of disease activity resulted from the combination of clinical visit, laboratory markers evaluation and physician's knowledge of the patient disease history. In the absence of a consensus, Aranow (26) found a better correlation between the SLEDAI and the PGA when the latter was assessed taking into account laboratory test results.

Face validity

According to the OMERACT, face validity is the degree to which the instrument appears to match with the target domain, according to experts (18). Face validity is satisfied when the instrument is considered able to capture what it should capture (i.e. disease activity). This property is reported across all articles selected through this systematic review (2-4, 9-13, 21-103).

Supporting the face validity property, PGA was defined "gold-standard" in 11 studies (2, 10-11, 21, 32, 49, 67, 76, 78, 84, 88) and in 32 it was used as the reference to which other activity scores were

compared, such as the SLEDAI (4, 10, 13, 25, 27-28, 31, 33, 35-36, 41, 46-47, 50-51, 53, 62, 65, 68, 72-73, 76, 81, 96-99), BILAG (4, 27, 35-36, 46, 65, 72-73, 81, 98), SLAM (4, 72, 76, 99), LAI (68, 88), patient global assessment (ptGA) (81, 83-84) and ECLAM (35). However, it was used as a single outcome measure only in two studies (49, 100) whilst in the majority the PGA was scored together with another instrument (typically the SLEDAI) (2, 9, 11-12, 21, 24, 30, 32, 34, 37-40, 44-45, 48-49, 55, 58-59, 61, 63-64, 66-67, 74-75, 80, 82, 86, 89-95, 103).

In 16 studies the PGA was used as a mean to assess changes in disease activity after treatment (3, 12, 22, 29, 40, 42-43, 52, 54, 56-57, 80, 85, 87, 95, 98, 101). In one open-label study (43) the decrease in PGA score was considered the primary endpoint. In 32 studies, disease activity measured by PGA was compared to changes in laboratory markers, with the aim to correlate clinical and serological features (9, 21, 30, 32, 34, 37-39, 44-45, 48-49, 55, 58-61, 63-64, 66-67, 69, 71, 74-75, 82, 86, 89, 91-94).

The PGA was integrated in composite indices, including the definition of LLDAS (12-13, 29, 34, 37, 39-40) and remission (29, 37) (table1). In 11 retrieved studies (10, 13, 33, 36, 45, 48, 50, 55-56, 65, 96) the PGA was part of SELENA Flare Index (SFI) (104) and in 10 studies (3, 29, 40, 46, 52, 60, 69, 80, 94, 98) it was part of the Systemic Responder Index (SRI) (3) (both discussed in responsiveness paragraph). In an epratuzumab trial, the absence of deterioration of PGA (not >10% worsening) was one of the items to achieve BICLA (BILAG-Based Composite Lupus Assessment) response (105). In one study, PGA was part of a modified score to assess disease activity in pregnancy (SLEPDAI) (51).

Construct validity

Construct validity is the degree to which the PGA relates to other instruments which measure the same concept (18). This may be explored through convergent and divergent validity. Data regarding divergent validity are lacking for PGA. Convergent validity is fulfilled indirectly in studies where the PGA is used as the gold standard to assess the construct validity of other indices. Correlations with other instruments measuring similar constructs should typically demonstrate coefficient (r) \geq 0.50 (106).

Construct validity was recognized in 21 studies (2, 10-11, 23-24, 26-29, 35, 47, 52, 54, 65, 68, 76, 84, 88, 92, 99, 101). The correlation with the SLEDAI was determined in 12 studies (fig. 2) (10, 23-24, 26, 28-29, 35, 54, 68, 76, 84, 99), with the SLAM in 4 studies (r comprised between 0.47 and 0.65) (35, 76, 84, 99), with LAI in 2 studies (r =0.64-0.75) (68, 84), with BILAG in 2 studies

($r=0.61-0.62$) (35, 84) with ECLAM in 2 studies ($r=0.58-0.65$) (35, 84).

Criterion validity

Criterion validity is defined as the degree to which the scores of an instrument adequately reflect “the truth” in the form of a “gold standard” (107). In the absence of a well-recognised gold standard for disease activity, criterion validity of PGA is established when it correlates with a measure that the author of the study defined *a priori* as the gold standard. In Fatemi et al. (35) the PGA correlated, although moderately, with the need for treatment change ($r=0.46$, $p<0.01$).

Criterion validity also refers to the degree to which an instrument predicts aspects and phenomena occurring in the future (108). In this sense, criterion validity of PGA is satisfied when scores correlate with phenomena subsequently influenced by disease activity such as quality of life measurements (HRQoL, SF-36, FACIT-Fatigue score, Lupus Impact Tracker (LIT), LupusPRO), biomarkers levels (complement fractions, ESR, auto-antibodies), treatment variations and damage assessments (SDI).

Twenty-nine studies (25, 31, 34-39, 41, 43, 48-49, 53-55, 59, 61-64, 74-75, 81, 82, 86, 92-93, 96-97) have assessed criterion validity of PGA (table 2).

Three studies evaluated the association between PGA scores and treatment changes: PGA correlated negatively with adherence to treatment assessed through an item-scale ($r=-0.31$, $p=0.11$) (34); clinically defined mild and moderate flares had a higher disease activity by PGA ($P<0.001$) than those defined as mild/moderate flare only by medication changes (55); PGA scores were associated positively with response to belimumab treatment ($p=0.039$) (43).

A $PGA>1$ was predictive of polymorphic light cutaneous eruption ($p=0.02$) (59) and correlated negatively with LLDAS attainment (37). A $PGA\geq 2$ correlated with risk of pregnancy loss (29% vs 8%, $p=0.005$) (49). No study has evaluated the correlation of PGA with damage measures.

3.2.2 Discrimination

Discrimination refers to whether the score (PGA-VAS) differentiates between situations of interest (18): discrimination of PGA measures the ability of the PGA-VAS to report a consistent score where no change in disease activity has occurred (reliability) and to detect change when a change in disease activity has occurred (sensitivity to change or responsiveness).

Reliability

Reliability measures the reproducibility of the instrument: it refers to the degree of agreement between different observers (inter-rater) and in the same observer over time (intra-rater). The quantification of reliability is expressed by a correlation coefficient. An acceptable reliability is indicated by values of ICC or weighted Kappa (k) in excess of 0.60 and a good reliability is in excess of 0.85 (20).

Inter-rater reliability (Inter-RR) of PGA is the ability to provide consistent scores in a stable population between two or more physicians who evaluate disease activity of the same patient. Inter-RR was assessed in 7 studies (4, 10-11, 36, 65, 68, 94), showing values ranging from 0.67 (68) to 0.96 (94). FitzGerald and Grossman (10) found a good Inter-RR in retrospective assessment of PGA (k=0.79).

A difference between Inter-RR of PGA assessed by an untrained physician (ICC=0.5-0.63) or a trained investigator (ICC=0.79-0.81) was found (36).

The four-point PGA (0: no flare, 1 mild, 2 moderate, 3 severe) showed the lowest Inter-RR in assessing flare (ICC=0.18), compared to that of BILAG-2004 (ICC=0.54) and SFI (ICC=0.21) (65).

Intra-rater Reliability (Intra-RR) is the ability to provide consistent scores in a stable population by the same assessor over time. It estimates how similar a given patient scores were at the two visits. PGA Intra-RR was assessed in 3 studies (10, 68, 94), ranging from 0.55 (68) to 0.88 (10).

Responsiveness

Responsiveness, or sensitivity to change, is the usefulness of a test to detect minimum clinically important differences (MCID) (20, 109). Responsiveness of PGA is the ability to detect worthwhile variations in disease activity over time, measuring worsening or improvements in SLE disease status.

The assessment of PGA responsiveness was performed in 10 studies (4, 23, 50, 58, 77-79, 81, 83-84) using different methods (110). Only in one study (4), PGA sensitivity was assessed comparing the change with an anchor (109), represented by the treatment sensitive index (TSI): PGA sensitivity resulted comprised between that of the BILAG (highest sensitivity) and SLEDAI (lowest sensitivity). More frequently, responsiveness was assessed correlating changes in PGA with changes in other scores (23, 50, 58, 77, 78, 81, 83), finding a significant correlation with variations in SLEDAI (r ranging from 0.39-0.66) (23, 77-78), SLAM (0.61) (77), LAI (0.56) (77),

ptGA (0.37) (77), SRI-50 (0.48) (78) and ESR ($p<0.0001$) (58), but not with C3, C4, circulating immunocomplexes and prednisone dose (77). One study showed a significant ability of the PGA in distinguishing between patients ($p<0.0001$), observers ($p<0.0001$), but not between visits (79). Ward et al (84) expressed the sensitivity in PGA scoring with the standardized response mean (SRM), demonstrating a very large effect size ($ES=2.23$) (110).

PGA responsiveness was used to assess flare (9): PGA was identified as the “gold standard” to rate the exacerbation of lupus activity (21, 67, 88), preliminarily defined by a change of ≥ 1.0 in a 0-3 VAS occurred since the last visit. To discriminate between severity of flares, PGA was incorporated in a composite index: the Sclera Flare Index (10) (table 1). Several definitions of MCID were retrieved: in SRI-4 a significant worsening was defined as an increase of $>10\%$ on the PGA-VAS (111) corresponding to ≥ 0.3 -points from baseline; Touma et al (80) considered worsening any increase in PGA from baseline; in the epratuzumab trial (87) a significant improvement was a 20% decrease in PGA score evaluated after 12 months of treatment.

In a post-hoc analysis of phase 3 Belimumab trials, improvements and no worsening in PGA were greater among SRI responders versus SRI non-responders ($p<0.001$) (52). Otherwise, a validation study of the SRI for juvenile SLE (60) showed that the exclusion of the BILAG or PGA from the SRI did not change the accuracy of the SRI in detecting improvement.

3.2.3 FEASIBILITY

Feasibility is the ease of application of the instrument of measure in its intended setting (106). Feasibility refers not to the quality of the outcome measure, but to aspects such as completion time, cost of an instrument, equipment, type and ease of administration. No study has evaluated the feasibility of the PGA in SLE to date.

4. DISCUSSION

The assessment of disease activity in SLE is particularly challenging. In this systematic review, we have analysed the measurement properties of the PGA, including the validity, reliability, responsiveness and feasibility.

In support of its face validity, the PGA was used to define disease activity score in all the 91 studies retrieved by the literature search, having a role as outcome measure as well as comparator to assess the validity of other indices. Nevertheless, despite the fact that the PGA was considered the reference in 39 studies involving other indices, it was used as the sole instrument in only 2 of them. This suggests that the role of the PGA is limited for disease activity assessment when used as single instrument. It is therefore desirable to use the PGA along with other tools (typically the SLEDAI) or to include PGA in a composite index (e.g., SFI, LLDAS, SRI, DORIS remission criteria) (3, 5, 10, 13-15, 104).

However, the PGA allows for the measurement of disease activity in a global way (content validity). It does not provide a predefined or limited list of disease manifestations or organ systems, therefore allowing to capture all the heterogeneous aspects of SLE disease activity. PGA is usually reported by experts as allowing an exhaustive coverage of the concept of disease activity in SLE (20, 108). Thanks to this feature, the PGA was included in composite indices with the aim to rate manifestations not included in glossary-based instruments such as the SLEDAI and BILAG (3) or for which a threshold has been defined (cytopenia).

It is noteworthy that the PGA correlates with several other instruments which measure disease activity. Construct validity is shown by the good correlation observed with the SLEDAI, BILAG, LAI and ECLAM (10, 23-24, 26, 28-29, 35, 54, 68, 76, 84, 99).

The PGA also showed a good predictive validity as it correlated significantly with measures of future outcomes, such as quality of life or laboratory exams but no study has currently evaluated its correlation with measures of damage.

Of note, the literature search revealed heterogeneous definitions of physician assessment of disease activity other than PGA (MD global (4, 70, 73), physician overall assessment [PHYOA] (85)). Different scores and lengths of visual evaluation were employed: the first was the 0-10 VAS suggested by Liang et al (8) and adopted in childhood SLE; the most common tool (the 0-3 VAS), was developed (68) to capture the concept of flare, and measured on 3 cm VAS in SRI (3) and 10 cm VAS in SFI (10, 104), but other scores (0-2, 0-4, 0-5, 0-7) (11, 53, 78, 80, 86-87) and lengths (8 cm, 15 cm) (10, 82-84) have also been used. Currently, no agreement has been reached upon which scale should be used: if a pointed scale with anchored values (0, 1, 2, 3) or a centimetric scale with all values comprised between 0.0 and 3.0. Moreover, there is an uncertainty as to whether the best timing of assessment is prior or after reviewing laboratory exams (26). The lack of standardized scoring, as well as the subjectivity of the physician judgments, can be an important source of heterogeneity, especially in trials.

Five studies have demonstrated good ICC values for reliability (all higher than 0.60 and ranging up to 0.97). Reliability was excellent when scored through a pointed scale, as the Likert scale which was anchored in unit numbers from 0 (not active) to 7 (most active) (Inter-RR ICC: 0.96; Intra-RR ICC: 0.88) (80) but lower when assessed through a centimetric VAS scale using values comprised between 0.0 and 3.0 (Inter-RR ICC: 0.67; Intra-RR ICC: 0.55) (68). This important heterogeneity in the anchoring of the PGA prevented us from performing a metanalysis of reliability data.

Physician training is very important. Even though the PGA showed optimal reliability, a very low Inter-rater reliability for flare using the PGA (ICC=0.18) was found in a single study (65), compared to that of the BILAG (ICC=0.54) or SFI (ICC=0.21). According to the authors, this difference was probably due to the greater familiarity of the physicians with the BILAG 2004 index. Moreover, a difference between Inter-RR of PGA assessed by an untrained physician (0.5-0.63) and a trained investigator (0.79-0.81) was found, suggesting the need for a PGA-scoring training or a standardization (36).

Responsiveness of the PGA was assessed through different methods (109, 111) showing a high sensitivity for detecting clinical variations (84). Changes in PGA correlated with changes of other disease activity indices (SLEDAI, SLAM, LAI, ptGA), laboratory exams (ESR), PROs (LIT) (23, 50, 58, 77-78, 81, 83) and response to treatment (4). These results enabled its use as a gold standard for assessing flare and defining flare severity in several studies (21, 67, 88).

No data were found regarding feasibility of the PGA. In most studies, the PGA was assessed by a rheumatologist experienced in SLE care or research and, as already stated, the ICC reliability was different between an untrained physician and a trained investigator (36).

Finally, the PGA enables the measurement of disease activity globally compared to a glossary-based index. At the same time, a global judgment does not allow the detailed capture of specific organ manifestations of the disease. Altogether, the PGA should ideally be used in a composite index or as part of a disease status definition.

In conclusion, the PGA was demonstrated to be a valid, responsive and feasible instrument, while its reliability was strongly impacted by the scale adopted, suggesting the major need for a standardization in its scoring.

Bibliography

1. Mikdashi J, Nived O. Measuring disease activity in adults with systemic lupus erythematosus: the challenges of administrative burden and responsiveness to patient concerns in clinical research. *Arthritis Res Ther*. 2015 Jul 20;17:183.
2. Barr SG, Zonana-Nacach A, Magder LS, Petri M. Patterns of disease activity in systemic lupus erythematosus. *Arthritis Rheum*. 1999;42(12):2682–8.
3. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, et al. Novel evidence-based systemic lupus erythematosus responder index. *Arthritis Care Res*. 2009;61(9):1143–51.
4. 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(9):1107–18.
5. Piga M, Floris A, Cappellazzo G, Chessa E, Congia M, Mathieu A, et al. Failure to achieve lupus low disease activity state (LLDAS) six months after diagnosis is associated with early damage accrual in Caucasian patients with systemic lupus erythematosus. *Arthritis Res Ther*. 2017;19:247.
6. Floris A, Piga M, Perra D, Chessa E, Congia M, Mathieu A, et al. Treatment target in newly diagnosed Systemic Lupus Erythematosus. *Arthritis Care Res* 2019 doi:10.1002/acr.24086.
7. Felten R, Sagez F, Gavand P, et al. 10 most important contemporary challenges in the management of SLE. *Lupus Science & Medicine* 2019;6:e000303. doi: 10.1136/lupus-2018-000303
8. Liang MH, Socher SA, Roberts WN, Esdaile JM. Measurement of systemic lupus erythematosus activity in clinical research. *Arthritis Rheum*. 1988 Jul;31(7):817–25.
9. Petri M, Genovese M, Engle E, Hochberg M. Definition, incidence, and clinical description of flare in systemic lupus erythematosus. A prospective cohort study. *Arthritis Rheum*. 1991 Aug;34(8):937–44.
10. FitzGerald JD, Grossman JM. Validity and reliability of retrospective assessment of disease activity and flare in observational cohorts of lupus patients. *Lupus*. 1999;8(8):638–44.
11. Touma Z, Gladman DD, Su J, Anderson N, Urowitz MB. A novel lupus activity index accounting for glucocorticoids: SLEDAI-2K glucocorticoid index. *Rheumatology*. 2018 Aug 1;57(8):1370–6.
12. Fanouriakis A, Adamichou C, Koutsoviti S, Panopoulos S, Staveri C, Klagou A, et al. Low disease activity—irrespective of serologic status at baseline—associated with reduction of corticosteroid dose and number of flares in patients with systemic lupus erythematosus treated with belimumab: A real-life observational study. *Semin Arthritis Rheum*. 2018 Dec 1;48(3):467–74.
13. Franklyn K, Lau CS, Navarra SV, Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). *Ann Rheum Dis*. 2016 Sep 1;75(9):1615–21.

14. van Vollenhoven R, Voskuyl A, Bertsias G, Aranow C, Aringer M, Arnaud L, 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 Mar;76(3):554–61.
15. Zen M, Iaccarino L, Gatto M et al . Prolonged remission in Caucasian patients with SLE: prevalence and outcomes.*Ann Rheum Dis*. 2015 Dec;74(12):2117-22.
16. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis*. 2019 Jun 1;78(6):736–45.
17. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009 Aug 18;151(4):264–9, W64.
18. Beaton DE, Maxwell LJ, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Instrument Selection Using the OMERACT Filter 2.1: The OMERACT Methodology. *J Rheumatol*. 2019 Aug;46(8):1028–35.
19. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010 Jul;63(7):737–45.
20. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford University Press; 2015. 10.1093/med/9780199685219.001.0001
21. Ho A, Barr SG, Magder LS, Petri M. A decrease in complement is associated with increased renal and hematologic activity in patients with systemic lupus erythematosus. *Arthritis Rheum*. 2001 Oct;44(10):2350–7.
22. Islam MN, Hossain M, Haq SA, Alam MN, Ten Klooster PM, Rasker JJ. Efficacy and safety of methotrexate in articular and cutaneous manifestations of systemic lupus erythematosus. *Int J Rheum Dis*. 2012 Feb;15(1):62–8.
23. Jesus D, Rodrigues M, Matos A, Henriques C, Pereira da Silva JA, Inês LS. Performance of SLEDAI-2K to detect a clinically meaningful change in SLE disease activity: a 36-month prospective cohort study of 334 patients. *Lupus*. 2019 Apr 1;28(5):607–12.
24. Györi N, Giannakou I, Chatzidionysiou K, Magder L, van Vollenhoven RF, Petri M. Disease activity patterns over time in patients with SLE: analysis of the Hopkins Lupus Cohort. *Lupus Sci Med*. 2017;4(1):e000192.
25. Quimby KR, Flower C, Hambleton IR, Landis RC, Hennis AJM. Comparison of the systemic lupus erythematosus activity questionnaire and the systemic lupus erythematosus disease activity index in a black barbadian population. *Int J Rheumatol*. 2013;2013:875369.
26. Aranow C. A pilot study to determine the optimal timing of the Physician Global Assessment (PGA) in patients with systemic lupus erythematosus. *Immunol Res*. 2015 Dec;63(1–3):167–9.
27. Askanase A, Li X, Pong A, Shum K, Kamp S, Carthen F, et al. Preliminary test of the LFA

- rapid evaluation of activity in lupus (LFA-REAL): an efficient outcome measure correlates with validated instruments. *Lupus Sci Med*. 2015;2(1):e000075.
28. Jesus D, Matos A, Henriques C, Zen M, Larosa M, Iaccarino L, et al. Derivation and validation of the SLE Disease Activity Score (SLE-DAS): a new SLE continuous measure with high sensitivity for changes in disease activity. *Ann Rheum Dis*. 2019;78(3):365–71.
 29. Parodis I, Emamikia S, Gomez A, Gunnarsson I, van Vollenhoven RF, Chatzidionysiou K. Clinical SLEDAI-2K zero may be a pragmatic outcome measure in SLE studies. *Expert Opin Biol Ther*. 2019 Feb;19(2):157–68.
 30. Merrill JT, Petri MA, Buyon J, Ramsey-Goldman R, Kalunian K, Putterman C, et al. Erythrocyte-bound C4d in combination with complement and autoantibody status for the monitoring of SLE. *Lupus Sci Med*. 2018;5(1):e000263.
 31. Gandhi N, Arora S, Sengupta M, Sequeira W, Jolly MM, Block JA. Validation of SIMPLE Index for Lupus Disease Activity. *J Clin Rheumatol*. 2018 Sep;24(6):313–8.
 32. Mahler M, Bentow C, O'Malley T, Ibarra C, Conklin J, Aure MAR, et al. Performance Characteristics of Different Anti-Double-Stranded DNA Antibody Assays in the Monitoring of Systemic Lupus Erythematosus. *J Immunol Res*. 2017;2017:1720902. doi: 10.1155/2017/1720902.
 33. Annapureddy N, Giangreco D, Devilliers H, Block JA, Jolly M. Psychometric properties of MDHAQ/RAPID3 in patients with systemic lupus erythematosus. *Lupus*. 2018 Jan 1;961203318758503.
 34. Mazur-Nicorici L, Sadovici-Bobeica V, Garabajiu M, Mazur M. Therapeutic adherence in patients with systemic lupus erythematosus: a cross-sectional study. *Romanian J Intern Med Rev Roum Med Interne*. 2018 Jun 1;56(2):109–15.
 35. Fatemi A, Raeisi A, Sayedbonakdar Z, Smiley A. Sensitivity analyses of four systemic lupus erythematosus disease activity indices in predicting the treatment changes in consecutive visits: a longitudinal study. *Clin Rheumatol*. 2018 Apr;37(4):955–62.
 36. Askanase AD, Nguyen SC, Costenbader K, Lim SS, Kamen D, Aranow C, et al. Comparison of the Lupus Foundation of America-Rapid Evaluation of Activity in Lupus to More Complex Disease Activity Instruments As Evaluated by Clinical Investigators or Real-World Clinicians. *Arthritis Care Res*. 2018;70(7):1058–63.
 37. Zen M, Iaccarino L, Gatto M, Saccon F, Larosa M, Ghirardello A, et al. Lupus low disease activity state is associated with a decrease in damage progression in Caucasian patients with SLE, but overlaps with remission. *Ann Rheum Dis*. 2018 Jan;77(1):104–10.
 38. Chaigne B, Chizzolini C, Perneger T, Trendelenburg M, Huynh-Do U, Dayer E, et al. Impact of disease activity on health-related quality of life in systemic lupus erythematosus - a cross-sectional analysis of the Swiss Systemic Lupus Erythematosus Cohort Study (SSCS). *BMC Immunol*. 2017 28;18(1):17.
 39. Golder V, Kandane-Rathnayake R, Hoi AY-B, Huq M, Louthrenoo W, An Y, et al. Association of the lupus low disease activity state (LLDAS) with health-related quality of life in a multinational prospective study. *Arthritis Res Ther*. 2017 20;19(1):62.

40. Parodis I, Sjöwall C, Jönsen A, Ramsköld D, Zickert A, Frodlund M, et al. Smoking and pre-existing organ damage reduce the efficacy of belimumab in systemic lupus erythematosus. *Autoimmun Rev.* 2017 Apr;16(4):343–51.
41. Schneider M, Mosca M, Pego-Reigosa J-M, Gunnarsson I, Maurel F, Garofano A, et al. Cross-cultural validation of Lupus Impact Tracker in five European clinical practice settings. *Rheumatol Oxf Engl.* 2017 01;56(5):818–28.
42. Forbess LJ, Bresee C, Wallace DJ, Weisman MH. Failure of a systemic lupus erythematosus response index developed from clinical trial data: lessons examined and learned. *Lupus.* 2017 Aug;26(9):909–16.
43. Vashisht P, Borghoff K, O'Dell JR, Hearsh-Holmes M. Belimumab for the treatment of recalcitrant cutaneous lupus. *Lupus.* 2017 Jul;26(8):857–64.
44. Khan A, Arbab-Zadeh A, Kiani AN, Magder LS, Petri M. Progression of noncalcified and calcified coronary plaque by CT angiography in SLE. *Rheumatol Int.* 2017 Jan;37(1):59–65.
45. Thanou A, Stavrakis S, Dyer JW, Munroe ME, James JA, Merrill JT. Impact of heart rate variability, a marker for cardiac health, on lupus disease activity. *Arthritis Res Ther.* 2016 02;18:197.
46. Devilliers H, Bonithon-Kopp C, Jolly M. The lupus impact tracker is responsive to changes in clinical activity measured by the systemic lupus erythematosus responder index. *Lupus.* 2017 Apr;26(4):396–402.
47. Antony A, Kandane-Rathnayake RK, Ko T, Boulos D, Hoi AY, Jolly M, et al. Validation of the Lupus Impact Tracker in an Australian patient cohort. *Lupus.* 2017 Jan;26(1):98–105.
48. Mok CC, Ding HH, Kharboutli M, Mohan C. Axl, Ferritin, IGFBP2 and TNFR2 as biomarkers in systemic lupus erythematosus. *Arthritis Care Res.* 2016 Sep;68(9):1303–9.
49. Mankee A, Petri M, Magder LS. Lupus anticoagulant, disease activity and low complement in the first trimester are predictive of pregnancy loss. *Lupus Sci Med.* 2015 Dec 9;2(1):e000095.
50. Giangreco D, Devilliers H, Annapureddy N, Block JA, Jolly M. Lupus Impact Tracker is responsive to physician and patient assessed changes in systemic lupus erythematosus. *Lupus.* 2015 Dec;24(14):1486–91.
51. Buyon JP, Kim MY, Guerra MM, Laskin CA, Petri M, Lockshin MD, et al. Predictors of Pregnancy Outcomes in Patients With Lupus: A Cohort Study. *Ann Intern Med.* 2015 Aug 4;163(3):153–63.
52. Furie R, Petri MA, Strand V, Gladman DD, Zhong ZJ, Freimuth WW, et al. Clinical, laboratory and health-related quality of life correlates of Systemic Lupus Erythematosus Responder Index response: a post hoc analysis of the phase 3 belimumab trials. *Lupus Sci Med.* 2014;1(1):e000031.
53. Mok CC, Kosinski M, Ho LY, Chan KL, Jolly M. Validation of the LupusPRO in Chinese Patients From Hong Kong With Systemic Lupus Erythematosus. *Arthritis Care Res.* 2015;67(2):297–304.

54. Ribi C, Trendelenburg M, Gayet-Ageron A, Cohen C, Dayer E, Eisenberger U, et al. The Swiss Systemic lupus erythematosus Cohort Study (SSCS) - cross-sectional analysis of clinical characteristics and treatments across different medical disciplines in Switzerland. *Swiss Med Wkly*. 2014;144:w13990.
55. Thanou A, Chakravarty E, James JA, Merrill JT. How should lupus flares be measured? Deconstruction of the safety of estrogen in lupus erythematosus national assessment-systemic lupus erythematosus disease activity index flare index. *Rheumatol Oxf Engl*. 2014 Dec;53(12):2175–81.
56. Mok CC, Ying SKY, Ma KM, Wong CK. Effect of raloxifene on disease activity and vascular biomarkers in patients with systemic lupus erythematosus: subgroup analysis of a double-blind randomized controlled trial. *Lupus*. 2013 Dec;22(14):1470–8.
57. Bello KJ, Fang H, Fazeli P, Bolad W, Corretti M, Magder LS, et al. Omega-3 in SLE: a double-blind, placebo-controlled randomized clinical trial of endothelial dysfunction and disease activity in systemic lupus erythematosus. *Rheumatol Int*. 2013 Nov;33(11):2789–96.
58. Stojan G, Fang H, Magder L, Petri M. Erythrocyte sedimentation rate is a predictor of renal and overall SLE disease activity. *Lupus*. 2013 Jul;22(8):827–34.
59. Foering K, Chang AY, Piette EW, Cucchiara A, Okawa J, Werth VP. Characterization of clinical photosensitivity in cutaneous lupus erythematosus. *J Am Acad Dermatol*. 2013 Aug;69(2):205–13.
60. Mina R, Klein-Gitelman MS, Nelson S, Eberhard BA, Higgins G, Singer NG, et al. Validation of the systemic lupus erythematosus responder index for use in juvenile-onset systemic lupus erythematosus. *Ann Rheum Dis*. 2014 Feb;73(2):401–6.
61. Karol DE, Criscione-Schreiber LG, Lin M, Clowse MEB. Depressive symptoms and associated factors in systemic lupus erythematosus. *Psychosomatics*. 2013 Oct;54(5):443–50.
62. Moorthy LN, Saad-Magalhães C, Sato JO, Len CA, Vasco MB, Appenzeller S, et al. Validation of the Portuguese Simple Measure of Impact of Lupus Erythematosus in Youngsters (SMILEY) in Brazil. *Lupus*. 2013 Feb;22(2):190–7.
63. Akhter E, Burlingame RW, Seaman AL, Magder L, Petri M. Anti-C1q antibodies have higher correlation with flares of lupus nephritis than other serum markers. *Lupus*. 2011 Oct;20(12):1267–74.
64. Mok CC, Ho LY, Leung HW, Wong LG. Performance of anti-C1q, antinucleosome, and anti-dsDNA antibodies for detecting concurrent disease activity of systemic lupus erythematosus. *Transl Res J Lab Clin Med*. 2010 Dec;156(6):320–5.
65. Isenberg DA, Allen E, Farewell V, D’Cruz D, Alarcón GS, Aranow C, et al. An assessment of disease flare in patients with systemic lupus erythematosus: a comparison of BILAG 2004 and the flare version of SELENA. *Ann Rheum Dis*. 2011 Jan;70(1):54–9.
66. Kiani AN, Vogel-Claussen J, Magder LS, Petri M. Noncalcified coronary plaque in systemic lupus erythematosus. *J Rheumatol*. 2010 Mar;37(3):579–84.
67. 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 Oct;44(10):2342–9.
68. Petri M, Hellmann D, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. *J Rheumatol*. 1992 Jan;19(1):53–9.
 69. Iaccarino L, Andreoli L, Bocci EB, Bortoluzzi A, Ceccarelli F, Conti F, et al. Clinical predictors of response and discontinuation of belimumab in patients with systemic lupus erythematosus in real life setting. Results of a large, multicentric, nationwide study. *J Autoimmun*. 2018;86:1–8.
 70. Brunner HI, Holland M, Beresford MW, Ardoin SP, Appenzeller S, Silva CA, et al. American College of Rheumatology Provisional Criteria for Global Flares in Childhood-Onset Systemic Lupus Erythematosus. *Arthritis Care Res*. 2018;70(6):813–22.
 71. Carmichael SJ, Day RO, Tett SE. A cross-sectional study of hydroxychloroquine concentrations and effects in people with systemic lupus erythematosus. *Intern Med J*. 2013 May;43(5):547–53.
 72. Mina R, Klein-Gitelman MS, Ravelli A, Beresford MW, Avcin T, Espada G, et al. Inactive disease and remission in childhood-onset systemic lupus erythematosus. *Arthritis Care Res*. 2012 May;64(5):683–93.
 73. Brunner HI, Mina R, Pilkington C, Beresford MW, Reiff A, Levy DM, et al. Preliminary criteria for global flares in childhood-onset systemic lupus erythematosus. *Arthritis Care Res*. 2011 Sep;63(9):1213–23.
 74. Vilá LM, Alarcón GS, McGwin G, Bastian HM, Fessler BJ, Reveille JD, et al. Systemic lupus erythematosus in a multiethnic US cohort, XXXVII: association of lymphopenia with clinical manifestations, serologic abnormalities, disease activity, and damage accrual. *Arthritis Rheum*. 2006 Oct 15;55(5):799–806.
 75. Vilá LM, Alarcón GS, McGwin G, Bastian HM, Fessler BJ, Reveille JD, et al. Systemic lupus erythematosus in a multiethnic cohort (LUMINA): XXIX. Elevation of erythrocyte sedimentation rate is associated with disease activity and damage accrual. *J Rheumatol*. 2005 Nov;32(11):2150–5.
 76. Uribe AG, Vilá LM, McGwin G, Sanchez ML, Reveille JD, Alarcón 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 Oct;31(10):1934–40.
 77. Fortin PR, Abrahamowicz M, Danoff D. Small changes in outpatients lupus activity are better detected by clinical instruments than by laboratory tests. *J Rheumatol*. 1995 Nov;22(11):2078–83.
 78. Touma Z, Gladman DD, Ibañez D, Urowitz MB. Development and initial validation of the systemic lupus erythematosus disease activity index 2000 responder index 50. *J Rheumatol*. 2011 Feb;38(2):275–84.
 79. 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 Aug;21(8):1468–71.

80. Touma Z, Urowitz MB, Taghavi-Zadeh S, Ibañez D, Gladman DD. Systemic lupus erythematosus disease activity Index 2000 Responder Index 50: sensitivity to response at 6 and 12 months. *Rheumatol Oxf Engl*. 2012 Oct;51(10):1814–9.
81. Lai J-S, Beaumont JL, Ogale S, Brunetta P, Cella D. Validation of the functional assessment of chronic illness therapy-fatigue scale in patients with moderately to severely active systemic lupus erythematosus, participating in a clinical trial. *J Rheumatol*. 2011 Apr;38(4):672–9.
82. Ward MM, Marx AS, Barry NN. Psychological distress and changes in the activity of systemic lupus erythematosus. *Rheumatol Oxf Engl*. 2002 Feb;41(2):184–8.
83. Ward MM, Marx AS, Barry NN. The rating scale preference measure as an evaluative measure in systemic lupus erythematosus. *Lupus*. 2000;9(9):696–701.
84. 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 Mar;27(3):664–70.
85. van Vollenhoven RF, Morabito LM, Engleman EG, McGuire JL. Treatment of systemic lupus erythematosus with dehydroepiandrosterone: 50 patients treated up to 12 months. *J Rheumatol*. 1998 Feb;25(2):285–9.
86. Enocsson H, Wetterö J, Skogh T, Sjöwall C. Soluble urokinase plasminogen activator receptor levels reflect organ damage in systemic lupus erythematosus. *Transl Res J Lab Clin Med*. 2013 Nov;162(5):287–96.
87. Strand V, Petri M, Kalunian K, Gordon C, Wallace DJ, Hobbs K, et al. Epratuzumab for patients with moderate to severe flaring SLE: health-related quality of life outcomes and corticosteroid use in the randomized controlled ALLEVIATE trials and extension study SL0006. *Rheumatol Oxf Engl*. 2014 Mar;53(3):502–11.
88. Ruiz-Irastorza G, Khamashta MA, Gordon C, Lockshin MD, Johns KR, Sammaritano L, et al. Measuring systemic lupus erythematosus activity during pregnancy: validation of the lupus activity index in pregnancy scale. *Arthritis Rheum*. 2004 Feb 15;51(1):78–82.
89. Nehring J, Schirmbeck LA, Friebus-Kardash J, Dubler D, Huynh-Do U, Chizzolini C, et al. Autoantibodies Against Albumin in Patients With Systemic Lupus Erythematosus. *Front Immunol*. 2018;9:2090.
90. Park D-J, Kang J-H, Lee K-E, Kang SW, Kwok S-K, Kim S-K, et al. Association of depression with socioeconomic status, anticardiolipin antibodies, and organ damage in patients with systemic lupus erythematosus: results from the KORNET registry. *Clin Exp Rheumatol*. 2018 Aug;36(4):627–35.
91. Arriens C, Hynan LS, Lerman RH, Karp DR, Mohan C. Placebo-controlled randomized clinical trial of fish oil's impact on fatigue, quality of life, and disease activity in Systemic Lupus Erythematosus. *Nutr J*. 2015 Aug 18;14:82.
92. Jolly M, Francis S, Aggarwal R, Mikolaitis RA, Niewold TB, Chubinskaya S, et al. Serum free light chains, interferon-alpha, and interleukins in systemic lupus erythematosus. *Lupus*. 2014 Aug;23(9):881–8.
93. Mok CC, Birmingham DJ, Leung HW, Hebert LA, Song H, Rovin BH. Vitamin D levels in

- Chinese patients with systemic lupus erythematosus: relationship with disease activity, vascular risk factors and atherosclerosis. *Rheumatol Oxf Engl*. 2012 Apr;51(4):644–52.
94. 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 May;38(5):868–73.
 95. Gordon C, Wallace DJ, Shinada S, Kalunian KC, Forbess L, Braunstein GD, et al. Testosterone patches in the management of patients with mild/moderate systemic lupus erythematosus. *Rheumatol Oxf Engl*. 2008 Mar;47(3):334–8.
 96. Kaya A, Goker B, Cura ES, Tezcan ME, Tufan A, Mercan R, et al. Turkish lupusPRO: cross-cultural validation study for lupus. *Clin Rheumatol*. 2014 Aug;33(8):1079–84.
 97. Navarra SV, Tanangunan RMDV, Mikolaitis-Preuss RA, Kosinski M, Block JA, Jolly M. Cross-cultural validation of a disease-specific patient-reported outcome measure for lupus in Philippines. *Lupus*. 2013 Mar;22(3):262–7.
 98. Navarra SV, Guzmán RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet Lond Engl*. 2011 Feb 26;377(9767):721–31.
 99. Manzi S, Rairie JE, Carpenter AB, Kelly RH, Jagarlapudi SP, Sereika SM, et al. Sensitivity and specificity of plasma and urine complement split products as indicators of lupus disease activity. *Arthritis Rheum*. 1996 Jul;39(7):1178–88.
 100. Sullivan KE, Suriano A, Dietzmann K, Lin J, Goldman D, Petri MA. The TNFalpha locus is altered in monocytes from patients with systemic lupus erythematosus. *Clin Immunol Orlando Fla*. 2007 Apr;123(1):74–81.
 101. Eudy AM, Siega-Riz AM, Engel SM, Franceschini N, Howard AG, Clowse MEB, et al. Effect of pregnancy on disease flares in patients with systemic lupus erythematosus. *Ann Rheum Dis*. 2018;77(6):855–60.
 102. Petri M, Howard D, Repke J. Frequency of lupus flare in pregnancy: The hopkins lupus pregnancy center experience. *Arthritis Rheum*. 1991;34(12):1538–45.
 103. Petri M, Perez-Gutthann S, Longenecker JC, Hochberg M. Morbidity of systemic lupus erythematosus: role of race and socioeconomic status. *Am J Med*. 1991 Oct;91(4):345–53.
 104. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. *Lupus*. 1999 Oct 1;8(8):685–91.
 105. Wallace DJ, Kalunian K, Petri MA, Strand V, Houssiau FA, Pike M, et al. Efficacy and safety of epratuzumab in patients with moderate/severe active systemic lupus erythematosus: results from EMBLEM, a phase IIb, randomised, double-blind, placebo-controlled, multicentre study. *Ann Rheum Dis*. 2014 Jan;73(1):183–90.
 106. Prinsen C a. C, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2018;27(5):1147–57.
 107. COSMIN terminology from a European Rheumatology perspective: a glossary for the

EULAR PRO Tool Box Initiative. Glossary [Internet]. [cited 2020 Jan 31]. Available from: <http://oiml.eular.org/glossary.cfm>

108. Drost EA. Validity and reliability in social science research. *Educ Res Perspect.* 2011;38(1):105.
109. Copay AG, Subach RB, Glassman DS et al. Understanding the minimum clinically important difference: a review of concepts and methods. *The Spine Journal* 7 (2007) 541–546
110. Middel B, Van Sonderen E. Statistical Significant Change Versus Relevant or Important Change in (Quasi) Experimental Design: Some Conceptual and Methodological Problems in Estimating Magnitude of Intervention-Related Change in Health Services Research. *Int J Integr Care.* 2002;2:e15. Epub 2002 Dec 17.
111. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol.* 1993 Mar;20(3):557–60.
112. Corzillius. M, Fortin P and Stucki G. Responsiveness and sensitivity to change of SLE disease activity measures. *Lupus* (1999) 8, 655±659

Table 1: Different definitions of disease activity according to PGA instrument.

Disease activity status	PGA values (scale)	PGA Definition	Study (reference)
Remission	0-3 VAS	0.5	(29, 37)
		0	(36, 72)
Low disease activity	0-3 VAS	≤1 (LLDAS)	(12-13, 29, 34, 37, 39-40)
		<1.5	(63)
Inactive disease	0-3 VAS	0	(3, 24, 26-27, 38, 41-42, 51, 54, 60, 62, 65, 87, 99)
		0	(72-75, 83)
		0	(80, 82, 94)
		0	(88)
Mild activity	0-3 VAS	1	(24, 26-27, 41-42, 60, 62, 65, 99)
		<1.0	(2)
		1	(88)
Moderate activity	0-3 VAS	2	(24, 26-27, 42, 60, 62, 65, 99)
		1.1 – 2.0	(2)
		2 – 2.5	(3)
Severe activity	0-3 VAS	3	(3, 24, 26-27, 41-42, 60, 62, 65, 99)
		2.1 – 3.0	(2)
		2	(88)
Flare	0-3 VAS	Increase of ≥1.0	
		- Last 93 days	(9)
Mild-moderate flare (SFI)	0-3 VAS	- From last visit	(10)
Severe flare (SFI)	0-3 VAS	Increase to >2.5	(10)
Clinical worsening	0-3 VAS	Increase of ≥0.3	(3)

In the second column the definitions were reported according to the VAS used in the study. VAS: Visual Analogue Scale. SFI: Selena Flare Index.

Table 2: criterion validity data reporting correlation coefficients between PGA and quality of life measures, laboratory markers and miscellaneous.

Table 2	Measure correlated with PGA	Study	Effect size	P
Quality of life measurements	Lupus Impact Tracker	(41)	NA	p<0.05
	SF-36	(38)		
	Physical function		r=-0.17	P<0.01
	Role physical		r=-0.23	P<0.001
	Bodily pain		r=-0.32	p<0.0001
	General Health		r=-0.08	NS
	Role emotional		r=-0.19	p<0.001
	Mental health		r=-0.16	p<0.05
	Vitality		r=-0.13	p<0.05
	Social Function		r=-0.13	p<0.05
	Physical Component Summary		r=-0.21	p<0.001
	Mental Component Summary		r=-0.14	p<0.05
	Simple Index	(31)	r=0.54	p<0.0001
	Smiley (Parent)	(62)	r=0.03	p=0.005
	HRQoL	(53)	r=-0.13	P=0.007
	Lupus Pro Symptoms	(53)	r=-0.14	P=0.0003
		(96)	r=-0.35	P=0.001
	Lupus Pro Medications	(97)	r=-0.39	P=0.0001
		(53)	r=-0.11	P=0.01
		(96)	r=-0.28	P=0.005
Depressive symptoms	(61)	NA	P<0.01	
CES-D	(82)	r=0.03	NS	
Modified CES-D	(82)	r=0.02	NS	
State Anxiety	(82)	r=0.04	NS	
FACIT-fatigue scale	(81)	r _{baseline} =-0.09	NA	
	(81)	r _{follow-up} =-0.29	NA	
LFA-REAL	(36)	r=0.79-0.81	P<0.001	
SLAQ flare	(25)	r=0.56	NA	
Biomarkers	anti-dsDNA	(64)	(93)(93)r=0.41	p<0.0001

	(92)	r=0.36	p<0.001
	(75)	r _{baseline} =0.680	p=0.001
	(75)	r _{all visits} =0.346	p=0.006
	(63)	NA	p=0.87
C3	(63)	NA	P=0.23
C4	(63)	NA	P=0.29
anti-C1q	(63)	NA	P=0.09
	(64)	R=0.43	P<0.001
Anti-Chromatin	(63)	NA	P=0.24
Anti-Ribosomal P	(63)	NA	P=0.12
MCP-1	(63)	NA	P=1.0
ICAM	(63)	NA	P=1.0
VCAM	(63)	NA	P=0.13
ESR	(54)	r=0.34	p<0.001
25(OH)D3	(93)	r=-0.20	p=0.003
anti-nucleosome	(64)	r= 0.36	p<0.001
lymphopenia <500	(74)	NA	p<0.0001
IL-6	(92)	r=0.25	p=0.03
presence of free light chains	(92)	r=0.46	p=0.001
suPAR	(86)	NA	p=0.2
Miscellaneous			
NC coronary plaques	(66)	NA	p=0.05

ICAM: intercellular adhesion molecule; VCAM: vascular cell adhesion molecule; MCP-1: monocyte chemoattractant protein-1, ESR: erythrocyte sedimentation rate; IL-6: Interleukin-6; NC: non calcified; suPAR: soluble urokinase plasminogen activator receptor; LFA-REAL: Lupus Foundation of America-Rapid Evaluation of Activity in Lupus; CES-D: Centers for Epidemiological Studies—Depression scale ; SLAQ: Systemic Lupus Erythematosus Activity Questionnaire; NA: Not acquired; NS: not significant.

Fig.1: Flow chart illustrating the literature search and study selection

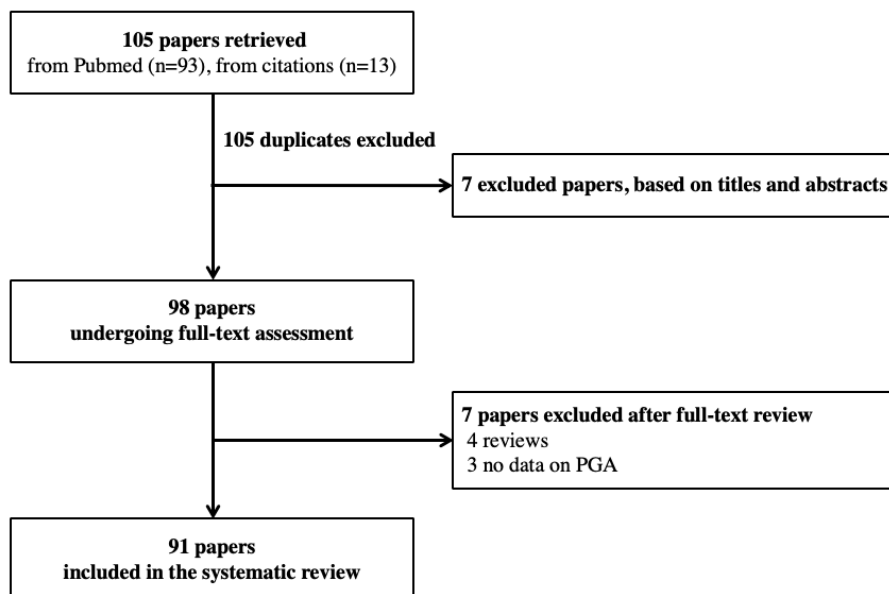
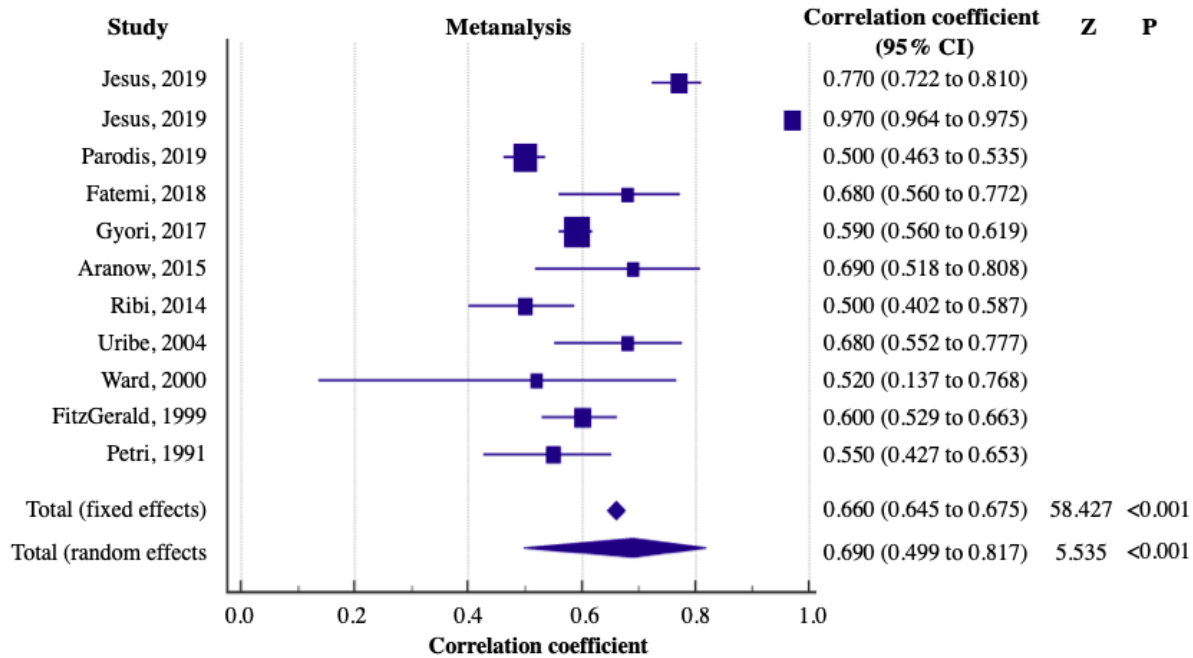


Fig. 2: Meta-analysis of the studies reporting data concerning the construct validity between PGA and SLEDAI. A good correlation was considered for a value higher than 0.60. The pooled correlation coefficients (95% CI) is given both for the fixed effects model and the random effects model. Random effects model gives a more conservative estimate considering the heterogeneity.



Test for heterogeneity

I² (inconsistency): 99.04%
 95% CI for I²: 98.80 to 99.23
 Significance level: p < 0.0001