UCSF
UC San Francisco Previously Published Works

Title
Longitudinal validation of the Brief Index of Lupus Damage.

Permalink
https://escholarship.org/uc/item/3f77k7cd

Journal
Arthritis care & research, 66(7)

ISSN
2151-464X

Authors
Katz, Patricia
Trupin, Laura
Rush, Stephanie
et al.

Publication Date
2014-07-01

DOI
10.1002/acr.22268

Peer reviewed

Longitudinal validation of the Brief Index of Lupus Damage (BILD)

Patricia Katz, PhD[1], **Laura Trupin, MPH**[1], **Stephanie Rush, BA**[1], and **Jinoos Yazdany, MD**[1]
[1]University of California, San Francisco

Abstract

Objective—To provide further validation of the Brief Index of Lupus Damage (BILD), assessing its sensitivity to change in disease status and ability to predict mortality risk.

Methods—Data were from the UCSF Lupus Outcomes Study (LOS; n=958), in which the BILD was administered twice, approximately 5 years apart. We examined disease activity and health care utilization among participants who completed the BILD twice (n=745). We identified increases in disease activity and utilization that would suggest a disease flare between the two BILD administrations and compared their occurrence by BILD score increases (0, 1, 2, 3, >3). Deaths were ascertained when annual interviews were attempted. Kaplan-Meier life-table analysis compared mortality rates for four groups of initial BILD scores (0, 1, 2,   3), and differences were tested using a log-rank test. Using Cox proportional hazard models, we estimated risk of death associated with higher BILD scores.

Results—BILD score increases were associated with spikes in disease activity (p=0.05) and physician visits (p=0.004) over baseline, and with hospitalizations (p<0.001) during the intervening years. Of those with BILD score increases >3, 84% were hospitalized prior to the second BILD. During follow-up there were 60 deaths (6.3%). BILD scores of 2 (hazard ratio: 6.1 [95% confidence interval 1.3–30.0]) and   3 (10.8 [2.5–46.2]) were associated with higher risk of death.

Conclusion—This analysis provides evidence of the BILD's predictive validity and ability to detect change. While not intended to replace clinical evaluation of disease damage, the BILD appears to be a useful tool for research.

Measuring damage due to systemic lupus erythematosus (SLE) has traditionally relied upon physician assessments such as the Systemic Lupus International Collaborating Clinics/ American College of Rheumatology Damage Index (SDI)[1]. Studies have supported the validity and reliability of the SDI, but because it requires a trained physician with either comprehensive knowledge of the patient or access to comprehensive medical records to complete a 41-item questionnaire, it is not feasible to use it in large-scale epidemiological or clinical research. We developed a patient-reported measurement of lupus disease damage, the Brief Index of Lupus Damage (BILD,) to address the need of these larger scale studies[2].

Corresponding author: Patricia Katz, PhD, University of California, San Francisco, Arthritis Research Group, 3333 California Street, Suite 270, San Francisco, CA 94143-0920, Phone: 415-476-5971, Fax: 415-476-9030, patti.katz@ucsf.edu.

The BILD was designed to be administered by trained interviewers. Our cross-sectional evaluation of the BILD among 900 individuals with SLE supported the content, criterion, and construct validity of the BILD. We found good agreement and moderately high Spearman's rank correlations between the SDI and the BILD. Higher BILD scores were also associated with longer disease duration, higher disease activity, poorer functional status, and increased health care utilization. Since the original publication of the BILD, self-administered versions have been tested and validated in English[3] and German[4]. Each of these exhibited psychometric characteristics similar to those we previously reported.

To date, the BILD has not been tested longitudinally to determine if it can predict important outcomes that might reflect changes in SLE-related damage. Our objective in this paper was to provide further validation of the BILD assessing its sensitivity to change in disease status and its ability to predict higher mortality risk.

Methods

Data source

Data were drawn from the UCSF Lupus Outcomes Study (LOS), which has previously been described in detail[5]. Briefly, LOS participants were recruited from an existing cohort of individuals with confirmed SLE diagnoses, based on medical chart reviews supervised by a rheumatologist. Except for the initial confirmation of SLE diagnostic criteria, the data for the LOS are primarily collected via annual standardized telephone interviews. Enrollment began in 2003 and continued through 2010, although 80% of participants were enrolled prior to 2006. The study was approved by the UCSF Committee on Human Research, and all participants gave written informed consent.

Brief Index of Lupus Damage (BILD)

The BILD was developed as a self-report proxy for the SDI. The SDI consists of 41 items covering 12 organ systems, designed to quantify cumulative organ damage due to SLE regardless of attribution. Thirty-one items can score 1 point, 6 items can score 2 points, and 1 item can score 3 points, yielding a potential score range of $0 - 46$[6]. The SDI is completed by a trained physician and requires either comprehensive knowledge of the patient or access to extensive medical records.

In developing the BILD, the goal was not to replicate the SDI item for item, but to develop a proxy that would be sensitive to differences in disease damage. At the same time, items needed to be comprehensible to a lay respondent. Therefore, the BILD does not replicate the SDI item for item, because either some manifestations were considered to be too rare to be likely to contribute meaningfully to the score (e.g., shrinking lung) or not likely to be interpreted with enough specificity to capture the concept of damage (e.g., alopecia). The instrument contains 28 items that capture information on 26 of the original 41 SDI items. Development of the BILD items and validation of the measure is described in detail elsewhere[2]. The BILD was first administered in 2007. Participants who joined the study after that time completed it as they enrolled. All participants in the 2012 interview

completed a second BILD. Thus, for most of the cohort, the two BILD scores were approximately 5 years apart, however the interval ranged from 11 – 70 months.

Variables

Disease status—To assess changes in disease status, we examined both disease activity and health care utilization. These measures were chosen because the manifestations included in the BILD would typically give rise to increases in symptoms or utilization at the time they occur. We therefore identified increases in symptoms or utilization that would suggest a flare in disease activity between the first and second administration of the BILD. Our assumption was that such flares could represent SLE-related events that could lead to increases in disease-related damage. Disease activity was measured with the Systemic Lupus Activity Questionnaire (SLAQ), a validated patient-reported measure of disease activity[7,8]. For utilization, both the number of physician visits and hospitalizations were examined. The SLAQ and utilization measures are collected annually and self-reported by participants.

Mortality—Deaths were ascertained when annual telephone interviews calls were attempted and recorded based on report from family members of the deceased, using the alternate phone contacts provided by subjects at enrollment if necessary. For an earlier mortality analysis of this cohort[9], a National Death Index (NDI) search was also conducted through 2006. At that time, only 3 cohort members who had been lost to follow-up were identified as deceased in the NDI. The percentage of participants lost to follow-up in any year (net of refusals and deaths) has remained very low throughout the study (0.5% annually).

Analysis

Sensitivity to change—Only participants who completed the BILD two times were included in this analysis (n = 745). We calculated changes in BILD scores, categorizing increases in scores as 0, 1, 2, 3, or >3 (higher scores represent more accrued damage). These categories were chosen based on the distribution of BILD change scores. Only increases in scores are possible because, by definition, damage cannot be reversed. We then examined increases from baseline in disease activity measured with the SLAQ, increases in the number of physician visits, and hospitalizations that occurred in the years between the two BILD scores. Increases in SLAQ and physician visits were defined as an increase of 0.5 standard deviation over baseline in at least one interview wave between BILD administrations[10], equivalent to a 4-point increase in SLAQ and 5 additional physician visits. For hospitalizations, we examined the presence of any hospitalization during the follow-up period and of any single wave with at least two hospitalizations. We tested differences among the BILD change groups using chi square tests for trend.

Mortality—A total of 958 participants completed the BILD instrument at least one time between 2007 and 2010; that group is included in the mortality analysis. For subjects who did not die during the study period, follow-up time (i.e., time to censorship) was based on the number of months from their baseline BILD score until their last interview, with the exception of 6 subjects who were not known to have died but who were not interviewed again after their baseline BILD. We assigned these individuals a follow-up time of 6

months, equivalent to half the average interval between interview waves. For deceased subjects, follow-up time was calculated as the number of months between the baseline BILD score and the date of death, if known. If the exact date of death was not reported, the follow-up time was based on the midpoint of the interval within which their death was known to have occurred. Fifteen participants had no known date of death; the average interval used to estimate follow-up time was 10 months (range 2–18 months).

The baseline BILD scores were divided into 4 categories corresponding to 0, 1, 2, and 3 points (higher scores represent more damage). Kaplan-Meier life table analysis was used to compare mortality rates by categories of the BILD, and differences were tested using a log rank test. Using Cox proportional hazard models, multivariate analyses of possible predictors of mortality were performed. In addition to the BILD categories, predictors examined included SLAQ, non-white ethnicity, education beyond high school, gender, age, and disease duration.

As a sensitivity analysis, we re-estimated the life table and proportional hazards analyses using both endpoints of the interval for subjects without a precise follow-up time. The results were virtually identical to the mid-point analysis reported here and are not shown. Additionally, because of a report that renal damage was the most important predictor of mortality within the SDI[11], we were interested to see whether the BILD was associated with mortality if we excluded the end stage renal disease (ESRD) items. Therefore, we re-calculated the BILD without the ESRD items and entered the revised BILD and a separate variable for ESRD into the Cox regression model described above.

SAS version 9.3 (Cary, NC) was used for the sensitivity to change analyses, and Stata v.12 for the mortality analysis.

Results

Among 958 subjects completing the BILD at the initial administration, the mean age was 49 years, 92% were female, 63% Caucasian, and mean disease duration was 16 years (Table 1). There were no differences between participants who were followed (n = 745) through the second BILD interview compared to those who dropped out prior to the 2012 interview but were not known to have died (n = 159) in gender, age, race, disease duration, baseline SLAQ score, number of doctor visits, or baseline BILD (data not shown). Those who did not remain in the study were, however, more likely to have been hospitalized in the baseline year and less likely to have completed college.

BILD scores ranged from 0 – 13 with a median of 2 (IQR 1 – 3) in the first administration. At the second administration, BILD scores ranged from 0 – 18 with a median of 2 (IQR 1 – 4). Half of the sample reported at least one additional manifestation of damage in the second BILD administration. The average annual change in BILD was 0.21 points.

Sensitivity to change in disease status

Changes in BILD ranged from 0 – 13 with a median of 1 (IQR 0 – 1). Of the 745 LOS participants with two BILD scores, 372 had no increase in BILD score. Of the remaining

participants 210, 84, 41, and 38 had increases of 1, 2, 3, and >3 points, respectively (Table 2). Overall, 37% of LOS participants had an increase in SLAQ in at least one wave, 43% had an increase in physician visits in at least one wave, 40% had one or more hospitalizations during the study period, and 14% had 2 or more hospitalizations in a single year. The prevalence of an increase in SLAQ scores was highest among participants with changes in BILD >3. There was a greater prevalence of increased physician visits among participants with a change in BILD >0, and the χ^2 for trend was significant (p = 0.004), reflecting a greater likelihood of an increase in physician visits among those with greater increases in BILD. Greater increases in BILD were associated with significantly greater likelihood of any hospitalizations during the study period or 2 hospitalizations reported at any interview wave (χ^2 for trend p < 0.0001). Over 80% of the people with BILD increases 3 were hospitalized between the two BILD assessments.

Prediction of mortality

During the follow-up period there were 60 deaths (6.3%) among the 958 LOS participants who had an initial BILD score. Mean follow-up time was 49 months (range 4–75). In comparing participants who died to those who survived, regardless of follow-up status, there were no differences in gender, race, education, or baseline SLAQ (data not shown). Those who died were significantly older, had longer disease duration, and more physician visits and hospitalizations in the baseline year.

For initial BILD scores of 0, 1, 2, and 3, the mortality rates were 0.9%, 2.0%, 6.3%, and 13.1%, respectively (Table 3). In Cox proportional hazards models adjusting for age, gender, disease duration, race/ethnicity, education and SLAQ score, a BILD score of 2 (hazard ratio [HR]: 6.1; 95% confidence interval [CI]: 1.3 – 30.0), and a BILD score 3 (HR: 10.8; 95% CI: 2.5 – 46.2) were associated with higher risk of death over the follow-up period. Race, education, gender and SLE disease activity were not associated with increased mortality risk in that model. Analyses of the revised BILD excluding the ESRD items showed a similar pattern, although hazard ratios were somewhat reduced (BILD = 2, HR: 4.9; 95% CI: 1.4 – 17.6; BILD score 3, HR: 6.7; 95% CI 1.9 – 23.1). For comparison, the hazard ratio for ESRD based on the renal disease items was 2.9 (95% CI: 1.4 – 5.9). Kaplan-Meier analysis also showed significantly higher mortality with increasing BILD scores (Log-rank chi square = 43.1 p<0.001; Figure 1).

Discussion

Our previous analyses provided evidence of the cross-sectional validity of the BILD as a brief patient-reported measure of lupus disease damage. In these analyses, we demonstrate the predictive validity of the BILD. The BILD was both sensitive to change in disease status, increasing after disease events that were likely to represent accrual of damage, and a strong predictor of mortality in this cohort.

In order for a health outcome measure to be useful in longitudinal studies, it must be able to detect real change when such change occurs. We found that BILD scores increased following rises in disease activity and physician visits and following hospitalizations that occurred between the first and second BILD administration. Each of these events can be

assumed to reflect disease manifestations or events that could cause accumulation of disease damage. In these analyses, there was a stronger relationship between increases in BILD scores and hospitalizations than between increases in BILD scores and increases in disease activity or physician visits. Hospitalizations may be a stronger reflection of disease events that could cause damage accrual, so the difference noted in these relationships might be expected.

We designed the BILD to capture the overall concept of damage in SLE for epidemiologic research. It is important to note that it is not a direct substitute for the SDI, since the BILD omits many items from the SDI that were either not suitable for patient self-report or were not informative because of their frequent over-reporting by patients. Instead, among a group of patients (rather than in any individual patient), the BILD is able to differentiate between those with high or low degrees of SLE damage and identify those with higher risks of mortality. Indeed, the mortality hazard ratios from the BILD are equivalent to, and in some cases stronger than, those reported for the SDI[12–14].

There are potential limitations to this study. Some individuals thought to be lost to follow-up may have died. Our prior search of the National Death Index showed that this was not common (only 3 of those lost to follow-up were found to have died), and even so, such occurrences would tend to bias our results toward negative results. As noted above, these analyses include only English-speaking, primarily well educated individuals. While we and others have begun to expand the administration to non-English-speaking and less well-educated groups, such individuals are not included in this analysis.

It is possible that some individuals over-reported disease manifestations, particularly individuals with scores at the high end of the scale. For example, the maximum BILD increase was 13. While it is possible that someone could be accurately reporting such an increase in disease manifestations, it is also possible that this represents measurement error. To minimize this effect, we did not use a linear form of BILD increases in our analyses, but categorized changes instead, with the largest increase group capped at 3. Additionally, the average annual increase in BILD score in our cohort was 0.21, a rate of change that is within the range of SDI changes noted in other cohorts: in the LUMINA cohort an average 1-year change of 0.4 was observed[15], and an inception cohort from the SLICC network reported an average annual change over 5 years of 0.17[16].

A reliable, valid, and sensitive patient-reported measure of SLE damage greatly expands the range of SLE research, permitting epidemiological and other observational studies to include this important factor both as a covariate and an outcome. The evidence we present in this manuscript adds to the evidence supporting the usefulness and validity of the BILD. Since our original publication of the instrument[2], a self-reported version has been developed and validated in a predominantly African-American lupus cohort with lower levels of education[3], and a German version has also been validated[4]. Although our original administration included only English-speaking individuals, we are now administering the BILD in Spanish, after a rigorous translation process.

In summary, while previous cross-sectional examination of the BILD showed concurrent validity, this analysis provides evidence of the BILD's ability to detect change and predictive validity. Importantly, the BILD differentiated between those who later had high mortality risk from their disease and those who did not. While not intended to replace clinical evaluation of disease damage, the BILD does appear to be a useful tool for research.

Acknowledgments

References

1. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus Erythematosus International Collaborating Clinics/American College of Rheumatology Damage Index for SLE. Arthritis Rheum. 1996; 39:363–369. [PubMed: 8607884]

2. Yazdany J, Trupin L, Gansky S, et al. The Brief Index of Lupus Damage (BILD): a patient-reported measure of damage in SLE. Arthritis Care & Research. 2011

3. Drenkard, C.; Yazdany, J.; Trupin, L., et al. Validity of a self-administered version of the Brief Index of Lupus Damage (BILD) in a predominantly black sysgtemic lupus erythematosus population in the United States. 10th International Conference on SLE; April, 2013; Buenos Aires, Argentina. 2013.

4. Chehab G, Sander O, Richter J, et al. Validation and evaluation of the German Brief Index of Lupus Damage (BILD) - a self-reported instrument to record damage in systemic lupus erythematosus. Lupus. 2013; 22:1050–1055. [PubMed: 23963433]

5. Yelin E, Trupin L, Katz P, et al. Work dynamics among persons with systemic lupus erythematosus. Arthritis Rheum (Arthritis Care Res). 2007; 57:56–63.

6. Romero-Diaz J, Isenberg D, Ramsey-Goldman R. Measures of adult systemic lupus erythematosus. Arthritis Care & Research. 2011; 63:S37–S46. [PubMed: 22588757]

7. Karlson E, Daltroy L, Rivest C, et al. Validation of a systemic lupus activity questionnaire (SLAQ) for population studies. Lupus. 2003; 12:280–286. [PubMed: 12729051]

8. Yazdany J, Yelin E, Panopalis P, Trupin L, Julian L, Katz P. Validation of the systemic lupus erythematosus activity questionnaire in a large obsevational cohort. Arthritis Rheum (Arthritis Care Res). 2008; 59:136–143.

9. Hersh AO, Trupin L, Yazdany J, et al. Childhood-onset disease as a predictor of mortality in an adult cohort of patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken). 2010; 62(8):1152–1159. [PubMed: 20235215]

10. Norman G, Sloan J, Wyrwich K. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003; 41:582–592. [PubMed: 12719681]

11. Danila M, Pons-Estel G, Zhang J, Vila L, Reveille J, Alarcón G. Renal damage is the most important predictor of mortality within the damage index: data from LUMINA LXIV, a multiethnic US cohort. Rheumatology. 2009; 48:542–545. [PubMed: 19233884]

12. Nived O, Jonsen A, Bengtsson A, Bengtsson C, Sturfelt G. High predictive value of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for survival in systemic lupus erythematosus. J Rheumatol. 2002; 29:1398–1400. [PubMed: 12136895]

13. Chambers S, Allen E, Rahman A, Isenberg D. Damage and mortality in a group of British patients with systemic lupus erythematosus followed up for over 10 years. Rheumatology. 2009; 48:673–675. [PubMed: 19359343]

14. Sutton E, Davidson J, Bruce I. The Systemic Lupus International Collaborating Clinics (SLICC) damage index: a systemic literature review. Semin Arthritis Rheum. [epub ahead of print].

15. Alarcon G, Roseman J, McGwin GJ, et al. Systemic lupus erythematosus in three ethnic groups. XX. Damage as a predictor of further damage. Rheumatology. 2004; 43:202–205. [PubMed: 12923289]

16. Urowitz M, Galdman D, Ibanez D, et al. Evolution of disease burden over five years in a multicenter inception systemic lupus erythemasus cohort. Arthritis Care & Research. 2012; 64:132–137. [PubMed: 21954226]

Significance and Innovation

- A reliable, valid, and sensitive patient-reported measure of SLE damage greatly expands the range of SLE research, permitting epidemiological and other observational studies to include this important factor both as a covariate and an outcome.

- Previous work established the content and criterion, validity of the BILD.

- This study demonstrates the ability of the BILD to detect change and predict higher mortality risk.
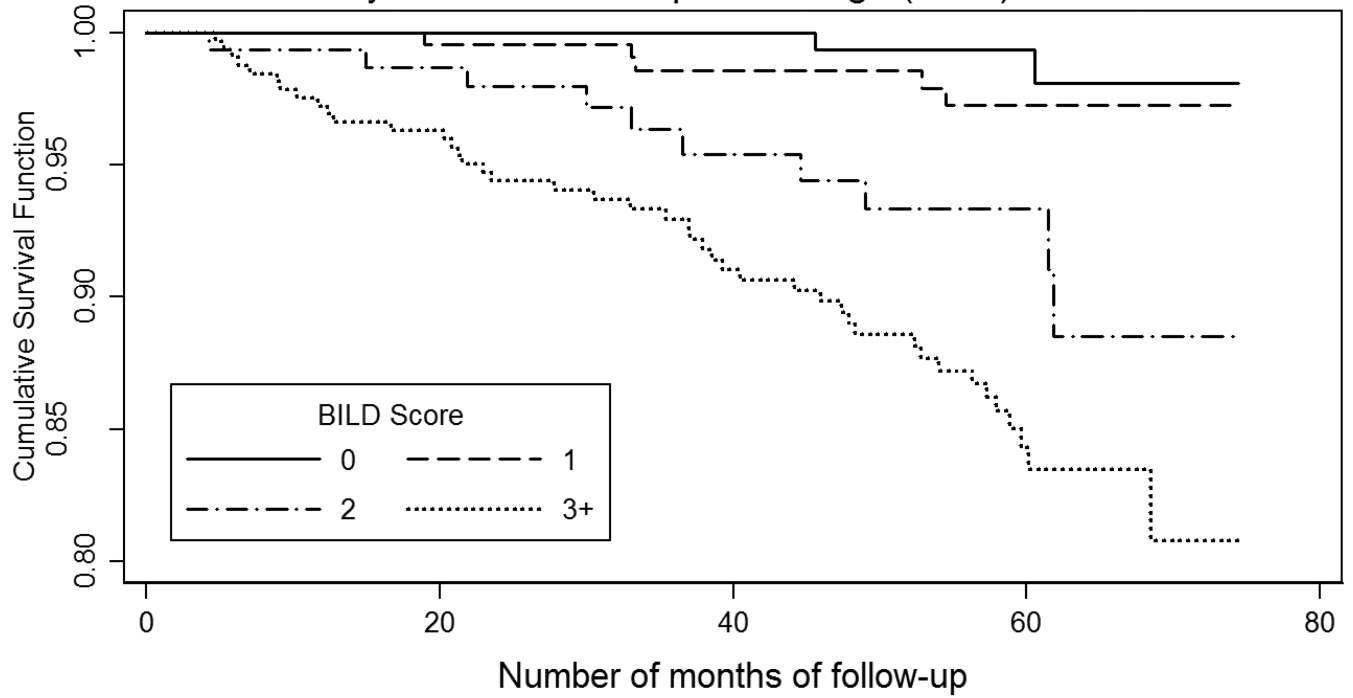
Figure 1.
Relationship between the Brief Index of Lupus Damage (BILD) and mortality: Results of Kaplan-Meier analysis.
Note: Log-rank chi square = 43.1 p<0.001

Table 1

Baseline characteristics among participants who completed one or two BILD interviews

	All participants with at least 1 BILD score (n = 958)	Participants with 2 BILD scores (n = 745)
Female, % (n)	92% (885)	93% (693)
Age, mean (SD) years	49 (13)	48 (13)
Race/ethnicity		
White, non-Hispanic	63% (599)	61% (457)
Hispanic	10% (95)	10% (78)
African American	11% (103)	10% (76)
Asian	10% (100)	11% (83)
Other/unknown	6% (61)	7% (51)
Education		
High School or less	16% (151)	15% (114)
Some College/AA/Trade	46% (437)	44% (324)
Bachelors degree or more	39% (369)	41% (306)
Disease duration, mean (SD) years	16 (9)	16 (9)
Medications during the past year		
Glucocorticoids (oral/IV)	58% (558)	56% (414)
Other immunosuppressants[*]	29% (274)	28% (211)
Only glucocorticoids	34% (329)	32% (241)
Both glucocordicoids and immunosuppressants	24% (229)	23% (173)
SLAQ score, mean (SD)	12 (8)	12 (8)
Annual number of physician visits, mean (SD)	16 (11)	16 (11)
Hospitalizations in baseline year, % (n)		
none	77% (739)	80% (613)
1	14% (138)	13% (86)
2 or more	8% (81)	6% (40)
Baseline BILD		
Median	2	1
IQR	1 – 3	0 – 3
Range	0 – 13	0 – 13
Follow-up BILD		
Median	-----	2
IQR		1 – 4
Range		0 – 18
Change in BILD		
Median	-----	1
IQR		0 – 1
Range		0 – 13
Average annual change in BILD		
Mean (SD)	-----	0.21 (0.30)

	All participants with at least 1 BILD score (n = 958)	Participants with 2 BILD scores (n = 745)
Range		0 – 2.4

*
immunosuppressants include azathioprine, mycophenolate mofetil, methotrexate, cyclosporine, leflunomide, cytoxan, etanercept, rituximab, infliximab, adalimumab, and abatacept

Table 2

Increases in disease activity and physician visits and hospitalizations associated with subsequent increases in BILD scores

| | Total (n=74) | Increase in BILD | | | | | | p* |
		0 (n=372)	1 (n=210)	2 (n=84)	3 (n=41)	>3 (n=38)	
SLAQ (at least 1 score > baseline by 0.5 SD)	37%	33%	40%	36%	39%	50%	.053
MD visits (at least 1 wave > baseline by 0.5 SD)	43%	36%	50%	44%	56%	50%	.004
Any hospitalizations during the study period	40%	29%	42%	52%	68%	84%	<.0001
Hospitalizations (at least 1 wave with ≥2)	14%	5%	19%	15%	29%	55%	<.0001

*
p value from χ² for trend.

Note: Increases in disease activity (i.e., SLAQ score) and utilization were selected to reflect increases in symptoms or utilization that would suggest a flare in disease activity between the first and second administration of the BILD that could lead to increases in disease-related damage.

Table 3

Mortality by initial BILD scores

	BILD score			
	0	**1**	**2**	**3**
Total BILD score				
Number subjects	225	247	158	328
Number of deaths	2	5	10	43
% mortality	0.9%	2.0%	6.3%	13.1%
Hazard ratio (95% CI), unadjusted	(reference)	2.4 (0.5, 12.2)	7.8 (1.7, 35.5)	15.5 (3.7, 63.8)
Hazard ratio (95% CI), adjusted[*]	(reference)	2.0 (0.4, 10.2)	6.1 (1.3, 30.0)	10.8 (2.5, 46.2)
BILD score omitting renal disease items				
Number of subjects	239	271	174	274
Number of deaths	3	7	13	37
% mortality	1.3%	2.6%	7.5%	13.5%
Hazard ratio (95% CI), adjusted[†]	(reference)	1.8 (0.5, 7.0)	4.9 (1.4, 17.6)	6.7 (1.9, 23.1)

[*] adjusted for age, sex, disease duration, race/ethnicity, education, and disease activity (SLAQ)

[†] adjusted for variables listed above as well as incident dialysis or renal transplant.