Organ Manifestations Influence Differently the Responsiveness of 2 Lupus Disease Activity Measures, According to Patients' or Physicians' Evaluations of Recent Lupus Activity

ERIKA R. CHANG, MICHAL ABRAHAMOWICZ, DIANE FERLAND, and PAUL R. FORTIN, for CaNIOS Investigators

ABSTRACT. ***Objective.*** To determine (1) which organ system manifestations contribute to the overall responsiveness of the Systemic Lupus Activity Measure (SLAM, revised 1991 with minor modifications as SLAM-R) and the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI); and (2) whether responsive items differ for physicians and patients.

Methods. Blinded data were obtained from repeated visits of 76 patients in the Study of Methotrexate in Lupus Erythematosus. At each visit, physicians and patients reported improvement, no change, or deterioration, and physicians then completed SLAM-R and SLEDAI. Items in SLAM-R and SLEDAI were grouped by organ system. The generalized estimating equations approach was used to measure associations between change in organ system activity and physician or patient perception of change in overall disease activity. The outcomes assessed, in separate analyses, were improvement and deterioration from the previous visit.

Results. Seventy-six patients contributed a total of 471 observations. The strongest correlates of physician-reported improvement were decreased constitutional, gastrointestinal (GI), and musculoskeletal involvement (components of SLAM-R), and decreased musculoskeletal (MSK) and central nervous system involvement (SLEDAI). Improvement reported by patients was most strongly associated with decreases in erythrocyte sedimentation rate and MSK and reticuloendothelial activity (SLAM-R), and in MSK activity (SLEDAI). Increased integument and MSK subscores (SLAM-R) and serosal and MSK subscores (SLEDAI) were associated with overall deterioration reported by physicians. Patient-reported deterioration was associated with increased GI subscores (SLAM-R) and with no changes in organ system involvement in SLEDAI.

Conclusion. Organ systems associated with reported change in overall SLE activity differed between SLAM-R and SLEDAI, between patients and physicians, and between each direction of change. (J Rheumatol 2002;29:2350–8)

Key Indexing Terms:
SYSTEMIC LUPUS ERYTHEMATOSUS RESPONSIVENESS DISEASE ACTIVITY

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by unpredictable flares and remissions, and by a diversity of clinical manifestations. It develops in people of all ages and both sexes, but primarily affects women of childbearing age[1]. Improved survival rates of patients with SLE have led to increasing interest in studying other outcomes, such as disease activity. Over 60 different measures of SLE activity with varying psychometric properties exist[2]. Two measures commonly used in North America are the Systemic Lupus Activity Measure (SLAM)[3], which was revised in 1991 with minor modifications as the SLAM-R, and the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI)[4]. Although SLAM-R and SLEDAI are both valid and reli-

From The Arthritis Centre of Excellence, the Division of Rheumatology, the Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontario; Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec; Division of Clinical Epidemiology, Montreal General Hospital, McGill University Health Centre, Montreal, Quebec; and the Canadian Network for Improved Outcomes in SLE (CaNIOS).

Supported in part by grants from the Canadian Arthritis Network, The Arthritis Society (Grant 95072), and the Natural Sciences and Engineering Research Council of Canada (Grant 105521-98). Dr. Abrahamowicz is a Health Scientist of the Canadian Institutes of Health Research. Dr. Fortin is a Senior Research Scholar of The Arthritis Society.

E.R. Chang, MSc, Research Associate, The Arthritis Centre of Excellence, University Health Network, University of Toronto; M. Abrahamowicz, PhD, Associate Professor of Biostatistics, Department of Epidemiology and Biostatistics, McGill University, Division of Clinical Epidemiology, Montreal General Hospital; D. Ferland, BScN, CaNIOS National Coordinator, The Arthritis Centre of Excellence, University Health Network, University of Toronto; P.R. Fortin, MD, MPH, Division of Rheumatology, Toronto Western Hospital Associate Professor of Medicine, Director of Clinical Research, Arthritis Centre of Excellence, University Health Network, University of Toronto.

Address reprint requests to Dr. M. Abrahamowicz, Division of Clinical Epidemiology, The Montreal General Hospital, 1650 Cedar Ave., Room L10-520, Montreal, Quebec H3G 1A4. E-mail: michal@michal.ri.mgh.mcgill.ca

Submitted February 6, 2002; revision accepted May 9, 2002.

able[3,5,6], less is known about their ability to detect change. Recent studies have indicated that although both are responsive to change, SLAM-R is slightly more so[6,7]. One of these studies also demonstrated that of the 2, SLAM is more sensitive to patient assessments of change. Our previous analyses support these findings[8]. In particular, score changes in SLAM-R, but not in SLEDAI, were shown to reflect both improvement and deterioration reported by patients. However, the responsiveness of both SLAM-R and SLEDAI was stronger for changes reported by physicians than by patients. This disparity may be caused by differences in the relevance to physicians or patients of changes in specific manifestations.

Our objectives were to determine (1) which organ system manifestations included in SLAM-R and SLEDAI contribute to the responsiveness of these instruments to patients' and physicians' assessments of changes in overall SLE activity; and (2) whether the type of organ system manifestations correlating with reported change differs between patients and physicians.

MATERIALS AND METHODS

Population. A secondary analysis was performed on blinded data obtained from the Study of Methotrexate in Lupus Erythematosus (SMILE). Conducted by the Canadian Network for Improved Outcomes in SLE (CaNIOS), this was a multicenter randomized controlled trial that compared the effects of methotrexate and placebo on disease activity in patients with lupus. Patients were included if they met the following criteria: ≥ 4 American College of Rheumatology (ACR) criteria for SLE, a SLAM-R score ≥ 8, and a Systemic Lupus International Collaborating Clinics/ACR Damage Index or SDI[9] score ≤ 15; taking stable doses of nonsteroidal antiinflammatory drugs, prednisone, or antimalarials; having no conditions such as World Health Organization class IV lupus nephritis that required treatment with other medications; using an effective method of birth control; and being the legal age of consent. After allocation to one of the 2 treatment arms, the study participants were assessed at monthly visits over a blinded phase lasting 12 months. Approval for this study was granted by the research ethics boards of all participating centers, and informed consent was provided by all recruited patients.

Data collection. At each visit, the physician and the patient recorded, independently of each other, their response to the question "Over the past month, has the lupus been...?" on a transition scale, with the possible answers "much worse," "worse," "no change," "better," and "much better." Such scales, although inherently subjective, have been used in previous research on SLE[7] and other diseases. This physician then completed the non-laboratory components of SLAM-R and SLEDAI. To ensure that the physician attending the patient was blinded with respect to the patient's treatment arm, another physician reviewed all laboratory measures of SLAM-R and SLEDAI. Since patients were assessed monthly during the followup period, each patient contributed multiple observations to the dataset. The data were entered in Medlog[10] and analyzed with SAS version 8[11].

Disease activity measures. SLAM-R includes 23 clinical and 7 laboratory items representing 11 organ systems, plus one "miscellaneous" item for scoring manifestations not listed elsewhere (Table 1). The content of the instrument is based on the frequency of appearance, ability to measure, and ease of operationalizing the manifestations. Each item is described, and *ad hoc* ascertainment and scoring rules for the "miscellaneous" item are recorded by the physician. Item scores depend on both the absence/presence and the severity of organ involvement, with higher scores indicating more severe manifestations. The maximum possible score is 84 if the miscellaneous item is included, and 81 if it is not. SLAM-R covers manifestations occurring during the preceding month.

SLEDAI[4] consists of 16 clinical and 8 laboratory items representing 9 organ systems (Table 1). The score weighting assigned to each item was derived from multiple regression modelling, with the dependent variable being physician global score for SLE activity and the independent variables being the various manifestations included. In contrast with SLAM-R, only the presence or absence of a manifestation is recorded, but the maximum possible score for a given item varies according to the perceived seriousness of the sign. For example, central nervous system (CNS) related items each score 8, whereas renal items each score 4, and hematological items each score only 1. In contrast to SLAM-R, only events occurring within the 10 days up to and including the day of measurement are recorded. The maximum possible SLEDAI score is 105, but scores higher than 46 are rare[4].

Classification of SLAM-R and SLEDAI items. Table 1 shows the grouping of items by organ system, with separate columns for SLAM-R and SLEDAI. With a few exceptions that are described below, this grouping was based on the subgroups defined in SLAM-R[3], and on the groupings used in the development of SLEDAI[4]. In our grouping of SLEDAI items, myositis was categorized with CNS manifestations, since a similar grouping was used in SLAM-R and it was felt that this manifestation reflected CNS involvement. In SLAM-R, we classified erythrocyte sedimentation rate (ESR) separately from the other items, and the other laboratory scores were divided into renal and hematological signs. Although vasculitis appears in both instruments, we grouped it differently in each, since it involves only skin blood vessels in SLAM-R, but any organ vessels in SLEDAI. Visual manifestations were classified as a CNS sign in SLEDAI but not in SLAM-R, as they include both anterior and posterior eye activity in SLEDAI, but only anterior activity in SLAM-R. For each group of items shown in Table 1, we calculated the maximum theoretical organ-specific subscore by summing the scores of all items within the group, using the same weights that are used in the instruments themselves.

Data cleaning. In both questionnaires, some items, particularly those involving manifestations detected by laboratory tests, were not scored, but a total score was still calculated. To avoid making assumptions about the presence or absence of a manifestation when no item score was recorded, we excluded visits from the analysis if either SLAM-R or SLEDAI was incomplete.

Previous research showed a systematic decrease in both SLAM-R and SLEDAI scores between the first 2 visits, regardless of the reported change in disease activity[7,8]. In contrast, the mean changes in both scores for all subsequent pairs of consecutive visits were very close to zero. This suggests that an apparent systematic decrease in SLAM-R and SLEDAI scores right after the baseline assessment reflects, at least partly, regression to the mean. This could occur because imposing a lower limit on baseline scores automatically implies that the observed mean score at the baseline will overestimate the true mean score. Specifically, in the presence of a random measurement error, imposing a lower limit will likely eliminate some patients for whom the observed baseline score (< 8) is lower than the "true" level of disease activity (≥ 8), while including some patients for whom the observed baseline score (≥ 8) overestimates the true activity (< 8). We chose therefore to omit the score change between the first and second visit from our analyses, to prevent it from biasing the responsiveness estimates.

Since few respondents reported the lupus being "much better" or "much worse," these categories were combined with "better" and "worse," respectively, resulting in a total of 3 transition categories: "better," "same," and "worse."

Estimation of associations between changes in SLAM-R or SLEDAI organ system subscores and change in overall SLE activity perceived by patients or physicians. SLAM-R and SLEDAI subscore changes were defined as the subscore at the present visit minus the subscore at the previous visit. We

Table 1. Categorization of items into organ system subgroups.

Organ System	Items (SLAM–R)	Maximum Subscore	Items (SLEDAI)	Maximum Subscore
CNS/neuromotor	Stroke; seizure; cortical dysfunction; headache; myalgia/myositis	14	Seizure; psychosis; organic brain syndrome; visual disturbances; cranial nerve; lupus headaches; CVA; myositis	60
Cardiovascular	Raynaud's; hypertension; carditis	7	NA	NA
Constitutional	Weight loss; fatigue; fever	8	Fever	1
ESR	ESR	3	NA	NA
Eye	Cytoid bodies; hemorrhages or episcleritis; papillitis or pseudotumor cerebri	9	NA (grouped with CNS)	NA
GI	Abdominal pain	3	NA	NA
Hematological	Hematocrit; white blood cells; lymphocyte count; platelet count	12	Thrombocytopenia; leukopenia	2
Immunological	NA	NA	Increased DNA binding; low complement	4
Integument	Oral/nasal, or periungal erythema, or malar rash, or photosensitive rash or nailfold infarct; alopecia; erythematous, maculopapular rash, or discoid lupus, or lupus profundus, or bullous lesions; vasculitis (leukocytoclastic vasculitis, urticaria, palpable purpura, livedo reticularis, ulcer, or panniculitis)	9	New rash; alopecia; mucous membrane ulcers	6
Musculoskeletal	Joint pain	3	Arthritis	4
Other	*Ad hoc* subscale	3	NA	NA
Pulmonary	Shortness of breath or pain	3	NA	NA
Renal	Serum creatinine; urine sediment	6	Urinary casts; hematuria, proteinuria; pyuria	16
Reticuloendothelial	Lymphadenopathy; hepato- or splenomegaly	4	NA	NA
Serosal	NA	NA	Pleurisy; pericarditis	4
Vascular	NA	NA	Vasculitis	8
Total		84		105

SLAM–R: Systemic Lupus Activity Measure-Revised[3]; SLEDAI: Systemic Lupus Erythematosus Disease Activity Index[4]; CNS: central nervous system; CVA: cerebrovascular accident; ESR: erythrocyte sedimentation rate; NA: not applicable.

modelled the probabilities of perceived improvement and deterioration as functions of changes in either SLAM-R or SLEDAI organ-specific scores. To increase the statistical power of our analyses, we included multiple measurements for each patient. To account for the correlation between subsequent measurements on the same patients, we employed the generalized estimating equations (GEE) approach to logistic regression[12], using an autoregressive order one [AR(1)] correlation structure[13]. SLAM-R and SLEDAI were assessed in different models. Separate analyses were carried out using either (1) decreased versus unchanged activity, or (2) increased versus unchanged activity, as the binary dependent variable. In each model, the outcome was defined based on the reported change in overall SLE activity from the previous visit, and the independent variables corresponded to the simultaneous changes in each organ system subscore of the instrument. Since all subscore changes were modelled simultaneously, a change in one subscore was automatically adjusted for changes in all other subscores, which allowed us to assess independent effects of particular organs.

In logistic regression, the regression coefficients and corresponding odds ratios (OR) represent the estimated effect of changing a given independent variable by 1 unit. However, because item weights in SLEDAI are preset and vary between organ groups according to perceived seriousness of the organ involvement[4], some subscores for SLEDAI organ groups change only by very specific increments. For example, arthritis and renal increase and decrease by multiples of 4 only. Therefore, when OR and 95% confidence intervals (CI) were calculated for SLEDAI item subgroups, the regression parameters and corresponding standard errors were first multiplied by the smallest possible change in the respective organ system subscores. Although the items conventionally grouped in the CNS organ system are each assigned a score of 8, the inclusion of myositis as a CNS manifestation in these analyses meant that it was possible for the CNS subscore to change by multiples of 4, and the OR for this organ system were calculated accordingly. Since it was possible for SLAM-R subscores to change by 1-point intervals, the OR and CI for their change were calculated directly from the regression parameters.

RESULTS

Study population. Eighty-six patients enrolled in SMILE and contributed 761 post-baseline visit pairs during the blinded phase of the study. Visit pairs were omitted from the analyses for the following reasons: the physician or patient transition score for the second visit in the pair was missing (58 visits); the visit had taken place at the patient's home so

SLAM-R and SLEDAI scores were not recorded (27 visits); or the instrument scoring was incomplete (205 visits). As a result, a total of 471 visit pairs, contributed by 76 patients, were used in these analyses. Baseline characteristics of the 76 patients are shown in Table 2. The number of visit pairs contributed by individual patients ranged from 1 to 12 (median 6, interquartile range 4–9). Patients reported improvement at 176 visits (37.4%), no change in 192 visits (40.8%), and deterioration in 109 visits (21.9%). In comparison, physicians reported improvement, no change, and deterioration at 161 (34.2%), 221 (46.9%), and 89 (18.9%) visits, respectively. The median score change when no change was reported was 0 for SLAM-R and SLEDAI, and the median score changes when either improvement or deterioration were reported were in the expected directions but did not exceed ± 2.

Tables 3 and 4 show the percentage of total visits in which increases and decreases in organ subscores were recorded in SLAM-R and SLEDAI, respectively. SLAM-R and SLEDAI subscore changes were generally small, with the median change being the smallest possible increment for all organ systems (data not shown). The proportion of visits in which change in either direction was recorded ranged from 0.2%, for the eye item in SLAM-R, to 46.7%, for the integument item group in SLAM-R, but for all organs and both scales the most prevalent category was "no change."

Associations between changes in organ system subscores and changes reported by patients and physicians. In Tables 5 to 8, the strength of the association between score change and reported change in overall disease activity is represented by odds ratios. The 2 columns of Table 5 show the OR for improvement reported by physicians and patients, respectively, corresponding to a 1-point decrease in subscores of SLAM-R. The effect of each organ-specific subscore change is adjusted for simultaneous changes in all other organ systems. The variable corresponding to changes in eye manifestations was omitted from the final models, because this type of change was too infrequent. As expected, the point estimates of most of these OR were greater than 1, indicating that the greater the decrease in a given subscore, the more likely it was that physicians and patients would report an improvement of disease activity. Decreased musculoskeletal (MSK) activity was the only change having a statistically significant association with improvement reported by both physicians and patients, while improved constitutional and GI activity were statistically significantly associated with improvement reported by physicians, and a decreased ESR score with patient-reported improvement. There were borderline significant associations between decreased ESR and GI activity and physician- and patient-reported improvement, respectively. The OR for association between change in reticuloendothelial activity and patient-reported improvement was statistically significant, but in an unexpected manner, since it suggests that a 1-point decrease in activity in this system would be associated with a lower, not higher, probability of patient-reported improvement.

Table 2. Demographic characteristics and baseline scores (76 patients contributed at least one visit pair to the analysis).

Characteristic	Median (IQR)	%
Current age, yrs	39.8 (33.9–48.3)	
Age at diagnosis	30.9 (24.8–42.5)	
Sex (female)		90.8
Marital status		
Single		26.3
Married		60.5
Separated, widowed, or divorced		13.2
Education (completed high school)		44.7
Baseline scores		
SLAM–R	12 (9–14)	
SLEDAI	10 (6–14)	
SDI	1 (0–2)	

SDI: Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index[9].
IQR: interquartile range.

Table 3. Frequency of organ system subscore changes in SLAM–R.

Organ System	Visits with Score Decrease, %	Visits with No Score Change, %	Visits with Score Increase, %
Constitutional	25.5	51.4	23.1
Cardiovascular	13.4	72.2	14.4
ESR	10.4	79.0	10.6
Eye	0.2	99.8	0
GI tract	6.2	87.7	6.2
Hematological	21.7	53.9	24.4
Integument	25.5	53.3	21.2
Musculoskeletal	22.7	56.9	20.4
Neurological	22.5	59.2	18.3
Ad hoc	7.0	87.1	5.9
Pulmonary	14.9	72.4	12.7
Renal	19.5	60.9	19.5
Reticuloendothelial	5.5	88.5	5.9

Table 4. Frequency of organ system subscore changes in SLEDAI.

Organ System	Visits with Score Decrease, %	Visits with No Score Change, %	Visits with Score Increase, %
CNS	4.5	90.7	4.8
Constitutional	2.8	94.7	2.6
Hematological	1.7	95.5	2.8
Immunological	10.4	78.1	11.5
Integument	18.1	65.2	16.8
Musculoskeletal	11.7	78.8	9.6
Renal	15.9	67.3	16.8
Serosal	1.9	96.0	2.1
Vasculitis	1.9	96.6	1.5

Table 5. OR for association between reported improvement in overall disease activity and 1-point decreases in SLAM-R subscores. Results of the GEE approach to logistic regression for repeated measurements. The OR for a 1-point score decrease in each organ system is adjusted for the effects of all other organs. Expected OR > 1.0.

Organ System	OR (95% CI) Physician	Patient
Neuromotor	1.19 (0.92–1.55)	1.01 (0.82, 1.25)
Cardiovascular	0.84 (0.60–1.19)	0.92 (0.74, 1.14)
Constitutional	1.70 (1.35–2.13)*	1.20 (0.91, 1.59)
ESR	1.42 (0.97, 2.08)†	1.52 (1.09, 2.13)*
GI tract	1.87 (1.03, 3.40)*	1.57 (0.99, 2.49)†
Hematological	1.10 (0.88, 1.37)	1.06 (0.89, 1.27)
Integument	1.06 (0.79, 1.41)	0.93 (0.73, 1.20)
Musculoskeletal	1.35 (1.01, 1.79)*	1.41 (1.03, 1.92)*
Ad hoc	1.17 (0.67, 2.04)	0.79 (0.43, 1.47)
Pulmonary	1.19 (0.77, 1.84)	0.94 (0.62, 1.42)
Renal	1.01 (0.80, 1.26)	0.96 (0.78, 1.17)
Reticuloendothelial	1.36 (0.85, 2.18)	0.68 (0.47, 0.99)*

† $p \leq 0.1$; * $p \leq 0.05$.

This effect is only marginally significant ($p = 0.0451$) and it is likely to represent an inflated type I error, due to multiple testing.

As an example of the clinical implications of these changes, a 1-point decrease in the MSK activity scale subscore of SLAM-R could mean either presence of objective inflammation without limited joint function where there had been inflammation with limited function at the previous visit, or absence of activity when arthralgia had been reported previously. The interpretation of the OR for organ systems containing 2 or more items was more complex, since either a manifestation could disappear and be replaced by another with a lower severity score, or its severity might decrease while the severity of other manifestations remained stable. A comparison between physician and patient point estimates reveals that the OR for physician-reported improvement tended to be greater, indicating that physicians' overall assessments were affected slightly more by changes in specific organ involvement.

Table 6 shows the OR for association between reported improvement in overall disease activity as assessed by either physicians or patients, and decreases in individual SLEDAI

Table 6. OR for association between reported improvement in overall disease activity and theoretical minimum decreases in SLEDAI subscores. Results of the GEE approach to logistic regression for repeated measurements. The OR for the smallest possible score decrease in each organ system is adjusted for the effects of change in all other organs. Expected OR > 1.0.

Organ System	Score Change on Which OR Is Based	OR (95% CI) Physician	Patient
CNS	–4	1.49 (1.09, 2.03)*	1.17 (0.82, 1.67)
Constitutional	–1	0.64 (0.25, 1.61)	1.16 (0.48, 2.83)
Hematological	–1	1.55 (0.35, 6.83)	1.29 (0.68, 2.42)
Immunological	–2	1.07 (0.65, 1.74)	1.08 (0.67, 1.75)
Integument	–2	1.38 (0.95, 2.00)†	1.10 (0.86, 1.39)
Musculoskeletal	–4	1.81 (1.09, 2.99)*	2.04 (1.21, 3.43)**
Renal	–4	1.22 (0.91, 1.64)	0.92 (0.70, 1.22)
Serosal	–2	0.45 (0.15, 1.34)	0.57 (0.21, 1.54)
Vasculitis	–8	0.56 (0.16, 1.93)	1.40 (0.28, 6.92)

†$p \leq 0.1$; *$p \leq 0.05$; **$p \leq 0.01$.

organ system subscores, after adjustment for change in other subscores. A decrease in one of the subscores signifies either the disappearance of manifestations from the previous visit, or, in the case of integument involvement, the continuing presence of activity with no worsening. Because it was only possible for SLEDAI subscores to change by intervals corresponding to the prespecified weights of individual items, the OR in Table 6 correspond to the smallest possible subscore decrease that could actually have occurred in that group of items. These OR were of roughly the same magnitude as those for 1-point decreases in the corresponding SLAM-R subscores. Decreased CNS and MSK activity showed a statistically significant association with improvement reported by physicians, while decreased integumentary activity had a borderline statistically significant association. In contrast, change in MSK activity was the only type having a statistically significant association with improvement reported by patients. However, this association was very strong, being in fact greater than that between decreased MSK activity and physician-reported change.

Tables 7 and 8 focus on increases in SLE activity. Table 7 shows the OR for the association between reported overall increases in SLE activity and 1-point increases in the score of individual SLAM-R organ systems, after adjustment for change in all other organ systems. The 1-point score increase indicated either onset of a manifestation or increase in its severity. Here, an OR > 1 indicated that as an organ system subscore increased, the assessor was more likely to report worsening of activity. For both patients and physicians, there were statistically significant and borderline significant associations. The change most strongly associated with patient-reported increase in activity was in GI involvement, while those for constitutional and integumentary manifestations were borderline significant. Worsening integument and MSK involvement showed the statistically strongest associations with deterioration reported by physicians.

Table 7. OR for association between reported deterioration in overall disease activity and 1-point increases in SLAM–R subscores. Results of the GEE approach to logistic regression for repeated measurement. The OR for a 1–point score increase in each organ system is adjusted for the effects of change in all other organs. Expected OR > 1.0.

Organ System	OR (95% CI) Physician	Patient
Neuromotor	0.81 (0.60, 1.09)	1.01 (0.77, 1.31)
Cardiovascular	1.04 (0.68, 1.61)	0.90 (0.65, 1.24)
Constitutional	1.13 (0.86, 1.48)	1.29 (0.96, 1.73)†
ESR	0.96 (0.61, 1.53)	0.86 (0.57, 1.31)
GI tract	0.68 (0.37, 1.25)	1.83 (1.04, 3.19)*
Hematological	0.97 (0.76, 1.23)	1.00 (0.81, 1.24)
Integument	1.65 (1.28, 2.12)**	1.23 (0.96, 1.57)†
Musculoskeletal	2.11 (1.44, 3.07)**	1.23 (0.91, 1.65)
Ad hoc	1.09 (0.65, 1.83)	1.16 (0.75, 1.78)
Pulmonary	1.26 (0.80, 2.00)	1.32 (0.77, 2.28)
Renal	1.26 (0.90, 1.76)	1.07 (0.82, 1.39)
Reticuloendothelial	1.07 (0.62, 1.87)	1.09 (0.68, 1.76)

† $p \leq 0.1$; * $p \leq 0.05$; ** $p \leq 0.01$.

In comparison to the findings for SLAM-R, increases in only a few of the SLEDAI organ system subscores were systematically associated with worsening of overall SLE activity reported by patients (Table 8). Only increased serosal activity showed a borderline significant association with deterioration recorded by patients on the transition scale. A higher probability of physician-reported deterioration, on the other hand, corresponded with increased arthritis and serositis, with borderline associations with fever and integument subscores. Further, the OR for association between subscore increase and reported change tend to be higher for physicians than for patients. These results suggest that different factors may have influenced the patients and physicians in their assessments of overall worsening of SLE activity.

The organ systems in which there was a statistically significant association, at the $\alpha = 0.05$ level, between change in the level of involvement and change reported by physicians or patients are shown in Table 9. Whereas changes in MSK activity were associated with patients' and physicians' perceptions quite consistently, the statistical significance of associations between other types of changes and improvement or deterioration recorded on the transition scales varied by assessor, direction of recorded change, and instrument.

DISCUSSION

Measurement of SLE activity is complicated by the variety of manifestations that may appear both in different people and in the same person over time. Instruments such as SLAM-R[3] and SLEDAI[4] have been developed that quantify activity in several organ systems and produce an aggregate disease activity score. Although these instruments were modelled after physician judgment of SLE activity at single points in time, their use in monitoring patients over time requires that we understand better how the instruments behave with respect to both physician and patient assessments of change in SLE activity.

This study evaluated the association between organ-specific changes in SLAM-R and SLEDAI scores and improvement or deterioration in overall SLE activity that was reported by physicians and patients. We found that the strength and statistical significance of the effect of changes in subscores for specific organs on the global assessments by physicians and patients differed between SLAM-R and SLEDAI. In addition, the relationships between changes in organ involvement and reported changes in overall disease activity depended on whether the assessor was the physician or the patient, and on whether improvement or deterioration was reported (Table 9). Overall, there did not seem to be any type of change in organ activity that was a correlate of

Table 8. OR for association between reported deterioration in overall disease activity and theoretical minimum increases in SLEDAI subscores. Results of the GEE approach to logistic regression for repeated measurements. The OR for the smallest possible score increase in each organ system is adjusted for the effects of change in all other organs. Expected OR > 1.0.

Organ System	Score Change on Which OR Is Based	OR (95% CI) Physician	Patient
CNS	4	1.16 (0.67, 2.02)	0.94 (0.60, 1.46)
Constitutional	1	4.48 (0.97, 20.76)†	1.59 (0.57, 4.46)
Hematological	1	2.10 (0.67, 6.53)	1.16 (0.41, 3.27)
Immunological	2	1.30 (0.65, 2.62)	1.30 (0.91, 1.87)
Integument	2	1.47 (0.97, 2.23)†	1.19 (0.89, 1.60)
Musculoskeletal	4	3.14 (1.69, 5.83)**	1.00 (0.58, 1.74)
Renal	4	1.15 (0.78, 1.71)	0.96 (0.68, 1.35)
Serosal	2	3.46 (1.28, 9.36)*	2.32 (0.91, 5.95)†
Vasculitis	8	2.65 (0.74, 9.47)	0.60 (0.15, 2.46)

†p ≤ 0.1; * p ≤ 0.05; ** p ≤ 0.01.

Table 9. Organ system changes with statistically significant association with physician- and patient-reported changes in overall SLE activity†.

	Better		Worse	
	Physician	Patient	Physician	Patient
SLAM-R	Constitutional	ESR	Integument	GI
	GI	MSK	MSK	
	MSK	Reticuloendothelial*		
SLEDAI	CNS	MSK	MSK	(none)
	MSK		Serosal	

†α = 0.05; * OR < 1.0.

reported change regardless of assessor, SLE activity instrument, and reported direction of change, although changes in MSK involvement showed a statistically significant association in all cases, except with patient-reported worsening in the logistic regression model using SLEDAI subscores as the independent variables. However, it did appear that, compared to patients, physicians' overall assessments of the changes in SLE activity showed more frequent and somewhat stronger associations with organ-specific changes.

Although some strong predictors of reported change were common to physicians and patients, the fact that responsiveness of some organ systems is unique to each suggests that different factors may have determined the importance of change perceived by patients and their physicians. Ward, *et al*[6] and Fortin, *et al*[7] postulated that the inclusion of subjective (patient-reported) manifestations in SLAM and SLAM-R, respectively, but not in SLEDAI, might account for the higher responsiveness of the SLAM to changes in patient global assessments of SLE activity. Because SLE manifestations such as pain and fatigue can be difficult to confirm objectively, their presence may not always be reflected in SLEDAI scores even if it affects patients. This suggestion was supported by a previous analysis in which we categorized manifestations in SLAM-R according to whether or not they could be observed directly by physicians[8]. We found that score changes in subjective manifestations enhanced the overall responsiveness of SLAM-R and were associated, independently of score changes in objective manifestations, with the responsiveness of SLAM-R to patient-reported changes in disease activity. The importance to patients of changes in subjective manifestations may thus explain some of the differences found in this study in the types of organ system changes associated with physician- and patient-reported overall change.

Differences in the item definitions and scoring systems of SLAM-R and SLEDAI may have accounted for some of the disparities in organ-specific responsiveness. SLEDAI records presence or absence of organ system-specific activity and reflects perceived severity of specific manifestations through preassigned item weights[4], with the most weight given to CNS involvement and the least weight allotted to fever and hematological activity. In addition, types of organ involvement included in this instrument are restricted to those that can be objectively confirmed; therefore, manifestations such as fatigue and pain in the absence of inflammation are not recorded. In contrast, SLAM-R weights each item equally, but includes a scoring gradient that measures the severity of a given manifestation. Also, it records manifestations that might not be observed directly by the physician and so must be reported by the patient. It is possible that changes too small to be detected by SLEDAI were still able to be recorded in SLAM-R. For example, a transition from lack of joint pain to arthralgia, or from objective inflammation to "limited function," would not have qualified as a change in activity on SLEDAI, yet, based on results in Tables 5 and 7 (47% and 80% increase in the probability of physician-reported improvement and deterioration, respectively, for a 1-point change in this item in

SLAM-R), both were important to observers. On the other hand, the inclusion of subjective manifestations in SLAM-R may have resulted in the recording of non-SLE related changes in the patient. Although activity recorded on SLAM-R was theoretically restricted to manifestations attributable to SLE, it was possible that the physician assessment was influenced by patient non-SLE related or psychosomatic complaints, or by a placebo effect from the clinical trial.

This study differed from others[6,7,14] in that it modelled how the probability of an observer reporting changes depends on a change in a given organ-specific component of the overall SLAM-R or SLEDAI, rather than on a change in the overall instrument score. The use of repeated measures increased the precision of estimates and reduced concerns about low statistical power, typical of SLE studies, which often suffer from small sample sizes. This allowed us to model the effects of changes in organ-specific scores while simultaneously adjusting for changes in all other organs.

Some limitations to this analysis should be discussed. First, it was not possible to directly evaluate the relationship between changes in renal, immunological, or hematological activity with reported changes in overall disease activity. This was because laboratory data, which may have revealed the treatment arms of SMILE participants, were not seen by the physicians evaluating global change in the patients. However, there may have been observable, although indirect, effects of activity in some of these systems on the patient, as changes in both renal and hematological scores were sometimes associated with perceived overall changes. The statistically significant but unexpected association between increased SLAM-R hematological scores and patient-reported deterioration may simply reflect inflated type I error, due to multiple testing.

Second, it should be noted that the transition scales we used have not been validated, and do not specify the reasons for changes in overall SLE activity that were reported by the physicians and patients. Although ideally the reasons were related to the organ system-specific changes recorded in SLAM-R and SLEDAI, we cannot exclude the possibility that increased or decreased organ involvement not detected by either instrument prompted the physician or patient to report improvement or deterioration in SLE. One inherent limitation of studies of responsiveness to change in SLE is the absence of an objective "gold standard" for clinically relevant change. Other researchers have evaluated disease activity instrument score changes with respect to either the transition scale used in the present analysis[7] or the changes in overall disease activity scores recorded on visual analog scales[6], or they have analyzed instrument score changes between 2 time points in which a change in overall disease activity was assumed to have occurred[15]. By using transition scores we may be able to avoid some of the "noise" from measurement error that may be generated with other approaches. However, future research should investigate the intrarater reliability of both assessments as well as the interrater reliability of physicians' responses.

In addition, we should also consider the possibility that participation in the trial might have altered the patients' or the physicians' perception of the relevance of changes in disease activity. For example, the patients might have been more likely than usual to feel they had improved. If this were the case, then we might observe an asymmetry in the responsiveness of SLAM-R and SLEDAI to perceived decreases and increases in disease activity.

Finally, the set of usable observations was substantially reduced by the absence of item scores from some of the observations, as a consequence of pending laboratory results. Although this may have affected the precision of some of the estimates, the exclusion of these observations is not likely to have biased the OR.

Generalization of these results to all patients with SLE and their physicians should be done cautiously. Because we performed a secondary data analysis, our study population was defined by the inclusion and exclusion criteria of the original randomized controlled trial. This meant that the initial SDI score had to be 15 or less, and the SLAM-R score had to be 8 or more. Also, patients were excluded if they were unable to comply with the treatment regimen. It is possible that there were systematic differences in perception of the relevance of change in disease activity by those with more organ damage, or by noncompliers.

The precision of some of the OR estimates, and the statistical power, may have been decreased by the lack of activity or change in activity of organ systems such as the eye, shown in Tables 3 and 4. Further work is needed, therefore, in other groups of patients with different patterns of change in organ involvement, to increase the precision of some of the results reported here and to assess to what extent the lack of statistically significant effect of changes in some organs may be due to insufficient frequency of such changes in our study. It may also be useful to investigate the effects of level of organ damage on the magnitude and statistical significance of associations between changes in specific manifestations and reported overall change, and to determine whether the pattern of change in organ-specific activity is related to patient compliance with treatment.

In conclusion, we have found that changes in organ systems that appear to be important determinants of the overall change in disease activity that is reported vary depending on the assessor, the instrument, and the direction of perceived overall change. Changes in some of the self-reported manifestations appear to be more meaningful to patients than to physicians. The differences between patient and physician evaluation, and the differences in the relevance to each of change in activity recorded by SLAM-R and SLEDAI, may partly explain why previous research[8] found that both instruments are more responsive to physi-

cians' than to patients' perceptions of change and why SLAM-R appears to be more responsive than SLEDAI. These results highlight the need for better communication between patients and physicians, and for inclusion of more patient-reported manifestations in instruments if they are to respond better to patient assessments of change in disease activity.

ACKNOWLEDGMENT

The authors thank Roxane du Berger and Marielle Olivier for their valuable assistance with SAS.

APPENDIX

Investigators and co-investigators in the Study of Methotrexate in Lupus Erythematosus (SMILE).

Principal Investigator: P.R. Fortin, MD, MPH, University Health Network, Toronto, Ontario (formerly at McGill University Health Centre, Montreal, Quebec).

Co-Investigators and Institutions (listed in order of number of patients entered into the study): M. Abrahamowicz, PhD; D. Ferland, RN, BScN; A.E. Clarke, MD, MSc; J. Penrod, PhD, McGill University Health Centre, Montreal, Quebec. D. Lacaille, MD, MHSc; J.M. Esdaile, MD, MPH; H.B. Stein, MD, Arthritis Research Centre of Canada, Vancouver, BC. D. Smith, MD, Ottawa Hospital, General Campus, Ottawa, Ontario. M. Zummer, MD; J-P. Mathieu, MD; L. Duchesne, MD; S. Mercille, MD; P. Dagenais, MD, PhD, Hôpital Maisonneuve-Rosemont, Montreal, Quebec. J.E. Pope, MD, MPH, St. Joseph's Hospital, London, Ontario (formerly at Victoria Hospital, London, Ontario). S. Edworthy, MD; S. Barr, MD; G. Morris, MD, Calgary Health Sciences Centre, Calgary, Alberta. M. Starr, MD, St. Mary's Hospital, Montreal, Quebec. V. Bykerk, MD, Mount Sinai Hospital, Toronto, Ontario (formerly at Credit Valley Hospital, Mississauga, Ontario). S. Bernatsky, MD; J. Canvin, MD; H.S. El-Gabalawy, MD; C. Peschken, MD, Winnipeg Health Sciences Centre, Winnipeg, Manitoba. A. Cividino, MD, McMaster-Chedoke Hospital, Hamilton, Ontario. J-L. Senécal, MD; J-P. Raynauld, MD; E. Rich, MD, Hôpital Notre-Dame, Montreal, Quebec. C. Kirk Osterland, MD; C.A. Yeadon, MD, Royal Victoria Hospital, Montreal, Quebec. A. Beaulieu, MD; S. Carette, MD, Centre Hospitalier de l'Université Laval, Quebec City, Quebec. G. Boire, MD, MSc, Centre Hospitalier de l'Université de Sherbrooke, Sherbrooke, Quebec. SMILE research assistants: R. Morrison, RN, MSN; E.S. Clark; K. Rangno, RN, Arthritis Research Centre of Canada, Vancouver, BC. P. Dale, Ottawa General Hospital, Ottawa, Ontario. D.L. Fenlon, BScN; W. Curran; L. Bere, RN, Victoria Hospital, London, Ontario. R. Talavera; B. Green, BN, BScN; E. Teixeira, RN, BN, Calgary Health Sciences Centre, Calgary, Alberta. J. O'Farrell, RN, Credit Valley Hospital, Mississauga, Ontario. S. Wood, RN; A. Huggard, RN, Winnipeg Health Sciences Centre, Winnipeg, Manitoba. J. Zahavich, RN, McMaster-Chedoke Hospital, Hamilton, Ontario. M. Prave, RN, Centre Hospitalier de l'Université Laval, Quebec City, Quebec.

REFERENCES

1. Hochberg M. The epidemiology of systemic lupus erythematosus. In: Wallace DJ, Hann BH, editors. Dubois' lupus erythematosus. Philadelphia: Lea & Febiger; 1993:49-57.
2. Liang MH, Stern S, Esdaile JM. Systemic lupus erythematosus activity. An operational definition. Rheum Dis Clin North Am 1988;14:57-66.
3. Liang MH, Socher SA, Larson MG, Schur PH. Reliability and validity of six systems for the clinical assessment of disease activity in systemic lupus erythematosus. Arthritis Rheum 1989;32:1107-18.
4. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. Arthritis Rheum 1992;35:630-40.
5. Petri M, Hellmann D, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. J Rheumatol 1992;19:53-9.
6. Ward MM, Marx AS, Barry NN. Comparison of the validity and sensitivity to change of 5 activity indices in systemic lupus erythematosus. J Rheumatol 2000;27:664-70.
7. Fortin PR, Abrahamowicz M, Clarke AE, et al. Do lupus disease activity measures detect clinically important change? J Rheumatol 2000;27:1421-8.
8. Chang E, Abrahamowicz M, Ferland D, Fortin P. Comparison of the responsiveness of lupus disease activity measures to changes in systemic lupus erythematosus activity relevant to patients and physicians. J Clin Epidemiol 2002;55:488-97.
9. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. Arthritis Rheum 1996;39:363-9.
10. Medlog Systems. Incline Village, NV: Medlog Systems; 1995.
11. SAS. Cary, NC: SAS Institute Inc.; 1999.
12. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13-22.
13. Jennrich RI, Schluchter MD. Unbalanced repeated-measures models with structured covariance matrices. Biometrics 1986;42:805-20.
14. Gladman DD, Goldsmith CH, Urowitz MB, et al. Sensitivity to change of 3 systemic lupus erythematosus disease activity indices: international validation. J Rheumatol 1994;21:1468-71.
15. Brunner HI, Feldman BM, Bombardier C, Silverman ED. Sensitivity of the Systemic Lupus Erythematosus Disease Activity Index, British Isles Lupus Assessment Group Index, and Systemic Lupus Activity Measure in the evaluation of clinical change in childhood-onset systemic lupus erythematosus. Arthritis Rheum 1999;42:1354-60.