Running Head: Damage Measurement in Childhood-Onset Lupus

Title: Measuring Disease Damage and its Severity in Childhood-Onset Systemic Lupus Erythematosus

Authors: Michael J. Holland MD[1]; Michael W. Beresford* MBChB, MRCP (UK), MRCPCH, PhD[2]; Brian M. Feldman MD, MSc, FRCPC[3]; Jennifer Huggins MD[1]; Ximena Norambuena MD[5]; Clovis A. Silva MD, PhD[6]; Gordana Susic MD [7]; Flavio Sztajnbok MD, MSc[8]; Yosef Uziel MD[9]; Simone Appenzeller MD, PhD[10]; Stacy P. Ardoin MD, MSc[11]; Tadej Avcin MD, PhD[12]; Francisco Flores, MD[13]; Beatrice Goilav MD[14]; Raju Khubchandani MD[15]; Marissa Klein-Gitelman MD, MPH[16]; Deborah Levy MD, MSc[3]; Angelo Ravelli MD[17]; Scott E. Wenderfer MD[18]; Jun Ying PhD [4]; Nicolino Ruperto MD, MPH[19]; Hermine I. Brunner MD, MSc, MBA[1] for the PRINTO and PRCSG Investigators.

*On behalf of the UK JSLE Study Group investigators.

Affiliations:

1: Cincinnati Children's Hospital Medical Center, Division of Rheumatology, Cincinnati, Ohio, USA; 2: Alder Hey Children's NHS Foundation Trust, Department of Paediatric Rheumatology Liverpool, UK and University of Liverpool, Institute of Translational Medicine, Liverpool, UK; 3: The Hospital for Sick Children, Division of Rheumatology, University of Toronto, Toronto, Canada; 4: University of Cincinnati College of Medicine, Cincinnati, Ohio, USA; 5: Hospital Dr. Exequiel Gonzalez Cortes, Paediatric Rheumatology, Santiago, Chile; 6: Pediatric Rheumatology Unit, Childrens' Institute, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, SP, Brazil; 7: Institute of Rheumatology, Belgrade, Division of Pediatric Rheumatology, Belgrade, Serbia; 8: Hospital Universitario Pedro Ernesto, Nucleo de Estudos da Saúde do Adolescente, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil; 9: Meir Medical Centre, Pediatric Rheumatology Unit, Department of Pediatrics, Kfar Saba and Sackler School of Medicine, Tel Aviv University, Israel; 10: University of Campinas Faculty of Medical Science, Rheumatology Unit, Campinas, Brazil; 11: Nationwide Children's Hospital, Section of Pediatric Rheumatology, Ohio State University Wexner College of Medicine, Columbus, Ohio, USA; 12: Ljubljana University Medical Center, Department of Pediatrics, University of Ljubljana Medical Faculty, Ljubljana, Slovenia; 13: Cincinnati Children's Hospital Medical Center, Division of Nephrology, Cincinnati, Ohio, USA; 14: The Children's Hospital at Montefiore, Division of Pediatric Nephrology, Albert Einstein College of Medicine, Bronx, New York, USA; 15: Jaslok Hospital and Research Center, Pediatric Rheumatology Clinic, Mumbai, India; 16: Lurie Children's Hospital, Division of Pediatric Rheumatology, Northwestern Feinberg School of Medicine, Chicago, Illinois, USA; 17: Istituto Giannina Gaslini, Pediatria II - Reumatologia, Genoa, Italy, and Università degli Studi di Genova, Dipartimento di Pediatria, Genoa,

Italy; 18: Baylor College of Medicine, Department of Pediatrics, and Texas Children's Hospital, Renal Section Houston, Texas, USA. 19: Istituto Giannina Gaslini, Pediatria II - Reumatologia, PRINTO, Genoa, Italy.

Grant/Financial Support:

Supported by a grant from the National Institutes of Health (5U01AR067166) and an Academic Research & Clinical award from the Cincinnati Children's Research Foundation; The Pediatric Rheumatology International Trials Organization: Supported by a grant from the European Union (contract no. QLG1-CT-2000-00514) and by IRCCS G. Gaslini, Genoa, Italy.

University of Liverpool: LUPUS UK and the National Institute of Health Research (NIHR)

Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2015/03756-4 to CAS) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 303422/2015-7 to CAS).

Corresponding Author:

Michael J. Holland.

3333 Burnet Avenue, MLC 4010, Cincinnati, Ohio, USA 45229-3026; 513-636-7982; 513-636-5990 (fax); mjholland@cmh.edu

Word Counts: Abstract: 249 words Body: 3657/3,800 words

ABSTRACT

Objectives: To describe the frequency and types of disease damage occurring with childhood-onset systemic lupus erythematosus (cSLE) as measured by the 41-item Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SDI), and to assess the SDI's ability to reflect damage severity.

Methods: Information for the SDI was prospectively collected from 1,048 cSLE patients. For a subset of 559 patients physician-rated damage severity measured by visual analog scale (MD-VAS_{damage}) was also available. Frequency of SDI-items, and the association between SDI summary-scores and MD-VAS_{damage} were estimated. Finally, an international consensus conference, utilizing nominal group technique, considered the SDI's capture of cSLE-associated damage and its severity.

Results: After a mean disease duration of 3.8 years, 44.2% (463/1048) of patients already had an SDI summary-score >0 (maximum: 14). The most common SDI items scored were proteinuria, scarring alopecia, and cognitive impairment. Although there was a moderately strong association between SDI summary-scores and MD-VAS_{damage} ($r_{Spearman}$ = 0.49; $p<0.0001$) in patients with damage (SDI summary-score >0), mixed effect analysis revealed that only four SDI items, each occurring in <2% of patients overall, were significantly associated with MD-VAS_{damage}. There was consensus among cSLE experts that the SDI in its current form is inadequate for estimating the severity of cSLE-associated damage.

Conclusion: Disease damage as measured by the SDI is common in cSLE, even with relatively short disease durations. Given the shortcomings of the SDI, there is a need to develop new tools to estimate the impact of cSLE-associated damage.

SIGNIFICANCE & INNOVATION

- In cSLE, damage is common, and occurs most frequently in the renal, cutaneous, neuropsychiatric, and musculoskeletal organ systems.
- By design, the SDI seeks to provide an enumeration of the types of damage present in cSLE, rather than directly measuring damage severity. Despite this, the SDI summary-score is often used as a stand-alone continuous outcome measure in research.
- An international group of pediatric rheumatologists and nephrologists experienced in cSLE reached unanimous consensus that a new measure or approach is needed to better capture damage severity in cSLE.

ACKNOWLEDGEMENTS

Cincinnati Children's Hospital Medical Center: Alan Watts, Kasha Wiley, Pinar Ozge Avar Aydin for data collection and cleaning. A special thank you to Susan Priest and the Pediatric Rheumatology Collaborative Study Group for their support in organizing the consensus conference. We would also like to thank the Pediatric Rheumatology International Trials Organization, and the University of Liverpool for the use of data and its collection by the UK JSLE Cohort Study and Repository, and the Clinical Research Network (CRN), including patients, and all center staff and the principal investigators of the UK JSLE Study Group.

INTRODUCTION

Disease activity describes theoretically reversible manifestations due to the inflammatory processes underlying systemic lupus erythematosus (SLE), while the term 'damage' is used to designate irreversible organ scarring or tissue degradation. However, a formal consensus definition of disease damage with SLE has not been published. Quantifying damage or measuring the severity of damage is an important consideration in gauging the overall outcome in SLE, particularly as disease-related mortality decreases (1, 2). Given the diversity of SLE-associated organ involvement, and in line with the development of disease activity indices for quantifying disease activity, a SLE damage index has been developed: the Systemic Lupus International Collaborating Clinics/ American College of Rheumatology Damage Index (SDI) (1).

Prior investigations support that the presence of SLE-associated damage that is scored by the SDI is associated with increased mortality in adults (2), and increased cumulative SLE activity in both adults and children (2-4). The developers of the SDI stressed that the index provides an enumeration of the presence or absence of damage types only, and that the SDI does not directly quantify damage severity (3). This is sensible because, for example, a patient with a small cataract and one with a debilitating stroke will both receive the same SDI summary-score of 1, provided each patient fulfills the definition of only one SDI item. Nevertheless, the SDI summary-score has been used as an independent, continuous outcome measure in statistical analyses (5).

There have been prior attempts to adapt the SDI to better measure damage severity by introducing item weightings. However, two prior attempts at item weightings using data gathered from adults with SLE did not meaningfully improve the association of SDI summary-scores with mortality, and therefore item weighting were not pursued further (2, 6). One prior

attempt was made to weight SDI items (based on SLEDAI item weights) in data from patients with childhood-onset SLE (cSLE); this approach did not improve prediction of damage using cumulative disease activity as a predictor, and was similarly not pursued (4).

In about 20% of patients with SLE the disease is diagnosed during childhood (cSLE), i.e. before 18 years of age (7). A prior international consensus process (8), focused on outcome measures in cSLE and juvenile dermatomyositis, defined a core set of variables to characterize cSLE-associated damage and its impact. Based on consensus, the agreed upon cSLE damage core set consists of the SDI as the current standard damage tool, a physician global damage assessment of damage [visual analog scale (VAS) or Likert scale], growth (height and weight), bodily development (menses, Tanner staging), and a measure of health-related quality of life.

The Outcome Measures in Rheumatology (OMERACT) collaborative has recently published a revised framework. The OMERACT Filter 2.0 builds on the OMERACT Filter 1.0, i.e. the use of outcome measures in rheumatology that are valid, discriminate conditions of interest, and are feasible (9). The OMERACT Filter 2.0 advocates development of core outcome measurement sets in rheumatology (10). There is new emphasis on the concept of multiple 'core areas' (death, disease impact, resource use, and disease manifestations) which may be addressed by these core outcome sets. Finally, the need to explicitly consider "perspective" is stressed, as well as the "context" in which an outcome measure will be used. The aspect of perspective (patient, physician, and/or society) and context appear particularly important when capturing damage severity. While the previously developed cSLE damage core set (8) touches on at least two of the OMERACT core areas for construct measurement, stand-alone use of the SDI is still commonplace when quantifying the amount of damage in clinical studies of both SLE and cSLE.

Based on the above, we sought to critically appraise the SDI when used in cSLE by: 1) delineating the frequency of SDI damage items in a large composite cohort under consideration of disease duration; 2) comparing the SDI summary-score to a physician global assessment of damage severity; and 3) to explore the possible impact of SDI item weightings to better capture damage severity as rated by the treating physician using statistical techniques. Subsequently, our findings and analyses were presented at a consensus conference held in April 2017, where an international panel of cSLE experts was asked to advise on approaches to measure damage severity in cSLE.

MATERIAL & METHODS

Patients

Longitudinal data from large, prospective cohorts of cSLE were reanalyzed. These were the United Kingdom Juvenile-onset SLE Cohort Study (UK, n= 350) (1), the cSLE cohort followed at Cincinnati Children's Hospital Medical Center (CCHMC, n= 139), and an international cohort (n= 559) assembled in Latin America, Australia, Asia, and various European countries by the Pediatric Rheumatology International Trials Organization (PRINTO) (8). The final composite study cohort included 1,048 patients. General demographic data were recorded, though ethnic/racial data were not collected for the PRINTO dataset due to legal restrictions. Approval was given by the Cincinnati Children's Hospital Medical Center Institutional Review Board for this secondary analysis. Training had been provided for the completion of the SDI and other disease measures of all cohorts considered in this study. All physicians completing instruments were experienced in the care of children with cSLE.

Damage measures and scales

Systemic Lupus International Collaborating Clinics / American College of Rheumatology Damage Index. Developed through a consensus process focused on adults with SLE, the SDI quantifies irreversible damage according to specific item definitions (1). The SDI captures damage in patients since their diagnosis with SLE, irrespective of whether damage is due to the SLE process itself, its treatment, or a co-morbid condition. Damage captured in the SDI is considered non-reversible if any given item has been present for at least 6 months continuously, or immediately for some events associated with acute organ damage, e.g. myocardial infarct (1).

The 41 items included in the SDI were selected based on experience with adult-onset SLE, with items mostly scored as being either present or absent. While there is no overall weighting system, some six SDI items (stroke, myocardial infarct, tissue loss, bowel infarct, avascular necrosis, or malignancy) can be scored twice, if two qualifying events occur at least 6 months apart from each other. The SDI item 'end stage renal disease' is always given a score of 3 when present for at least 6 months continuously. Scoring is cumulative: once an item qualifies for scoring in the SDI, that item is always scored moving forward, even if it subsequently resolves or is corrected (2). The SDI-summary-score is the simple sum of the item scores; thus a summary-score of 0 is assigned to patients who have *never* met criteria for any listed damage item.

Visual analog scale of disease damage severity. Treating physicians contributing data to the PRINTO cohort rated damage severity on a 10-cm VAS (MD-VAS$_{damage}$) for their patients. The

following anchor statements were included at each end of the 10-cm scale: 0, no damage; 10, very severe damage. This rating was done either prior to or after completing the SDI.

Statistical analyses

Descriptive analyses included frequencies for categorical variables as well as means and standard deviations for numerical variables. Comparison of numerical demographic features and SDI summary-scores between cohorts was done by a fixed effect model, with differences of post-hoc means between groups corrected for multiple comparisons using the Tukey's method. Damage item frequency was calculated based upon the SDI score at the last follow-up visit available for each patient and, per SDI instructions, included all SDI item ever scored for a given patient. Contingency table analysis compared item frequencies between groups of patients (CCHMC, UK, PRINTO). Significant differences between item frequencies were based on chi-square analyses or Fisher exact testing, where appropriate. Given variation in disease duration, for purposes of statistical comparison of item frequency between cohorts, only patient SDI scores with total disease durations of 4.5 years or less were considered. The 4.5 year cut-off reflects the 70th percentile of follow-up in the cohort with the shortest follow-up (PRINTO).

To assess the relationship of SDI summary-scores to physician-perceived severity of damage (MD-VAS$_{damage}$), we calculated the Spearman correlation coefficients from patient visits with available data. Correlation with MD-VAS$_{damage}$ was then separately assessed only for patient visits with known damage (SDI scores > 0). This was done to capture damage severity, which seems relevant only with the presence of some damage (SDI summary-score >0). To evaluate the effect of multiple comparisons, correlations were also calculated separately using only the first,

and then last visit for each patient. Finally, individual damage items (present vs. absent) were assessed for their associations with the MD-VAS_{damage} ratings using logistic regression models and the GEE method in computation. Computations were performed using SAS 9.3 (Cary, NC) or EXCEL (version 2013, Redmond, WA).

Consensus Conference

From April 23rd through 25th 2017, an international consensus conference of physicians with expertise in cSLE was held in Cincinnati, Ohio. The expert group consisted of thirteen physicians (10 pediatric rheumatologists and 3 pediatric nephrologists) with profound experience in the care of cSLE. Guided by an experienced moderator (BMF), nominal group technique was used to facilitate discussion and consensus formation. Consensus was defined a-priori as ≥75% agreement among participating experts. Prior to opening discussion, results of the analyses (see below) and a review of relevant literature were presented

RESULTS

Patients

As expected, most patients included in this study were female (82.9%), without significant gender differences between cohorts. While UK and CCHMC cohort data were collected after 2006, the PRINTO cohort data were completed by 2004. There were significant racial differences present between the UK and CCHMC cohorts. Patient age at diagnosis and mean total disease duration were significantly higher in the CCHMC cohort when compared to the other datasets

(CCHMC vs other cohorts; mean age: 13.9 years vs. 12 years; p<0.0001, mean duration: 5.15 years vs. 3.5-3.8 years; p<0.0001). Additional details are shown in **Table 1.**

Overall disease damage as measured by the SDI

Among the 1,048 patients, a total of 585 (55.8%) lacked disease damage (SDI summary-score = 0) at the time of the last follow-up, which occurred, on average, 3.8 years post diagnosis with cSLE (**Table 1**)**.** The proportion of patients without damage at last follow-up was highest in the UK cohort (77.7%). Mean SDI summary-scores significantly differed between the UK and PRINTO cohorts, and PRINTO and CCHMC cohorts, but not between the CCHMC and UK cohorts (**Table 1**). Overall, the mean SDI summary-score increased incrementally and closely related to increasing disease duration (**Figure 1, Panel 1**).

Common and less common types of disease damage with cSLE

The frequency of SDI items at the final follow-up visit in each cohort is summarized in **Table 2.** The three most commonly encountered SDI items were long-standing nephrotic-range proteinuria, scarring alopecia, and chronic cognitive impairment in all three cohorts. Irrespective of disease duration, the four most commonly damaged organ systems were the neuropsychiatric, kidney, skin, and musculoskeletal (**Figure 1, Panel 2**). There were four SDI items that were present in fewer than three patients (3/1,048 = 0.3%): angina, myocardial infarction, mesenteric insufficiency, and tendon rupture. An additional 12 SDI items were present in <1% of the study population. No patient had the SDI item pulmonary infarction scored. (**Table 2)**.

Relationship between physician-rated damage severity and SDI summary-scores

Ratings of MD-VAS$_{damage}$ were available for 1,793 of a total of 1820 patient visits with SDI scores. Of these, 1,245 (69.4%) were without damage (SDI summary-score = 0), and 548 (30.5%) with some damage (SDI summary-score > 0).

Damage-free (SDI summary-score=0) patients are expected to have a MD-VAS$_{damage}$ of 0. However, physicians considered 24.2% (301/1245) of "damage-free" patients as having some damage (MD-VAS$_{damage}$ >0). As shown in **Figure 2, Panel 1**, with SDI summary-scores of 0, MD-VAS$_{damage}$ ratings were <1 in 94% of the visits.

Figure 2, Panel 2 provides an overview of the MD-VAS$_{damage}$ and SDI summary scores throughout the range of observed values. Only 5.7% (31/548) of patients with SDI summary-scores of >0 were considered "damage-free" by their treating physician (MD-VAS$_{damage}$=0).

The correlation between SDI summary-score and MD-VAS$_{damage}$ overall was strong ($r_{spearman}$ 0.71; p<0.0001)**,** though when narrowed to include only visits with some damage (SDI >0) the correlation was only moderate ($r_{spearman}$ 0.496; p<0.0001). Correlations were similar when considering only the first, or last visit for each patient, ranging from 0.66-0.72 for all SDI scores, versus 0.45-0.54 for only those visits with SDI summary scores >0.

In exploratory analysis using mixed effect modeling, we aimed at identifying SDI-items that importantly influence physician-rated damage severity. This analysis revealed that only four of the 41 SDI items were significantly associated with the MD-VAS$_{damage}$. They were the SDI items pulmonary fibrosis, shrinking lung syndrome, chronic pericarditis, and extensive cutaneous scar. Of note, all of these SDI-items had an overall frequency of <2% in the composite study cohort.

Consensus Conference

Neither the SDI nor the proposed pediatric adaptation of the SDI (11) were considered adequate to measure damage severity in cSLE. There was consensus (100%) that a separate measure or approach to capture damage severity of cSLE was needed, and that the OMERACT Filter 2.0 framework should be used for its development (83% agreement). Consensus was reached around a definition of cSLE-associated damage and damage severity as a first step toward improving measurement of damage-related constructs.

Damage associated with cSLE was defined as "*Impairment of anatomy or physiology that may be associated with scarring, may accumulate, and is not completely reversible. Damage may be caused by disease, adverse effects of medication, or associated comorbidity. In children this may lead to stunted cognitive, and physical development*" (83% consensus).

Based on consensus (77% agreement) damage severity was defined as follows: "*Severity of damage is measured by the organs involved, and the extent of anatomical and physiological derangement as judged by the expected impact on mortality, degree of support required, activity limitation, restriction in social participation, and patient-centered quality of life*."

DISCUSSION

We examined the patterns and severity of disease damage in a large composite cohort of cSLE patients with the goal to appraise the ability of the SDI to capture disease damage and its severity. Based on detailed statistical analysis and expert consensus, the SDI was considered in inadequate for quantifying the impact of disease damage in cSLE.

Although cSLE damage patterns have previously been described in detail (4, 7, 12), our analyses seem unique given the number of cSLE patients included, and the analysis of physician global assessments of damage severity. Damage patterns were largely similar across disease cohorts, with SDI-items in the neuropsychiatric, renal, musculoskeletal and skin systems most frequently encountered. Dissimilarities in damage patterns were identified for the frequency of chronic muscle atrophy or muscle weakness across cohorts. Reasons for these differences are unknown but could include differences in access to and use of medications, as well as differences in the relevant healthcare systems and demographics. Unfortunately, insufficient longitudinal data were available to identify the driving factors for the observed differences. Only two patients in our cohort encountered malignancies (type unknown), and this observation is in line with the malignancy risk based on an earlier epidemiological study in cSLE (13), supporting the representativeness of this composite cohort. Further, the rarity of most SDI items in our cohort in line with prior studies in adults (14).

Both SDI organ domain scores and total SDI summary-scores increased with disease duration. Indeed, SDI summary-scores were closely correlated with disease duration. This supports our prior research where disease duration and cumulative disease activity were closely correlated with each other, but the cumulative burden of disease activity with cSLE was the better predictor of SDI summary-scores in a smaller cSLE cohort (4). A close relationship between damage accrual as measured by the SDI and disease duration supports the construct validity of the SDI for measuring damage.

The SDI-summary-score is the simple sum of its mostly un-weighted SDI-item scores. Thus one might expect that a higher score corresponds to more severe damage, especially because

there are seven “high-impact” items that receive an additional score (myocardial infarction, stroke, end-stage renal disease, avascular necrosis, malignancy, significant tissue loss, and gastrointestinal infarction/resection) if occurring repeatedly. While true item weightings have been incorporated in other scales used for cSLE and SLE (4, 15), our findings imply that more sophisticated item weightings will not improve the ability of the SDI to capture cSLE-associated damage severity. This is because there were only statistically significant associations between the MD-VAS Damage summary score and a few, rarely-endorsed, SDI items. This observation is similar to that of earlier studies using other statistical approaches, which also found little value in item weightings to improve the SDI’s ability to reflect damage severity, or mortality (2, 4, 6).

Although disease damage accumulated with cSLE was substantial, many of the SDI-items were rarely scored. Indeed, nearly half of SDI items were encountered in <1% of cSLE patients. Although it might be tempting to eliminate items with extremely low prevalence to improve the feasibility of the scale when used in cSLE, we do not think this is advisable as item reduction will not improve the construct validity of the SDI, nor its ability to capture damage severity. Further, SLE and cSLE are highly heterogeneous in their phenotypes and certain SDI items, such as myocardial infarctions, are known to occur more commonly with higher age and longer disease durations than those captured by our study. There is also broad agreement that the inclusion of rarely scored items is warranted when estimating damage with other multi-system diseases such as inflammatory immune-mediated myositis and vasculitis, respectively (16, 17).

If the SDI was a good measure of cSLE-associated damage severity, children with SDI summary-scores of 0 (best possible value) should consistently have an MD-VAS$_{damage}$ rating of 0 (best possible value). However, over 20% of children without cSLE-associated damage received a

MD-VAS$_{damage}$ >0. While the exact reasons for this observation are unknown, this could simply be a reflection of so-called “end-scale aversion”: it has been recognized that raters tend to fail to provide the best or worse possible rating on a VAS (18, 19). Actually, almost all MD-VAS$_{damage}$ rating with SDI summary-scores of 0 were in the range of 0 to <1.

Another explanation for MD-VAS$_{damage}$ ratings exceeding 0 in patients without damage may be that certain pediatric specific damage is not considered in the SDI currently. Indeed, there is consensus among pediatric rheumatologists that pubertal development and growth are important aspects that deserve consideration when measuring cSLE-associated damage (8). This is reflected in the proposal of a pediatric version of the SDI (pSDI) where two items are added to the traditional SDI (11), namely reduced growth and delayed development. A problem with these cSLE-specific items might be that both items can resolve, which could be perceived to violate the concept of irreversibility of disease damage as defined by the SDI. Pediatric experts involved in the delineation of the cSLE damage core set considered this as acceptable, given the more pronounced tissue regeneration in children (8). An example to be considered is the resolution of some bone erosions in pediatrics.

The concept of reversibility of some damage is now clearly stated in the consensus definition of cSLE-associated damage. Notably, some of the current SDI items can also “resolve” clinically, such as nephrotic range proteinuria or seizures, but the resultant SDI item scores are maintained (2). Whether it would be more sensible to reduce the summary-score when growth and development delays have resolved or to continue considering them in a damage summary-score will need further evaluation, should the validation of the pSDI be pursued.

For the SDI to be considered a robust measure of cSLE-associated damage severity, one would expect a strong correlation between the SDI summary-score and the MD-VAS_{damage} rating. However, our data suggest that this is not the case, which is in line with our prior research (6). Indeed, the lack of a strong association of damage severity with commonly encountered SDI-items likely reflects the SDI originators' caution that the index provides only an enumeration of damage items, rather than quantifying severity (1). Hence, one might suggest the use of the term "summary-count," rather than summary-score, to emphasize the distinction in future research. This may imply that it is inadvisable to present means and standard deviations of the SDI summary-score, or to use the SDI summary-score as a continuous variable in statistical models.

The limitation of the SDI to measure damage severity is not surprising because SDI items vary considerably in their impact on patient function, and the need for medical interventions. For example, a stroke resulting in hemiparesis will likely have greater impact on patient function than a cataract, despite the equal summary-score contribution of each item. Indeed, physician perception of damage severity is likely influenced by both the type of damage, and the resultant prognosis and/or health care utilization. This might indicate that there is a need for Likert scaling of SDI items to better capture the extent or impact of damage encountered. For example, a stroke that does not result in long-term clinical deficits would receive a lower item score than a stroke causing mild localized paresis, which in turn would be scored less than a stroke that renders the patient unable to walk or to speak.

Perspective is also a key factor when estimating the severity of cSLE damage. This can be exemplified by scarring alopecia, which has an extremely limited impact on mortality, but could

be devastating from a patient's viewpoint. These considerations are further complicated by the known resilience of children (11).

The short-comings of the SDI were highlighted by deliberations of cSLE experts during the recent consensus conference. In accordance with the OMERACT Filter 2.0 (10), the importance of perspective when measuring the severity or impact of cSLE-associated damage is reflected in the newly developed provisional consensus definition of damage severity: "*as judged by the expected impact on mortality, degree of support required, activity limitation, restriction in social participation, and patient-centered quality of life*."

Limitations of our study include the lack of physician damage severity ratings from all patients included in the composite cohort. Use of physician global assessments may be considered problematic in itself (20), with potential inter-rater and intra-rater variation, which we were unable to assess. Nonetheless, over 1,200 MD-VAS_{damage} ratings were available for analysis, and VAS are widely used in medical research in general and rheumatology in particular.

Taken together, the analysis of a large cSLE data set confirms the construct validity of the SDI as a valid measure to "count' damage events in children and adolescents with cSLE. The SDI in its current form is ill suited to accurately quantify the severity of cSLE-associated damage, even if item-weightings were revisited in an effort to improve the algorithm used to calculate the SDI summary-score. We propose an international effort, supported by the relevant professional organizations, to capture the severity of impact of disease damage of patients with SLE from childhood through adulthood.

LITERATURE CITED

1. Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Sanchez-Guerrero J, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. Arthritis & Rheumatology. 1996;39:363-9.

2. Gladman D, Urowitz MB**.** The SLICC/ACR damage index: progress report and experience in the field. Lupus. 1999;8:632-7.

3. Becker-Merok A, Nossent HC**.** Damage accumulation in systemic lupus erythematosus and its relation to disease activity and mortality. The Journal of rheumatology. 2006;33:1570-7.

4. Brunner HI, Silverman ED, To T, Bombardier C, Feldman BM**.** Risk factors for damage in childhood-onset systemic lupus erythematosus: Cumulative disease activity and medication use predict disease damage. Arthritis & Rheumatology. 2002;46:436-44.

5. Livingston B, Bonner A, Pope J**.** Differences in autoantibody profiles and disease activity and damage scores between childhood-and adult-onset systemic lupus erythematosus: a meta-analysis. Seminars in arthritis and rheumatism, 2012. Elsevier: 271-80.

6. Brunner HI, Feldman BM, Urowitz MB, Gladman DD**.** Item weightings for the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Disease Damage Index using Rasch analysis do not lead to an important improvement. The Journal of rheumatology. 2003;30:292-7.

7. Watson L, Leone V, Pilkington C, Tullus K, Rangaraj S, McDonagh JE, et al. Disease activity, severity, and damage in the UK juvenile-onset systemic lupus erythematosus cohort. Arthritis & Rheumatology. 2012;64:2356-65.

8. Ruperto N, Ravelli A, Murray KJ, Lovell DJ, Andersson-Gare B, Feldman BM, et al. Preliminary core sets of measures for disease activity and damage assessment in

juvenile systemic lupus erythematosus and juvenile dermatomyositis. Rheumatology. 2003;42:1452-9.

9. Boers M, Brooks P, Strand CV, Tugwell P**.** The OMERACT filter for Outcome Measures in Rheumatology. J Rheumatol. 1998;25:198-9.

10. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino M-A, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. Journal of clinical epidemiology. 2014;67:745-53.

11. Gutiérrez-Suárez R, Ruperto N, Gastaldi R, Pistorio A, Felici E, Burgos-Vargas R, et al. A proposal for a pediatric version of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index based on the analysis of 1,015 patients with juvenile-onset systemic lupus erythematosus. Arthritis & Rheumatology. 2006;54:2989-96.

12. Ravelli A, Duarte-Salazar C, Buratti S, Reiff A, Bernstein B, Maldonado-Velazquez MR, et al. Assessment of damage in juvenile-onset systemic lupus erythematosus: A multicenter cohort study. Arthritis Care & Research. 2003;49:501-7.

13. Bernatsky S, Clarke AE, Labrecque J, von Scheven E, Schanberg LE, Silverman ED, et al. Cancer risk in childhood-onset systemic lupus. Arthritis research & therapy. 2013;15:R198.

14. Sung Y-K, Hur NW, Sinskey JL, Park D, Bae S-C**.** Assessment of damage in Korean patients with systemic lupus erythematosus. The Journal of rheumatology. 2007;34:987-91.

15. Gladman DD, Ibañez D, Urowitz MB**.** Systemic lupus erythematosus disease activity index 2000. The Journal of rheumatology. 2002;29:288-91.

16. Sultan SM, Allen E, Cooper RG, Agarwal S, Kiely P, Oddis CV, et al. Interrater reliability and aspects of validity of the myositis damage index. Ann Rheum Dis. 2011;70:1272-6.

17. Suppiah R, Flossman O, Mukhtyar C, Alberici F, Baslund B, Brown D, et al. Measurement of damage in systemic vasculitis: a comparison of the Vasculitis Damage Index with the Combined Damage Assessment Index. Ann Rheum Dis. 2011;70:80-5.

18. Bleichrodt H, Johannesson M**.** An experimental test of a theoretical foundation for rating-scale valuations. Medical Decision Making. 1997;17:208-16.

19. Torrance GW, Feeny D, Furlong W**.** Visual analog scales: do they have a role in the measurement of preferences for health states? : Sage Publications Sage CA: Thousand Oaks, CA; 2001.

20. Taylor J, Giannini EH, Lovell DJ, Huang B, Morgan E**.** Lack of Concordance in Inter-Rater Scoring of the Provider's Global Assessment of Children with Juvenile Idiopathic Arthritis with Low Disease Activity. Arthritis care & research. 2017.

FIGURE LEGENDS:

Figure 1, Panel 1

The relation of SDI summary-scores and disease duration in the cohort (n=1048) is shown for patients up to 10 years since the diagnosis with cSLE. Analysis of variance supports a strong linear relationship (R-square = 0.9; p-value <0.0001)

Figure 1, Panel 2

The relationship of presence of damage in the nine organ domains of the SDI is shown for disease durations up to 6 years. Diabetes and malignancies were very rare and are omitted from depiction. Lines are moving averages of annual values. The most commonly damaged organ systems were the neuropsychiatric and the musculoskeletal, renal and skin

Figure 2, Panel 1

The relationship between physician-rated severity of damage (MD-VAS_{Damage}) and the SDI summary-score in patients with cSLE are depicted, with focus on MD-VAS_{Damage} ratings of 0, 1 or 2. Despite SDI scores of 1 or 2, some physicians provided damage severity ratings of 0. Conversely, only 76% of patients with a SDI summary-scores of 0 received the best possible MD-VAS_{Damage} rating.

Figure 2, Panel 2

The relationship between physician-rated severity of damage (MD-VAS_{Damage}) and the SDI summary-score in patients with cSLE are depicted. With higher SDI summary-scores more commonly higher MD-VAS_{Damage} ratings are assigned.

Table 1: Demographics and Summary Score by Cohort

Cohort:	*All Cohorts*	*UK (1)*	*Cincinnati (2)*	*PRINTO (3)*	*P-value (1) vs. (2)*	*P-value (1) vs. (3)*	*P-value (2) vs. (3)*
Total N:	1048	350	139	559		N/A	
Female (%)‡	869 (82.9%)	291 (83.1%)	118 (84.9%)	460 (82.3%)			
Age at Diagnosis†	12.2 (3.1)	12 (3.32)	13.9 (2.82)	12 (2.95)	<0.0001	0.86	<0.0001
Disease Duration†	3.81 (2.98)	3.82 (3.17)	5.15 (3.75)	3.46 (2.57)	<0.0001	0.08	<0.0001
SDI Summary-score†	0.94 (1.56)	0.56 (1.06)	0.99 (1.59)	1.16 (1.76)	0.0006	0.0001	0.30
Patient N (%) with SDI summary-score = 0‡	585 (55.8%)	272 (77.7%)	80 (58.5%)	233 (41.7%)		<0.0001	

† Listed values are Mean (Standard Deviation). P-values are from post-hoc analysis using fixed effect models and correcting for the Tukey's method and only listed if p-value are<0.1. N/A: Not applicable. ‡ Listed values are Frequency (in %); P-values are from Chi square tests and only listed if p-values are <0.1

Table 2. Comparison of SDI Item Frequency by Cohort**

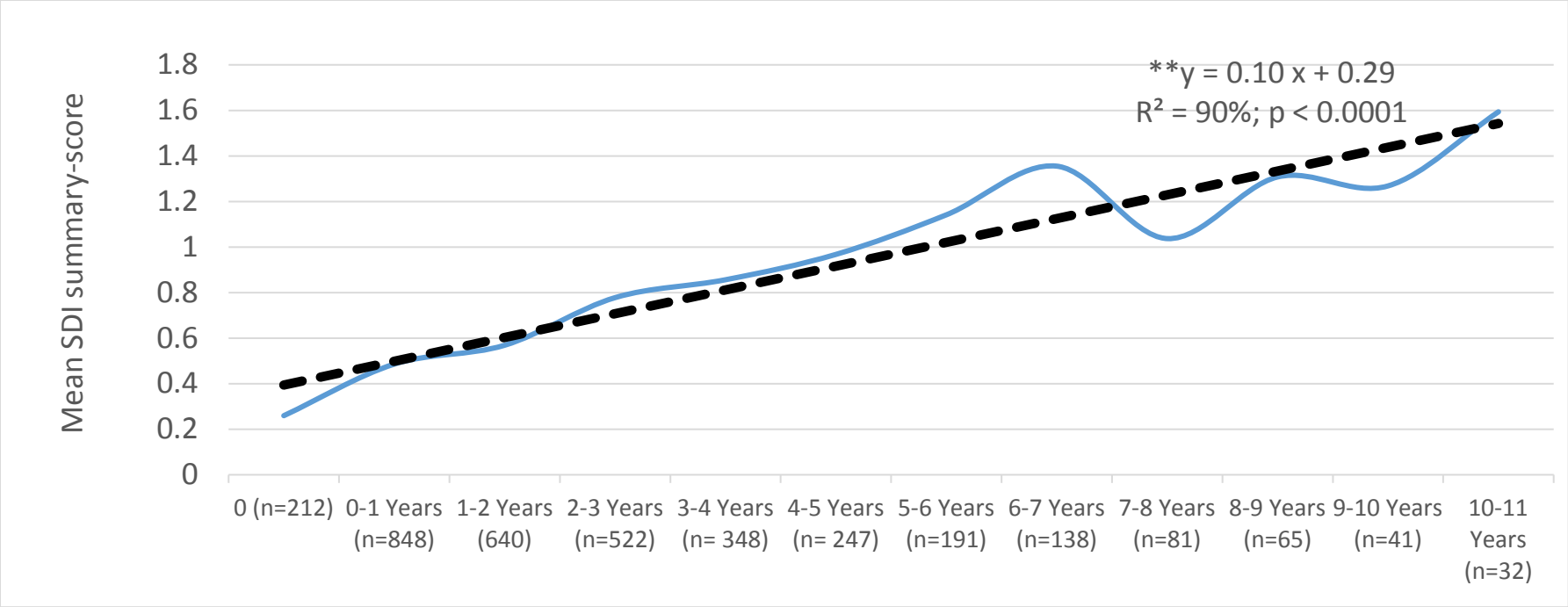
	Overall	*UK*	*CCHMC*	*PRINTO*	*P-value*
Ocular					
Cataracts	4.2	2	6.5	5	0.38
Retinal change	1.7	1.1	0.7	2.3	0.24
Neuropsychiatric					
Cognitive impairment	8.4	3.4	12.2	10.7	0.0004*
Seizures requiring therapy	3.9	2.3	2.2	5.4	0.09*
Cerebrovascular accident	2.9/ 0.4	2 / -	3.6/ 1.4	3.4/0.4	0.014†
Cranial/peripheral neuropathy	2.4	2.3	0.7	2.9	0.33†
Transverse myelitis	0.4	0.3	0.7	0.4	0.29†
Renal					
Estimated GFR‡ < 50%	3.6	1.1	4.3	5	0.034*
Proteinuria ≥ 3.5 gm/day	9.6	4.3	10.8	12.7	0.0038*
End-stage renal disease	1.3	0.6	3.6	1.3	0.14†
Pulmonary					
Pulmonary hypertension	0.4	0.9	0.7	0	0.41†
Pleural fibrosis	1.3	0.3	0	2.3	0.18†
Shrinking lung	0.8	0.3	0	1.3	0.82†
Pulmonary fibrosis	0.9	0.6	0.7	1.3	0.6†
Pulmonary infarction	0	0	0	0	-
Cardiovascular					
Angina/artery bypass	0.1	0	0	0.2	1†
Myocardial infarction	0.1/ 0.1	0/ -	0/ -	0.2/ 0.2	-
Cardiomyopathy	0.9	0.3	0.7	1.3	0.21†
Valvular disease	0.9	0.9	0	1.1	0.75†
Pericarditis/pericardectomy	1.6	0.9	0.7	2.3	0.91†

	Overall	*UK*	*CCHMC*	*PRINTO*	*P-value*
Peripheral vascular					
Claudication for 6 months	0.3	0	0	0.5	0.67†
Minor tissue loss	2.4/ -	0.9/ -	1.4/ -	3.6/-	0.28†
Significant tissue loss ever	0.7/0.3	0.3/ -	0.7/ 0.7	0.9/0.4	0.11†
Venous thrombus swelling, ulcer/stasis	2.5	1.7	0.7	3.4	0.81†
Gastro Intestinal -					
Bowel infarction or resection	1.3/ 0.3	1.1	1.4/ 0.7	1.4/0.4	0.88†
Mesenteric insufficiency	0.1	0.3	0	0	0.63†
Peritonitis	0.2	0	0	0.4	0.63†
Stricture/upper GI tract surgery ever	0.4	0.6	0	0.4	1†
Pancreatic insufficiency	0.5	0.3	0	0.7	1†
Musculoskeletal					
Muscle atrophy/weakness	7.9	4	3.6	11.4	0.0056
Deforming/erosive arthritis	4.7	2.9	3.6	6.3	0.19
Osteoporosis with fracture	2.6	1.4	2.2	3.4	0.39†
Avascular necrosis	2.7/1.0	0.6/ -	10.8/ 5	2/0.4	0.055†
Osteomyelitis	0.4	0	0.7	0.5	0.63†
Ruptured tendons	0.1	0.3	0	0	0.41†
Skin					
Scarring chronic alopecia	9.1	11.7	6.5	8.2	0.06
Extensive scar/panniculum	1.7	1.4	1.4	2	0.82†
Skin ulceration for 6 months	2.7	1.7	0	4.1	0.27
Premature gonadal failure	2	0.6	0.7	3.2	0.24†
Diabetes	0.9	1.1	1.4	0.5	1†
Malignancy	0.2	0.3	0.7	0	-

‘** Values are % of patients with SDI item score of 1 / 2, or for ESRD a score of 3; ‡glomerular filtration rate; *P-values are from logistic models after adjusting for disease duration; † p-values are from Fisher’s exact test.

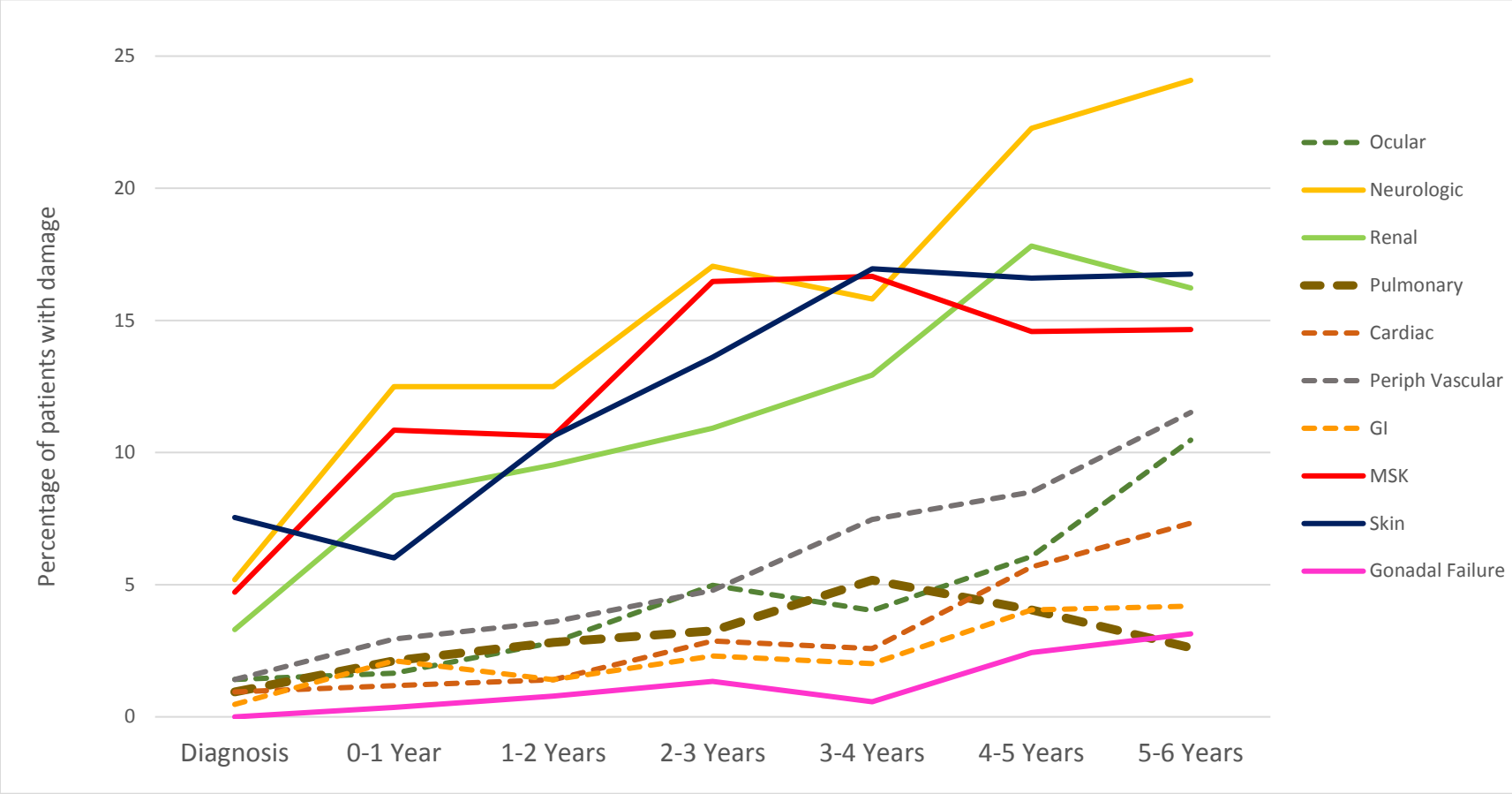
Figure 1: Damage accumulation over time

Panel 1: Observed SDI summary-scores by disease duration for up to 10 years



** Function of linear trendline including information about model fit (R^2)

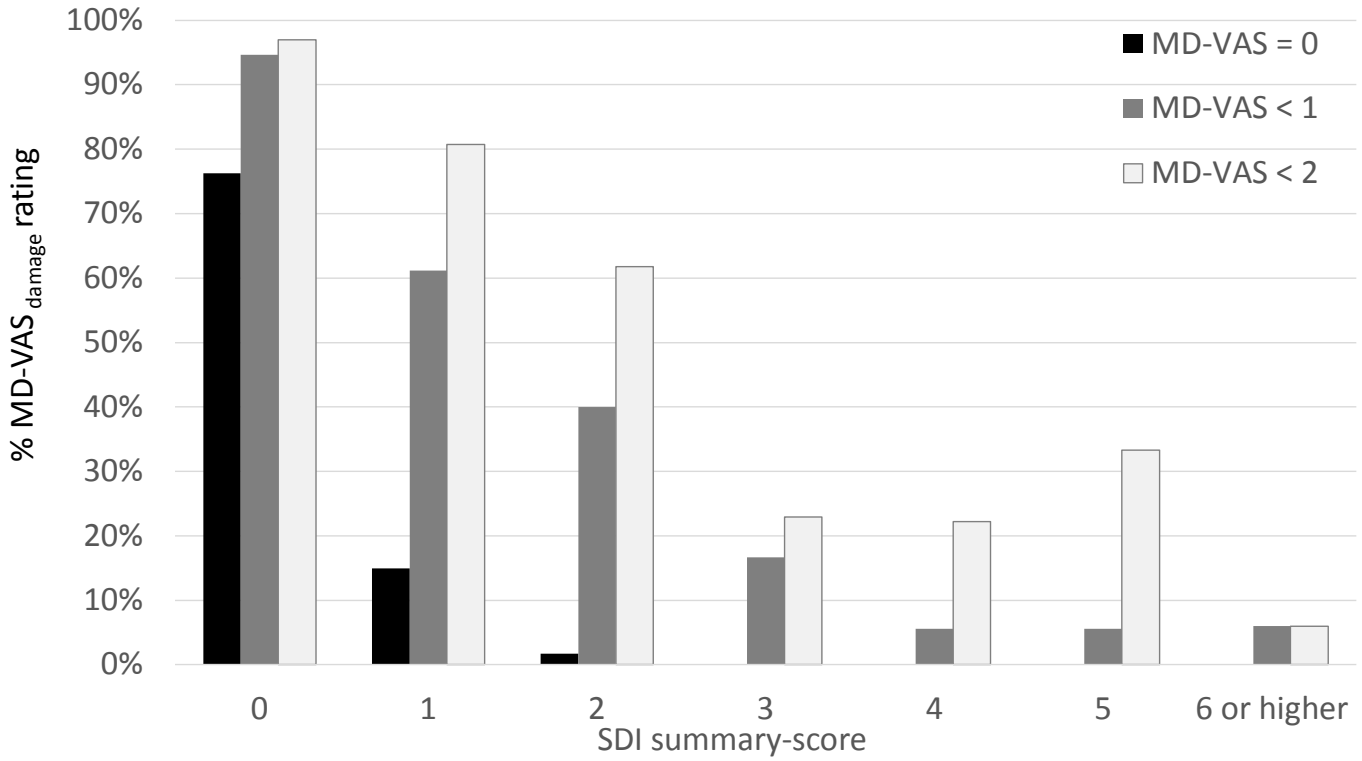
Panel 2: Observed SDI organ system involvement by disease duration for up to 6 years*



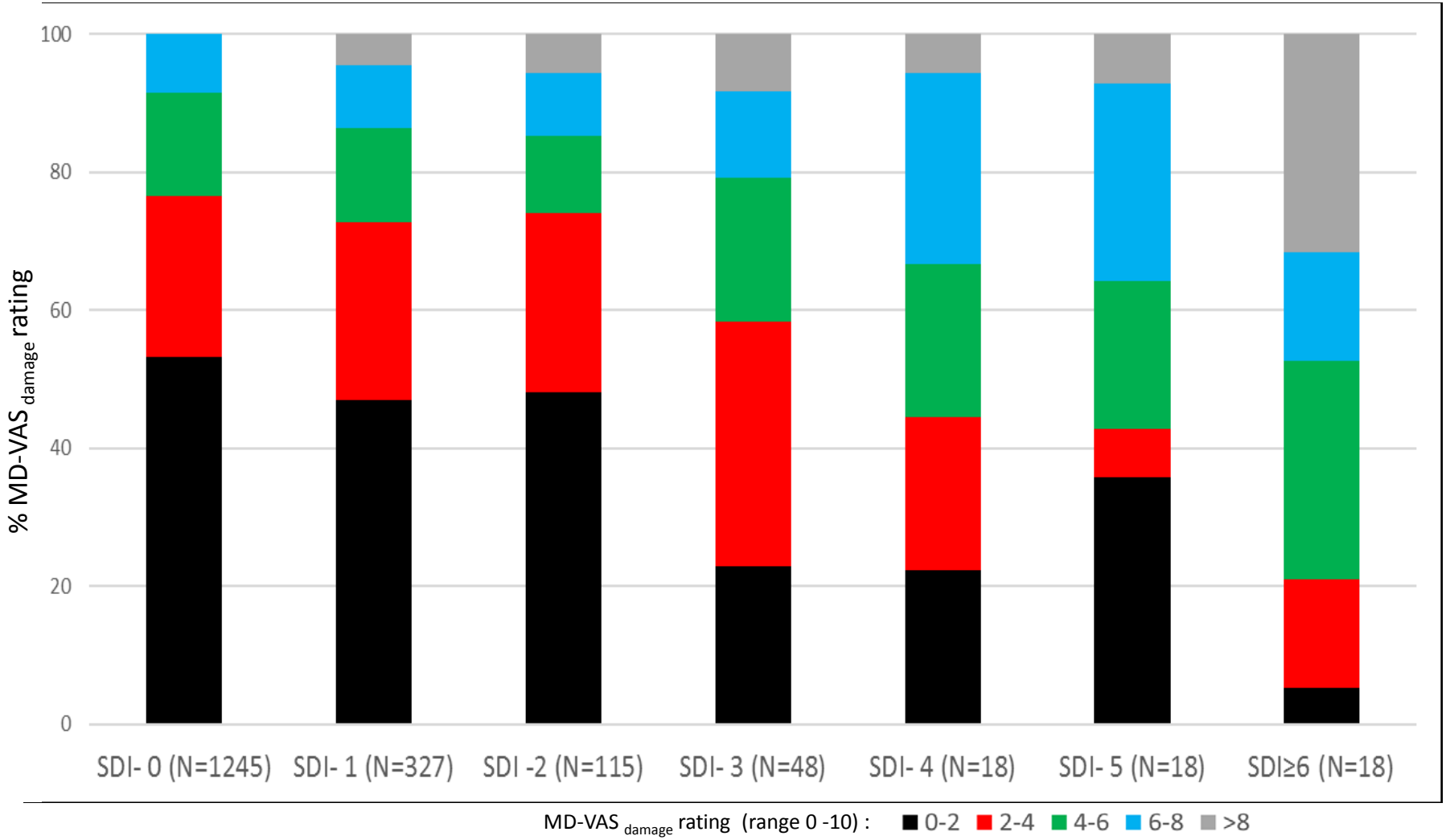
'* Malignancies and diabetes was present in < 0.4% of the composite cohort and both SDI organ systems are excluded from the figure`

Figure 2: SDI summary-scores and physician rating of perceived damage severity in cSLE.

Panel 1: MD-VAS$_{Damage}$ ratings of up to 2 (range 0 -10)



Panel 2: MD-VAS$_{Damage}$ ratings for SDI scores up to 12



OLD FIGURE or A