Clinical measurement of lupus nephritis activity is inferior to biomarker-based activity assessment using the renal activity index for lupus nephritis in childhood-onset systemic lupus erythematosus

Najla Aljaberi 1, Scott E Wenderfer, Arjun Mathur, Tingting Qiu, Steffy Jose, Angela Merritt, James Rose, Prasad Devarajan, Bin Huang, Hermine Brunner

INTRODUCTION
Renal involvement in childhood-onset SLE (cSLE), that is, lupus nephritis (LN), is associated with a significant burden of morbidity, with 10%–30% of patients with cSLE progressing to end-stage renal disease within 15 years of diagnosis.3 Indeed, renal damage is among the most important predictors of mortality in cSLE.2 3 Conventional tools, such as urine sediment and proteinuria are inferior to kidney biopsy to diagnose LN and determine the degree of renal inflammation, that is, LN activity.4 5 Although kidney biopsies are considered the gold standard to accurately capture the degree of activity inflammation with LN,6 7 the renal domain score of the SLE disease activity index (rSLEDAI) is among the most commonly used clinical measures of LN activity.3 8 This is of concern given that studies in adults with LN have shown that very low rSLEDAI scores...
can be present despite the persistence of renal inflammation as seen on kidney biopsy.\(^9\)\(^{10}\) Thus, additional tools or tests are needed to help clinicians monitor LN activity accurately in a clinical setting.

Urinary biomarkers have been in the spotlight as potential diagnostic and prognostic tools of LN activity.\(^11\)\(^{12}\) Most studies investigating urinary biomarkers in LN are focused on previously known protein targets. However, unbiased discovery proteomics has led to the identification of novel biomarkers that are potentially more clinically relevant. Based on the results of proteomics, the renal activity index in lupus (RAIL) was developed as a non-invasive tool to evaluate LN activity.\(^13\) The RAIL is a composite measure that is calculated from the natural log transformed concentrations of six urinary biomarkers, namely neutrophil gelatinase-associated lipocalin (NGAL), ceruloplasmin, monocyte chemoattractant protein-1 (MCP-1), adiponectin, haemopexin and kidney injury molecule-1 (KIM-1). Prior research showed that scores that are standardised for urine creatinine concentrations (RAIL-Cr) are 92% accurate in identifying high LN activity in a cohort of paediatric patients and young adult patients with LN.\(^13\) Furthermore, RAIL scores improve about 3 months prior to clinically observed response to LN therapy;\(^14\) hence can be used as early indicator of treatment response or LN flare. Validity of the RAIL biomarker to correctly capture LN throughout paediatric and adult age ranges has been confirmed.\(^14\)

The objectives of this study were to investigate in an independent cohort the potential of the RAIL biomarkers to estimate renal inflammation as seen on kidney biopsy; and to compare the performance RAIL and RAIL-Cr with that of the rSLEDAI in a paediatric clinical setting.

**METHODS**

**Patients and samples**

Patients participating in this study constitute a convenience sample of patients with cSLE followed at Cincinnati Children’s Hospital or Texas Children’s Hospital (n=115). All patients met the American College of Rheumatology criteria for the classification of SLE prior to age 18 years.\(^15\) Patients were excluded if they had renal disease not attributed to SLE.\(^16\) The following information was extracted from their medical records at the time of urine sample collection: age, gender, systemic blood pressure (SBP), diastolic blood pressure (DBP), estimated glomerular filtration rate (eGFR),\(^17\) urine protein-to-creatinine ratio (UPCR), urine albumin-to-creatinine ratio (UACR) and disease activity as measured by the SLE disease activity index (SLEDAI; possible range 0–105). Renal domain score of the SLEDAI (rSLEDAI; possible range 0–16) was used calculated as a clinical measure of LN activity,\(^3\) and the sum of the remaining SLEDAI items was used to estimate extrarenal disease activity (extra-renal SLEDAI). Reports of kidney biopsies performed in the cSLE population were obtained and included the scores of the National Institute of Health-activity index (NIH-AI) and National Institute of Health-chronicity index (NIH-Cl).\(^18\)\(^{20}\) Kidney biopsies were performed in all patients with proteinuria of at least 500 mg/day not explained otherwise, and histological findings were interpreted by one of two experienced nephropathologists. As previously suggested, renal inflammation as seen on kidney biopsy (LN-activity status) was classified as inactive, low-moderate and high LN activity for NIH-AI scores of <2, 2–10 or >10, respectively.\(^10\)\(^13\)\(^14\) Patients in whom a kidney biopsy was never clinically indicated were considered as not having renal involvement with cSLE (no-LN). Patients contributed one or more urine samples to the study. Urine samples for biomarker measurement were collected mostly on the day of the rSLEDAI measurement (maximum: 4 weeks).

**Biomarker measurement**

The supernatant from spun spot-urine samples was stored at 4°C within 1 hour of collection, followed by long-term storage at −80°C within 24 hours, and batch testing of the RAIL biomarkers. NGAL, ceruloplasmin, MCP-1, adiponectin, haemopexin, KIM-1 and urine creatinine were all assayed as previously reported by our group.\(^7\)\(^{13}\)\(^{14}\) Laboratory technicians performing biomarker assays were blinded to clinical information.

**Statistical analysis**

Calculation of the RAIL score, or RAIL-Cr score, requires that the raw concentrations of urine creatinine (in mg/mL) or the urine biomarkers (in ng/mL for NGAL, ceruloplasmin, haemopexin, adiponectin; in pg/mL for KIM-1, MCP-1) to be natural log-transformed. The RAIL score is calculated as follows: −4.29−0.34×ln(NGAL)−0.06×ln(ceruloplasmin)+0.89×ln(MCP-1)+0.18×ln(adiponectin)−0.65×ln(haemopexin)+0.62×ln(KIM-1). The RAIL-Cr score necessitates each additional natural-log transformed biomarker levels to be divided by the natural-log transformed value of urine creatinine prior to applying the RAIL algorithm. RAIL and RAIL-Cr scores can assume negative and positive values, with higher, more positive scores indicating a higher degree renal inflammation.\(^13\)

Descriptive analysis included medians and IQRs for continuous and categorical variables, respectively. Only clinical data collected within 1 month and urine samples collected within 3 months of the kidney biopsy date were associated with histological activity as seen on kidney biopsy (LN-activity status). Clinical characteristics, RAIL scores, RAIL-Cr scores and select traditional measures of patients with cSLE with LN at the time of initial kidney biopsy and at the initial study visit of patients without LN (no-LN) were compared. We also compared traditional LN measures (SBP, DBP, eGFR, UPCR, UACR, NIH-AI, NIH-Cl), biomarker concentrations, RAIL and RAIL-Cr scores by LN-activity status (inactive: NIH-AI ≤1, low-moderate: NIH-AI 2–10, high: NIH-AI >10) to the no-LN group.

Logistical regression and receiver operating characteristic curve (ROC) analyses for all patients with at least
Childhood lupus

One biopsy was done to test the diagnostic accuracy assessments of the RAIL, RAIL-Cr and rSLEDAI to capture high LN activity status (NIH-AI >10) versus lower LN activity status (NIH-AI ≤10). Values of area under the ROC curve (AUC) can be interpreted as outstanding, excellent, good, fair or poor for values of 0.91–1.0, 0.81–0.90, 0.71–0.80, 0.61–0.70 and <0.6, respectively.21 The areas under the ROC curves between measures (RAIL, RAIL-Cr, rSLEDAI) were compared using the test by DeLong et al.22

Linear regression models were used to evaluate the ability for biomarker scores and/or rSLEDAI to predict NIH-AI and beta-coefficient for the standardised predictor and corresponding confidence were estimated and plotted; R-squares are reported. Lastly, only considering RAIL scores from urine samples collected >3 months of a kidney biopsy and follow-up urine samples of the no-LN group, we compared rSLEDAI scores to median RAIL scores measured at the time of kidney biopsy by LN-activity status. All analyses were conducted as two-sided tests with p<0.05 as statistically significant using SAS V.9.4 (SAS, Cary, North Carolina, USA).

RESULTS

Demographics and clinical characteristics

There was a total of 115 patients with cSLE and 299 samples available for this analysis (figure 1). The minimum age in the cohort was 7 years. Baseline characteristics of the

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Baseline clinical characteristics of study cohort by LN status*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
<td>LN (n=68)</td>
</tr>
<tr>
<td>Clinical characteristics</td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>17 (14.5, 20)</td>
</tr>
<tr>
<td>Female</td>
<td>55 (80.88%)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>31 (46.97%)</td>
</tr>
<tr>
<td>Black</td>
<td>32 (48.48%)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (4.55%)</td>
</tr>
<tr>
<td>Ethnicity (Hispanic)</td>
<td>21 (30.88%)</td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>120 (110, 131.5)</td>
</tr>
<tr>
<td>DBP (mm Hg)</td>
<td>74 (63.5, 82)</td>
</tr>
<tr>
<td>Use of ACEIs/ARBs</td>
<td>14 (20.59%)</td>
</tr>
<tr>
<td>Urine biomarker scores</td>
<td></td>
</tr>
<tr>
<td>Total RAIL score</td>
<td>0.51 (−0.69, 1.6)</td>
</tr>
<tr>
<td>Total RAIL-Cr score</td>
<td>0.68 (−0.68, 1.44)</td>
</tr>
<tr>
<td>Traditional cSLE measures</td>
<td></td>
</tr>
<tr>
<td>NIH-AI score</td>
<td>4 (2, 12)</td>
</tr>
<tr>
<td>NIH-CI score</td>
<td>1 (0, 3)</td>
</tr>
<tr>
<td>Total SLEDAI score</td>
<td>13.5 (4, 22.5)</td>
</tr>
<tr>
<td>Extrarenal SLEDAI score</td>
<td>5.5 (2, 12)</td>
</tr>
<tr>
<td>Renal SLEDAI score</td>
<td>5 (0, 12)</td>
</tr>
</tbody>
</table>

*Continuous variables are demonstrated as median (25th, 75th percentile), categorical variables are demonstrated as frequency (percentage). †P values are generated from Kruskal-Wallis test and χ² test.

ACEI, ACE inhibitor; ARB, angiotensin receptor blocker; DBP, diastolic blood pressure; LN, lupus nephritis; NIH-AI, NIH-activity index; NIH-CI, NIH-chronicity index; RAIL, renal activity index for lupus; RAIL-Cr, RAIL standardised for urine creatinine; SBP, systolic blood pressure; SLEDAI, SLE disease activity index.
study cohort are shown in table 1. There were 47 patients without renal involvement with cSLE (no-LN group) and 68 patients with LN, among them 32 with one or more visits at the time of a kidney biopsy. Indeed, 39 urine samples were collected at the time of kidney biopsies and NIH-AI scores ranged from 0 to 17. Median time since kidney biopsy was 2 days (IQR: 1–10.5). Besides having, as expected, significantly higher renal activity (rSLEDAI scores), patients with LN had markedly higher extra-renal disease activity. A minority of patients with LN were prescribed ACE inhibitor or an angiotensin receptor blocker as part of LN management.

Urinary biomarker levels and traditional LN measures at the time of kidney biopsy
Among the traditional measures of LN considered in this study, only UPCR and UACR were significantly higher among patients with LN than without renal involvement (table 2). Notably, only the UACR but not the UPCR differed significantly with LN activity (as per kidney biopsy), and the UACR was more closely related to the NIH-AI score than the UPCR (Pearson’s correlation coefficient (r=0.47; p<0.0001 vs r=0.26; p=0.004)).

Select individual urine biomarker levels (ceruloplasmin, haemopexin) included in the RAIL were significantly lower without LN than with LN, even if inactive (table 2). Both RAIL and RAIL-Cr scores significantly differed (p<0.019) among patients with different LN-activity status (inactive, low-moderate or high). As shown in table 2 and online supplemental figure S1, median (IQR) scores of the RAIL and RAIL-Cr increased with higher categories of LN activity. There was a trend towards lower RAIL and RAIL-Cr scores in patients without LN compared with those with inactive LN but differences did not reach statistical significance. Patients with inactive LN activity status had lower RAIL (median (IQR): −0.04 (−1.13, 1.00)) and RAIL-Cr scores (median (IQR): 0.02 (−1.24, 0.01)). Patients with high LN activity status had higher RAIL (median (IQR): 2.02 (1.03, 2.68)) and RAIL-Cr scores (median (IQR): 1.83 (1.08, 2.43)).

The rSLEDAI scores did not discriminate well patients with different categories of LN activity (NIH-AI score <2 vs 2–10 vs >10; p=0.76). However, median (IQR) scores of the rSLEDAI for patients with inactive LN (NIH-AI ≤1) were significantly lower than those with NIH-AI ≥2 (median (IQR): 4 (2, 10) vs 12 (10, 16); p=0.0014). Similar results for the rSLEDAI and RAIL (or RAIL-Cr) were observed when considering all 39 visits and a 3-month window around the time of available kidney biopsies (online supplemental figures S2 and S3).

Figure 2 presents ROC curve for predicting high LN activity using RAIL, RAIL-Cr and rSLEDAI scores. The RAIL performed showed good accuracy in capturing high LN activity (NIH-AI >10) and performed best (AUC=0.79), followed by the RAIL-Cr (AUC=0.73) and the rSLEDAI (AUC=0.63), although differences in AUC values between measures did not reach statistical significance (AUC: RAIL vs rSLEDAI; p=0.26). Using all the kidney biopsy results (n=39) provided similar findings (see online supplemental figure S4). The distribution of RAIL and RAIL-Cr scores in relation to NIH-AI scores over the available range of 0–17 is shown in figure 3.

Clinical monitoring for active LN
There were 246 urine samples from patients with or without LN that were collected >3 months apart from the date of a kidney biopsy. As shown in figure 4, in these urine samples not collected around the time of a kidney biopsy, the RAIL scores of patients with rSLEDAI scores of 0 or 4 supported the presence of at least low LN activity (NIH-AI >2) in 70% (n=171) of the samples. In this same category of patients with low rSLEDAI, 6% (n=16) of RAIL scores were suggestive of high activity LN (NIH-AI >10). Negative RAIL scores were present in 46% (n=114) of patients with rSLEDAI scores of 0 or 4.

DISCUSSION
We aimed to compare the accuracy of the RAIL and the renal domain-score of the rSLEDAI. While the RAIL provided a good estimate of the degree of inflammation with LN, rSLEDAI score proved to be a poor surrogate for biopsy-proven high LN activity. Furthermore, RAIL scores of patients with LN with rSLEDAI scores of 0 should raise the suspicion of the presence of at least minimal disease activity.

Like the amounts of urine albumin, concentrations of the RAIL biomarker could be influenced by the hydration status of the patient, although we have shown in the past that they are produced in the kidney rather than being filtered from the blood. In the current study, we confirm our prior research findings in that standardisation by urine creatinine does not improve the accuracy of the RAIL in capturing LN activity in paediatric cohorts. Notably, we have shown in the past that standardisation by urine creatinine but not urine albumin improves the accuracy with which the RAIL biomarkers reflect LN activity in adults with LN.

Many urinary biomarkers have been proposed for monitoring LN and have been studied internationally. We have validated many of these biomarkers in the past, found them reflective of LN activity, and associated with LN histological findings. However, we were able to confirm that the six RAIL biomarkers were sufficient to capture LN activity accurately. Although we have firmly established NGAL as a biomarker of LN activity, using the current study design, NGAL levels did not add to the discrimination of the various categories of LN activity given that NGAL is a highly responsive biomarker and urine levels quickly decline with treatment. Conversely, especially levels of MCP-1 and adiponectin continued to differentiate between categories of LN activity, suggesting that these urine biomarkers normalise over a longer time frame.

Both patients with inactive LN and patients with cSLE without LN are expected to have rSLEDAI scores of 0. In
### Table 2  Traditional renal measures and biomarker concentrations at the time of kidney biopsy by LN activity category

<table>
<thead>
<tr>
<th>Variables*</th>
<th>(a) NIH-AI &lt;2 (n=8)</th>
<th>(b) NIH-AI 2–10 (n=12)</th>
<th>(c) NIH-AI &gt;10 (n=12)</th>
<th>(d) No LN (n=47)</th>
<th>(abcd) P value†</th>
<th>(abc) P value</th>
<th>(a) versus (d) P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Urine biomarkers‡ and RAIL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAL (ng/mL)</td>
<td>3.19 (2.66, 3.6)</td>
<td>3.89 (3, 4.48)</td>
<td>3.56 (3.16, 3.78)</td>
<td>3.59 (3.15, 3.88)</td>
<td>0.3130</td>
<td>0.2413</td>
<td>0.2518</td>
</tr>
<tr>
<td>Ceruloplasmin (ng/mL)</td>
<td>8.11 (6.66, 9.5)</td>
<td>9 (8.11, 9.64)</td>
<td>8.62 (7.63, 9.32)</td>
<td>3.76 (2.99, 4.98)</td>
<td>&lt;0.0001</td>
<td>0.6420</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MCP-1 (pg/mL)</td>
<td>6.34 (5.72, 7.04)</td>
<td>7.4 (6.79, 7.87)</td>
<td>8.33 (7.59, 8.84)</td>
<td>5.9 (4.41, 6.92)</td>
<td>&lt;0.0001</td>
<td>0.0103</td>
<td>0.3518</td>
</tr>
<tr>
<td>Adiponectin (ng/mL)</td>
<td>4.1 (3.79, 4.51)</td>
<td>4.84 (4.19, 5.89)</td>
<td>6.28 (4.73, 7.1)</td>
<td>4.8 (3.6, 5.69)</td>
<td>0.0096</td>
<td>0.0053</td>
<td>0.1974</td>
</tr>
<tr>
<td>Haemopexin (ng/mL)</td>
<td>7.21 (6.64, 8.13)</td>
<td>8.38 (7.4, 8.73)</td>
<td>7.92 (7.54, 8.15)</td>
<td>6.49 (6.03, 6.89)</td>
<td>&lt;0.0001</td>
<td>0.1930</td>
<td>0.0378</td>
</tr>
<tr>
<td>KIM-1 (pg/mL)</td>
<td>6.8 (5.94, 7.37)</td>
<td>7.79 (6.71, 8.31)</td>
<td>7.75 (6.9, 8.5)</td>
<td>6.27 (4.93, 7.11)</td>
<td>0.0007</td>
<td>0.1866</td>
<td>0.3771</td>
</tr>
<tr>
<td>Urine creatinine (mg/mL)</td>
<td>0.74 (0.62, 1.48)</td>
<td>0.79 (0.56, 1.15)</td>
<td>1.48 (0.74, 2.5)</td>
<td>1.21 (0.5, 2.07)</td>
<td>0.3344</td>
<td>0.2306</td>
<td>0.4034</td>
</tr>
<tr>
<td>RAIL score</td>
<td>−0.04 (−1.13, 1)</td>
<td>0.69 (0.09, 1.46)</td>
<td>2.02 (1.03, 2.68)</td>
<td>−0.21 (−1.69, 1.12)</td>
<td>0.0027</td>
<td>0.0190</td>
<td>0.8299</td>
</tr>
<tr>
<td>RAIL-Cr score</td>
<td>0.02 (−1.24, 1.01)</td>
<td>0.9 (0.05, 1.45)</td>
<td>1.83 (1.08, 2.43)</td>
<td>−0.34 (−1.5, 0.96)</td>
<td>0.0003</td>
<td>0.0176</td>
<td>0.6849</td>
</tr>
<tr>
<td><strong>Traditional LN measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>122.5 (110, 127.5)</td>
<td>125 (115, 136.5)</td>
<td>122.5 (118.5, 133)</td>
<td>118 (108, 124)</td>
<td>0.2008</td>
<td>0.6556</td>
<td>0.5425</td>
</tr>
<tr>
<td>DBP (mm Hg)</td>
<td>63.5 (59, 77.5)</td>
<td>77.5 (68, 89)</td>
<td>76.5 (66.5, 87)</td>
<td>70 (61, 75)</td>
<td>0.1083</td>
<td>0.3301</td>
<td>0.6946</td>
</tr>
<tr>
<td>eGFR</td>
<td>105.5 (90, 146.5)</td>
<td>97 (49, 128)</td>
<td>82 (59, 133.5)</td>
<td>109 (103, 130)</td>
<td>0.1498</td>
<td>0.4488</td>
<td>0.5008</td>
</tr>
<tr>
<td>UPCR</td>
<td>1.39 (1.05, 3.79)</td>
<td>4.49 (2.44, 9.21)</td>
<td>2.17 (0.95, 3.83)</td>
<td>0.14 (0.09, 0.23)</td>
<td>&lt;0.0001</td>
<td>0.0898</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UACR</td>
<td>0.12 (0.07, 0.19)</td>
<td>0.33 (0.21, 0.48)</td>
<td>0.63 (0.14, 2.56)</td>
<td>0.01 (0.01, 0.02)</td>
<td>&lt;0.0001</td>
<td>0.0191</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>NIH-AI</td>
<td>0.5 (0, 1)</td>
<td>4.5 (4, 7.5)</td>
<td>13.5 (12.5, 15)</td>
<td>–</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIH-CI</td>
<td>0 (0, 0)</td>
<td>1.5 (0, 3)</td>
<td>1 (1, 3)</td>
<td>–</td>
<td>0.0119</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rSLEDAI score</td>
<td>4 (2, 10)</td>
<td>12 (12, 16)</td>
<td>12 (8, 16)</td>
<td>0 (0, 0)</td>
<td>&lt;0.0001</td>
<td>0.0058</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*Values are median (25th, 75th percentile).
†P values are generated from Kruskal-Wallis test.
‡Concentration of urine proteins are natural log transformed.
§Standardisation by urine creatinine.

DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate as per modified Schwartz formula; KIM-1, kidney injury molecule 1; LN, lupus nephritis; MCP-1, monocyte chemotactic protein 1; NGAL, neutrophil gelatinase-associated lipocalin; NIH-AI, National Institute of Health activity index; NIH-CI, National Institute of Health chronicity index; SBP, systolic blood pressure; UACR, albumin-to-creatinine ratio; UPCR, urine protein-to-creatinine ratio.
In this study, the level of ceruloplasmin significantly differed between patients with inactive LN and those without LN (no-LN), despite both having rSLEDAI-scores of 0. Ceruloplasmin is massively excreted in the urine with active LN, but in our prior research levels of urinary ceruloplasmin were not significantly different between the inactive LN group and the no-LN group. Additional studies will be needed to confirm this observation in an independent larger cohort.

In our cohort, we found higher extrarenal SLEDAI scores in patients with active LN compared with no-LN. We are limited by the small sample size of patients with no-LN with high extrarenal SLEDAI, but our previous work showed that the RAIL score is not influenced by extrarenal SLEDAI scores.

Proteinuria and albuminuria are traditional measures of LN activity but may reflect renal damage in addition to renal inflammation. We have shown in the past that isolated traditional biomarkers of LN, including proteinuria cannot identify patients with high disease activity with high accuracy. The current study also expands on our prior observations and on those of other investigators that the rSLEDAI lacks the accuracy to monitor LN activity well. In this context, it seems important to point out that in cSLE, the renal domain score of The British Isles Lupus Assessment Group (BILAG) index does not seem to offer a better clinical alternative. Furthermore, although in this study the UACR seemed to perform somewhat better than the UPCR in capturing LN activity, differences were small and unlikely clinically important.

Given the limitations of the traditional clinical measures of LN activity, that is, the rSLEDAI, the renal domain score of the BILAG, the UPCR and UACR, we propose to consider surveillance of LN to be informed by RAIL scores. Based on the findings of this study, a RAIL score equal to or under −1.6 provides strong support for the absence of LN activity, while scores >2 seem to indicate high LN activity. Scores between these values were common in our study and should raise the suspicion of ongoing mild-to-moderate LN activity. For RAIL scores within this range, active LN is likely, especially in conjunction with higher rSLEDAI scores.

Owing to the predictive ability of the RAIL biomarkers to foresee LN flare, increasing RAIL score should prompt at least close monitoring of the patients. Despite numerically lower RAIL scores in the group of patients without known LN, we confirm our prior studies that the RAIL does not discriminate well patients without LN from those with inactive LN. Nonetheless, this study further demonstrates
the potential of the RAIL score in supporting the surveillance of LN.

The performance of the RAIL and RAIL-Cr in capturing LN activity in this study was not as good as what was reported by us in our prior studies. The exact reasons contributing to this observation are unknown. In our prior research we restricted the time window to 4 weeks and this study we allowed for 3 months. Supported by our exploratory analyses (data not shown), we suspect that the more liberal definition of a urine sample that is collected at the time of biopsy contributed to the inferior performance of the RAIL in this validation study compared with our prior studies. Hence, we provide a conservative estimate of the benefits of measuring the RAIL biomarkers in clinical settings

A strength of our study is the prospective collection of clinical characteristics and standardised laboratory testing which adds to the quality of the results presented. Furthermore, many of the observations in this study are consistent with those of other investigators in larger LN cohorts.

In conclusion, we confirm the validity of the RAIL as a biomarker-based non-invasive measure of LN activity in cSLE with LN. Our results support the benefits of routine measurement of the RAIL biomarkers to monitor LN.

Figure 4 Distribution of renal activity index for lupus (RAIL) (A) and RAIL standardised for urine creatinine (RAIL-Cr) (B) scores in relation to SLE disease activity index (rSLEDAI) scores from 246 samples from patients collected >3 months before/after kidney biopsy. Box-Whisker plots of RAIL score (y-axis, A) or RAIL-Cr (y-axis, B) by rSLEDAI (x-axis) in patients with lupus nephritis (LN) (red colour) and no-LN (blue colour), where medians, IQR, maximum, minimum and outliers are shown. The number of urine samples (N%) for each box is outlined above the x-axis along with the percentage out of a total of 151 LN samples and 95 no-LN samples. Based on data in figure 3, the grey dash lines indicate the RAIL scores corresponding to NIH-Al of 2 and 10.

Author affiliations
1Department of Pediatrics, United Arab Emirates University College of Medicine and Health Sciences, Al Ain, UAE
2Renal Section, Texas Children’s Hospital, Houston, Texas, USA
3Division of Rheumatology, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio, USA
4Division of Biostatistics and Epidemiology, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio, USA
5Division of Nephrology and Hypertension, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio, USA

Competing interests
We thank Rashmi D. Sahay and Chunyan Liu for participation in sample size estimation and an earlier analysis of the data. Qing Ma was essential for assaying the RAIL biomarker and Ms. Theresa Hennard, Catherin Robben and Megan Quinlan-Waters supported the assembly of the materials and data for this study. We are indebted to Ms. Lorie Lyrink for managing the biobanking of the urine samples used for this study. The project described was supported by the National Institutes of Arthritis and Musculoskeletal Skin Diseases under Award - Number F30AR078631. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH

Contributors NA, HB and PD conceptualised the study’s questions and design. NA, SEW, AMa, AMe and SJ planned data extraction and acquired the data. JR and BH performed the statistical analyses and contributed to the interpretation of the results. HB supervised all stages of this study and is the guarantor of this project. All authors provided key contribution to the study and manuscript and approved the final draft of this manuscript.

Funding This work has been supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institutes of Health (NIH) under award numbers P30 AR078631 and P30 AR070549, a Pilot and Feasibility Project from the Pediatric Center of Excellence in Nephrology (NIH P50 DK096418; PD) and the CCTST at the University of Cincinnati is funded by the National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) programme, grant 2UL1 TR001425-05A1.

Competing interests SEW has a consultation agreement with Bristol-Myers Squibb in 2020 for an unrelated project on paediatric lupus nephritis.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study was approved by the CCHMC institutional review board (#2008-0635).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data are available to be shared on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those...
of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**
Najla Aljaberi http://orcid.org/0000-0002-1390-7934
Hermine Brunner http://orcid.org/0000-0001-9478-2987

**REFERENCES**


