Agreement between multi-dimensional and renal-specific response criteria in patients with juvenile systemic lupus erythematosus and renal disease

N. Ruperto[1], A. Bazso[1], A. Pistorio[2], A. Ravelli[1,3], G. Filocamo[1], H.G. Hernandez Huirache[1], A.L. Rodriguez Lozano[1], A.B. Pringe[1], I. Vilca[1], A. Martini[1,3], for the Paediatric Rheumatology International Trials Organisation (PRINTO)

[1]IRCCS G. Gaslini, Pediatria II, Reumatologia, PRINTO, Genova, Italy; [2]IRCCS G. Gaslini, Servizio di Epidemiologia e Biostatistica, Genova, Italy; [3]Dipartimento di Pediatria, Università degli Studi, Genova, Italy.

Abstract

Objective

To evaluate change over time and level of agreement of renal-specific and multi-dimensional measures in juvenile systemic lupus erythematosus (SLE) with renal disease.

Methods

An analysis was made of 205/557 children with baseline 24-hour proteinuria ≥0.5 g. Data were collected at baseline, 6-, 12- and 24-month intervals. Using the Systemic Lupus International Collaborating Clinics (SLICC) renal index (change in proteinuria and urine sediment) as gold standard, responsiveness and discriminative ability analyses were used to identify key renal and multi-dimensional disease activity and damage measures for the evaluation of response to therapy. We also evaluated the kappa agreement between SLICC renal index and PRINTO/ACR juvenile SLE criteria (change in proteinuria, physician and parents evaluations, disease activity, health related quality of life [HRQOL]).

Results

Children with renal disease compared to children without renal disease, had a lower female rate and higher disease activity/response rate (p-values <0.01) but similar damage levels. Large responsiveness (standardised response mean ≥0.8) and statistical significant discriminative ability with the SLICC renal index 4 levels of response (improved, partially improved, stable and worsened) were observed for renal specific measures (proteinuria, urine sediment, renal sub-scores, p<0.0001) and for multi-dimensional variables (disease activity level and physician evaluation p<0.001). Agreement between the SLICC renal index and PRINTO/ACR criteria was moderate (0.57; 95% confidence intervals: 0.44-0.71).

Conclusion

We propose to incorporate multi-dimensional measures (physician and parents' evaluations, disease activity and HRQOL), in addition to renal specific measures, in future clinical trials in juvenile SLE with renal involvement.

Key words

Juvenile systemic lupus erythematosus, core set, response to therapy, disease activity

Nicolino Ruperto, MD, MPH
Anna Bazso, MD
Angela Pistorio, MD, PhD
Angelo Ravelli, MD, Professor
Giovanni Filocamo, MD, PhD
Hayde G. Hernandez Huirache, MD
Ana Luisa Rodriguez Lozano, MD
Alejandra Beatriz Pringe, MD
Iris Vilca, MD
Alberto Martini, MD, Professor

This study was upported by a grant from the European Union (contract no. QLG1-CT-2000-00514) and by IRCCS G. Gaslini, Genoa, Italy; Drs Bazso and Vilca were recipients of EULAR scientific training bursaries; Dr Pringe, was a recipient of a fellowship from the European Union ALFA program (contract no. AML/B7-311/97/0666/II-0246-FI).

Please address correspondence and reprint requests to:
Nicolino Ruperto, MD, MPH,
Paediatric Rheumatology International Trials Organisation (PRINTO),
IRCCS G. Gaslini, Università di Genova,
Pediatria II - Reumatologia,
Largo Gaslini 5,
16147 Genova, Italy.
E-mail:
nicolaruperto@ospedale-gaslini.ge.it

Received on August 13, 2009; accepted in revised form on January 18, 2010.

Competing interests: none declared.

Introduction

Systemic lupus erythematosus (SLE) is a disease characterised by protean clinical manifestation and unpredictable course. Renal involvement is one of the most important and severe clinical manifestations and has been the target of several clinical trials in both adult (1-13) and juvenile onset SLE (3).

All trials conducted in SLE nephritis have used different organ-specific renal criteria to evaluate response to therapy, making it difficult the comparisons of their results. The lack of standardised and validated criteria for the assessment of response to therapy in SLE with renal disease prompted the American College of Rheumatology (ACR) and other to propose organ-specific criteria for proliferative and membranous renal disease in SLE clinical trials (14, 15). More recently, the Systemic Lupus International Collaborating Clinics (SLICC) has proposed a renal activity/response index (SLICC renal index) for the evaluation of response to therapy in SLE with kidney involvement (16, 17). The SLICC renal index is based on the evaluation of the change over time of a renal score which comprises 24-hour proteinuria and urine sediment (red and white blood cells); measures included in the score have been selected through regression analysis using as gold standard the evaluation of a panel of experts.

In the paediatric field, the Paediatric Rheumatology International Trials Organisation (PRINTO) (18), supported by a grant from the European Union, undertook a multinational effort, based on a large-scale prospective data collection and consensus evaluation as gold standard, that was aimed at developing and validating measures and criteria to evaluate response to therapy in juvenile SLE; these are now known as the PRINTO core set (19, 20) and the PRINTO/ACR criteria (21) respectively. In the PRINTO core set there is one measure specific for the evaluation of renal disease (24-hour proteinuria) and, in addition to 4 multi-dimensional assessments (physician's global assessment of disease activity, global disease activity index, parent's global assessment of the overall child's well-being and health-related quality of life [HR-QOL]) that are relevant for all juvenile SLE subtypes (*e.g.* neurologic, haematologic, etc.) but that also contain specific parameters for patients with renal involvement (*e.g.* the renal sub-scores of the disease activity tools). The inclusion of 24-hour proteinuria in the PRINTO juvenile SLE core set was thought to be important since it is the most severe subtype of juvenile SLE is renal disease (22-24). According to the PRINTO/ACR criteria patients are classified as responders if they demonstrate at least 50% improvement from baseline in any 2 of the 5 PRINTO core set measures, with no more than 1 of the remaining worsening by more than 30%. The approach chosen by PRINTO was therefore to develop multi-dimensional criteria (20, 25-29), rather than organ specific criteria (*e.g.* renal, neuropsychiatric, haematological). However it is not clear if a similar approach can be used in a protean disease like SLE since the advantages/disadvantages of using organ-specific or multi-dimensional criteria in SLE nephritis have not yet been tested.

The goal of this project was to evaluate the change over time (up to 2 years of follow-up) of renal-specific and multi-dimensional measures in children with juvenile SLE with renal disease. A secondary goal was to evaluate the level of agreement between multi-dimensional criteria (PRINTO/ACR criteria) and renal specific criteria (SLICC renal index). Our overall hypothesis was that multi-dimensional criteria are valid alternatives to renal-specific criteria for the evaluation of response to therapy in children with juvenile SLE with renal disease.

Patients and methods

The PRINTO juvenile SLE database contains data related to 557 patients consecutively enrolled who: 1) had SLE by the 1997 revised ACR classification criteria (30, 31); 2) were under 18 years of age; and 3) were in an active phase of their disease, defined as the either the need to start corticosteroid therapy and/or a new immunosuppressive medication or to undergo a major increase in the dosage of ongoing corticosteroid and/or immunosuppressive drugs. These inclusion criteria were chosen to

select patients with recent disease onset or disease flare in order to try to closely mimic a clinical trial where only patients with active disease, either at onset or during flare, are enrolled.

For the purposes of this analysis we included the subgroup of children with renal disease, defined by the presence of a baseline 24-hour proteinuria ≥0.5 g. We did not collect random spot urinary protein:creatinine ratio whose use in clinical trials is considered unreliable in patients with high protein excretion (32, 33).

All patients were assessed at baseline, 6, 12 and 24 months thereafter, as previously described (20), for demographic, clinical (physician's global assessment of the child's overall disease activity on a 0 to 10 cm visual analogue scale [VAS]), the renal sub-score (only items related to renal parameters) and the total score of 3 disease activity tools (the Systemic Lupus Erythematosus Disease Activity Index [SLEDAI], the Systemic Lupus Activity Measures [SLAM], and the European Consensus Lupus Activity Measurement [ECLAM]), (34-37) all valid alternatives to the British Isles Lupus Assessment Group (BILAG) index (38), blood pressure values (one reading per patient with appropriate cuff size) standardised according to gender, age and height (39), standardised (20) laboratory tests (serum complement fractions C_3 and C_4, complete blood cell count, Westergren erythrocyte sedimentation rate [ESR], glomerular filtration rate [GFR] according to the Schwartz formula (14, 40)), parent's assessments (child's overall well-being on a 0 to 10 cm VAS, and HRQOL through the physical summary score [PhS] of the parent version of the Child Health Questionnaire [CHQ]) (41, 42). We also collected damage data through the renal sub-score (only items related to renal parameters) and the total score of the SLICC damage index (DI) and the physician's global assessment of the patient's overall disease damage on a 0 to 10 cm VAS (43). Renal biopsy classes were not available in this dataset. Of note, the data at 12, and 24 months have been analysed and reported for the first time in this paper.

In the rest of this paper we will refer to measures to indicate the individual variables used to evaluate response (*e.g.* 24-hour proteinuria, urine sediment) and criteria for the related definitions (*e.g.* SLICC renal index, PRINTO/ACR juvenile SLE criteria).

Evaluation of response criteria in SLE trials in patients with renal disease

Several measures and criteria to evaluate response to therapy in SLE are reported in the literature, but none of them has been formally validated with prospectively collected *ad hoc* data and no agreement exists about which criteria performs best in the context of clinical trials. All criteria retrieved were tested in the PRINTO dataset. Since it was not possible to test exactly all the measures as reported in the literature, in some cases we have modified the original version (for example, urine protein:creatinine ratio was substituted with 24-hour proteinuria, serum albumin was not available), as detailed in Table I. Each patient was classified as "improved" or "not improved" according to different criteria reported in the literature. The time frame used in most of the clinical trials in the literature was 6 months.

Choice of the gold standard

In order to compare the statistical properties of the different variables (renal-specific and multi-dimensional measures) present in the PRINTO dataset, there was the need to choose a proper standard against which to evaluate the measures and criteria tested. The SLICC renal index (16, 17), the ACR criteria (14), and the PRINTO/ACR juvenile SLE criteria (19-21) used as gold standard a consensus panel evaluation. We chose as gold standard for this work the SLICC renal index which identifies 4 level of mutually exclusive response: complete response (if the SLICC renal score is >0 at baseline and equal to 0 at follow-up), partial response (if the baseline score is greater than the follow-up score, but the follow-up score is not equal to 0), stable (if the follow-up score is equal to the baseline score), and worsening (if the follow-up score is greater than the baseline score). The SLICC renal score ranges from 0 to 15 and was computed as follows: 24/hour proteinuria 0.5–1 g (3 points), >1–3 g (5 points), >3 g (11 points); urine red blood cell (RBC) count >10/high power field (hpf) (3 points) and urine white blood cell (WBC) count >10/hpf (1 point) (score range 0–15). Since in the PRINTO dataset urine sediment was available as absent or >5 RBC or >5 WBC, (14), we attributed a score of 3 if RBC were >5/hpf and 1 point if the WBC were >5/hpf.

The PRINTO/ACR juvenile SLE criteria was used as prototype of multi-dimensional criteria; it classifies patients into 2 mutually exclusive categories of responders and non responders as detailed in the introduction.

The statistical comparison of the measures tested proceeded in subsequent steps, as described below.

Change over time and responsiveness

The statistical properties of all measures (renal and multi-dimensional) used to evaluate renal disease in the PRINTO database included: feasibility or practicality of the measures, determined by evaluating the percentage of missing values; face and content validity, based on the results of a previous consensus conference (19); responsiveness, determined by measuring the ability of each variable to detect clinically important change between baseline and 6, 12 and 24 months through the standardised response mean (SRM) (44, 45) calculated as the absolute mean change in score divided by the standard deviation (SD) of that score. An SRM value <0.5 was considered small, ≥0.5<0.8 moderate, and values ≥0.8 representing a large effect (46, 47).

Discriminative ability of the measures used to evaluate response to treatment

We evaluated the relationship between the absolute changes of the measures analysed and the SLICC renal index 4 levels of response as gold standard. The purpose of this analysis was to evaluate if the measures (renal and multi-dimensional) were able to discriminate patients who improved from patients who did not.

Table I. List of criteria for the evaluation of response to therapy tested.

Original criteria Short name (year publication). Description (references)	Criteria tested with modification in ***bold italics***	Response rate n. (%)
RENAL-SPECIFIC CRITERIA SLICC renal activity/response index (2008): 0.5-1 g/24 hour proteinuria (3 points), >1-3 g/24 hour proteinuria (5 points), >3 g/24 hour proteinuria (11 points), urine RBC >10/hpf (3 points), and urine WBC >10/hpf (1 point) (16, 17) (15, 16)	Same as original but urine **RBC >5/hpf** (3 points), and urine **WBC >5/hpf** (1 point)	171/203 (84.2%)
ACR GFR 25 (2006). 25% increase if baseline GFR is abnormal (14)	Same as original	97/180 (53.9%)
ACR proteinuria 50 (2006). ≥50% reduction in the urinary protein: urinary creatinine ratio (14)	***≥50% reduction in 24 hrs proteinuria***	131/180 (72.8%)
ACR sediment (2006). Change from active to inactive sediment (14)	Same as original	99/179 (55.3%)
ACR remission (2006). urinary protein: urinary creatinine ratio of < 0.2 and GFR 90 and inactive urinary sediment (14)	***urine protein <0.5 g/24h*** and GFR (Schwartz) <90 and inactive sediment	87/180 (48.3%)
Remission Illei (2007). <1 g/24 hour proteinuria, inactive sediment (<10 RBC/hpf and no cellular casts) and stable serum creatinine (1)	<1 g/24 hour proteinuria and ***inactive urinary sediment (14)*** and stable serum creatinine	114/180 (63.3%)
Remission Baca (2006), Illei (2001). <1 g/24 hour proteinuria and < 10 RBC/hpf, 0 cellular casts (3, 7)	<1 g/24 hour proteinuria and ***inactive urinary sediment (14)***	111/178 (62.4%)
Remission Grootscholten (2007). <0.5 g/24 hour proteinuria, <10 RBC/hpf, serum creatinine < 130% of the lowest serum creatinine (2)	<0.5 g/24 hour proteinuria, ***inactive sediment (14)***, serum creatinine < 130% of the lowest serum creatinine	68/183 (37.2%)
Remission Urowitz (2007). No RBC casts or hemegranular casts, hematuria or pyuria in the absence of other causes, or proteinuria (<0.5 g 24 hrs proteinuria,or ≥3+ on dipstick), or an abnormal renal biopsy with active lupus nephritis (55)	Same as original with ***inactive urinary sediment as per ref (14)***. Kidney biopsy not available	75/180 (41.7%)
Remission Ginzler (2005). Return to within 10% of normal values of serum creatinine, proteinuria and urine sediment (4)	Same as original with ***inactive urinary sediment as per ref (14)***	57/183 (31.1%)
Remission partial Ginzler (2005). 50% improvement in abnormal serum creatinine, proteinuria and urine sediment, without worsening (within 10%) of any measurement (4)	Same as original with ***inactive urinary sediment as per ref (14)***	80/180 (44.4%)
Remission Contreras (2004). Decrease in the UP:UC ratio to <3 in pts with baseline nephrotic range proteinuria (UP:UC ration ≥3) or by 50% in patients with sub-nephrotic proteinuria accompanied by either an improvement in baseline serum creatinine level ≥25% or a stable serum creatinine level that was within 25% of baseline (6)	Same as original ***with 24 hour proteinuria*** instead of urinary protein:creatinine ratio	161/183 (88%)
Remission Chan (2000). <0.3 g/24 hour proteinuria, normal urinary sediment, normal serum albumin, and values for both serum creatinine and GFR that were worsened <15% (5;8)	Same as original ***without albumin***	32/182 (17.6%)
Remission Gourley (1996). <1 g/24 hour proteinuria <10 dysmorphic RBC/hpf, 0 cellular casts, and without doubling of serum creatinine (9)	<1 g/24 hour proteinuria, ***inactive urinary sediment (14)*** without doubling of serum creatinine	87/181 (48.1%)
Euro-lupus (2002). Absence of primary response (based on change in serum creatinine, 24-hour proteinuria, albumin) or glucorticoid resistant flare or serum creatinine doubling(10;11) Note See table I in ref (10) for more details	Same as original ***without albumin and glucorticoid resistant flare***	139/178 (78.1%)
Creatinine doubling (1983-2007). Patients without doubling of serum creatinine considered as responders (2, 6, 12, 13, 51)	Same as original	195/202 (96.5%)*
ESRD (2002-2004). Patients with d ESRD (6;10;52) considered as responders	Same as original	199/202 (98.6%)*
MULTI-DIMENSIONAL CRITERIA PRINTO/ACR 50. 2 of any 5 improved ≥ 50%, max 1 worsened by >30% (19-21)	Same as original	167/205 (81.5%)
PRINTO/ACR 50. 2 of any 5 improved ≥ 50%, max 1 worsened by > 30% (19-21)	With ECLAM sub-score instead of ECLAM total score and proteinuria could not worsen	166/205 (81%)
PRINTO/ACR 50. 2 of any 5 improved ≥ 50%, max 1 worsened by > 30% (19-21)	Proteinuria could not worsen	162/205 (79%)
Physician. Attending physician evaluation of response to therapy (improved, stable, worsened) (19-21)	Same as original	144/179 (80.4%)
Parents. Parent evaluation of response to therapy(improved, stable, worsened) (19-21)	Same as original	147/177 (83.1%)

RBC: red blood cells, HPF high power field; UP/UC ratio: urinary protein:urinary creatinine ratio; ESRD: end stage renal disease.
*Number of patients who had doubling of serum creatinine and ESRD in the initial 6 months (7 patients had doubling of serum creatinine, 3 patients had ESRD).

Agreement multi-dimensional and renal-specific criteria

Agreement between the SLICC renal index (as gold standard) and the PRINTO/ACR juvenile SLE criteria was evaluated by means of the kappa statistics (48, 49), with the following cut-offs proposed by Landis & Koch (50): 0.01–0.2 = slight; 0.21–0.4 = fair; 0.41–0.6 = moderate; 0.61–0.8 = substantial; 0.81–1 = almost perfect agreement. We also re-evaluated 2 modified versions of PRINTO/ACR criteria (21); in the first modification we added the contingency that 24-hour proteinuria could not worsen (*e.g.* a patient with worsening in proteinuria but with improvement in the remaining 4 PRINTO core set variables is counted as not improved), and in the second we also substitute the ECLAM total score with ECLAM renal sub-score. The purpose of this analysis was to see if the PRINTO/ACR, as prototype of multi-dimensional criteria, could be an alternative to criteria based only on renal parameters. We expected a fair to moderate correlation between the 2 types of indexes since the construct that underlines the evaluation is different (multi-dimensional evaluations in the PRINTO/ACR criteria and renal parameters in the SLICC renal index).

Statistics

The data were reported as means and SD, medians (1^{st}–3^{rd} quartile) or absolute numbers and percentages. All comparisons at follow-up were made with respect to baseline. *P*-values refer to Wilcoxon's test for paired data, Kruskal-Wallis for continuous data and Chi-square/Fisher exact tests for categorical data as appropriate. In order to avoid multiple comparisons' errors, all p values were adjusted according to Bonferroni's correction.

The data were entered in an Access XP database and analysed with Excel XP (Microsoft), XLSTAT 6.1.9 Addinsoft, Statistica 6.0 (StatSoft, Inc), and Stata 7.0 (Stata Corporation).

Results

Table I reports the list of criteria derived from the literature and schematically divided into 2 main categories. The first group can be collectively named as renal-specific criteria (also called organ-specific) based on change of key specific measures used to evaluate renal disease; the prototype of these criteria are the SLICC renal index (16, 17) and the ACR criteria (14). The second group refers to multi-dimensional evaluations by the physicians, the parents or combination of different multi-dimensional domains whose prototype is the PRINTO/ACR juvenile SLE criteria (19-21).

The criteria/measures tested and their original versions are reported in Table I. The last column shows the frequency of response of the measures/criteria tested in the PRINTO dataset. There was a wide variability in the response rate, with positive response criteria ranging from 17.6% of remission criteria by Chan *et al*. (5, 8) to 88% of the remission criteria by Contreras *et al*. (6) whereas negative response criteria (creatinine doubling or ESRD) (2, 6, 10, 12, 13, 51, 52) identified >95% of patients as responders (6, 10, 52).

Demographic and clinical characteristics

Of the 577 children with juvenile SLE in the database, 215 with proteinuria ≥0.5 g 24-hour at baseline were identified. Of these 215, 10 patients were excluded from the analysis: 6 because they were lost to follow-up and 4 were deceased before the first 6-month follow-up (1 for severe multi-organ failure, 1 for severe pancreatic and disseminated intravascular coagulation, 1 with sepsis, 1 for unknown reasons). The final sample available for the analysis was, therefore, 205/557 (37%).

Table II shows the comparison of the demographic features at baseline between the 205 patients with renal disease included in the present analysis and the 352 patients without renal disease who were excluded. Children with renal disease at baseline had a lower female rate (75.1 *vs*. 86.6%), and an higher number of ACR criteria at diagnosis (6 *vs*. 5). Patients with renal disease also showed a higher level of disease activity, as documented by the median physician's global assessment of disease activity and the ECLAM, the SLEDAI or the SLAM. The response rates according to PRINTO ACR 50, 70, 90 or 100 criteria were statistically significantly higher for children with renal disease when compared to children without renal disease at baseline. Of the 167 PRINTO/ACR 50 responders 142 (85%) had proteinuria improved by at least 50%. There were no differences in the 2 cohorts for age at onset/study entry, disease duration, the parameters assessed by parents or for the damage level.

Change over time and responsiveness

Table III shows the descriptive characteristics, at baseline, 6, 12 and 24-month, for the measure analysed. The percentage of missing data was uniformly less than 10%, demonstrating that most variables had good feasibility (data not shown).

At baseline children have high median values for all renal parameters. The SLICC renal score was >0 for all 205 children, 24-hour proteinuria was >1 g in 133 (65%) and in nephrotic range (≥3.5 g/day) in 49 (24%), abnormal serum creatinine (>1.2 mg/dl) was observed in 28 (14%) with 10 (5%) patients having values greater than 2, abnormal GFR (Schwartz formula) <90 ml/min was observed in 55 (27%). Systolic and diastolic blood pressure were higher than the 95^{th} percentile in 28% of the patients (see also standardised z-scores in Table). Similar high baseline median values were present for the multi-dimensional measures by the physicians and the parents, and for the total scores of 3 disease activity tools the SLEDAI, SLAM and ECLAM; 108/190 (57%) children had poor HRQOL has defined by a CHQ physical well being (PhS) score <30 (2 SD below the mean of healthy controls). Damage levels were unremarkable as expressed by the low median values of the renal sub-score and total score of the SLICC DI (156/200, 78% with a score equal to 0) and the physician's global assessment of patient's overall disease damage.

At the 6-month follow-up a statistically significant change was observed for all measures. The SLICC renal score was >0 in 121 (60%) children, 24-hour proteinuria in nephrotic range in 17 (8%),

Table II. Comparison of the 205 patients with renal disease (24-hour proteinuria ≥ 0.5 g at baseline) versus the 352 excluded patients. Values are medians and first and third quartiles (1st; 3rd Q) unless otherwise indicated.

Variables (min-max) ↑ higher worse; ↓ lower worse	n.	Children with renal disease median (1st; 3rd Q)	n.	Children without renal disease median (1st; 3rd Q)	*p*-values
Female - n. (%)	199	154 (75.1%)	352	309 (86.6%)	0.0006
Age at onset (years)	205	12.4 (10.1; 14.4)	352	12.2 (10.2; 13.9)	0.77
Age at study entry (month 0) (years)	205	14.3 (11.9; 15.9)	352	13.9 (12.0; 15.6)	0.16
Disease duration (years)	205	0.6 (0.2; 2.6)	352	0.5 (0.2; 2.2)	0.43
Number of ACR criteria at diagnosis	205	6.0 (5.0; 7.0)	352	5.0 (4.0; 6.0)	0.0006
SLEDAI total score (0-105 score) ↑	204	22 (16; 28.5)	351	14 (9; 19)	<0.0001
SLAM total score (0-90 score) ↑	205	17 (11; 22)	352	14 (10; 19)	0.0325
**ECLAM total score(0-10 score) ↑*	205	7.0 (5.0; 9.0)	352	5.0 (4.0; 7.0)	<0.0001
**Physician's global assessment of the patient's overall disease activity (0-10 cm) ↑*	205	6.4 (5.0; 8.2)	350	5.4 (3.0; 7.8)	0.0003
**Parent's global assessment of the overall child's well-being (0-10 cm) ↑*	193	4.6 (2.2; 7.2)	332	4.3 (1.3; 6.7)	0.31
**CHQ Physical summary score (PhS) (40-60 score) ↓*	190	27.6 (10.8; 43.0)	307	32.1 (15.0; 43.4)	0.05
SLICC DI total score (0-49) ↑	200	0 (0; 1)	350	0 (0; 1)	0.40
Physician's global assessment of the patient's overall disease damage (0-10 cm) (0-10 cm) ↑	198	0 (0; 1.2)	345	0.1 (0; 1.1)	0.75
PRINTO/ACR 50 criteria	205	167 (81.5%)	352	251 (71.3%)	0.008
PRINTO/ACR 70 criteria	205	143 (69.8%)	352	184 (52.3%)	<0.0001
PRINTO/ACR 90 criteria	205	95 (46.3%)	352	96 (27.3%)	<0.0001
PRINTO/ACR 100 criteria	205	55 (26.8%)	352	57 (16.2%)	0.003

*Measures in italics are in included in the PRINTO core set for the evaluation of response in juvenile SLE (20).

serum creatinine (>2 mg/dl) in 6 (3%) patients, abnormal GFR (<90 ml/min) in 33 (17%). A total of 3/197 (2%) had a CHQ physical well being (PhS) score <30 (2 SD below the mean of healthy controls). Good responsiveness to clinical change, with a large SRM (≥0.8) was demonstrated, among the renal measures, by the SLICC renal score, by the 3 renal sub-score of the SLEDAI, SLAM and ECLAM and by C_3 while responsiveness was poor or moderate for the remaining renal parameters (24-hour proteinuria, C_4, GFR, serum creatinine, blood pressure and SLICC DI renal subscore). The SRM of the renal parameters rose when we restricted the analysis to the subgroup of patients with greater renal impairment at baseline (data not shown). All the remaining multi-dimensional measures showed SRM ≥ 1 with the highest values, observed for the total score of the SLEDAI, ECLAM, SLAM and physician's evaluation of disease activity (range 1.36-1.46). Poor responsiveness was showed by the SLICC DI total score and the physician's global assessment of patient's overall disease damage.

At 12 and 24 months 6/133 (4.5%) and 5/123 (4%) patients had serum creatinine >2 mg/dl and 4/138 (2.9%) and 4/128 (3.1%) patients were in end-stage renal disease. A greater increase with higher SRM values was observed at 12 months where also 24-hour proteinuria showed large SRM (0.89). After 24 months SRM values were similar to the values observed after 12 months of treatment confirming that most of the effect of treatments was reached at 6-12 months interval.

Discriminative ability of the measures used to evaluate response to treatment

Table IV shows the relationships between the median absolute changes in renal and multi-dimensional measures and the 4 categories of response (improved, partially improved, stable and worsened) of the SLICC renal index. Among the renal measures, the SLICC renal score, 24-hour proteinuria, urine sediment, the renal sub-score of the SLEDAI, SLAM and ECLAM and, to a lower extent, C_3 were able to discriminate between the 4 response levels of the SLICC renal index. Similarly, among the additional measures the median total score of the SLEDAI, SLAM, ECLAM, the physician's evaluation of disease activity, and the SLICC DI total score were also able to discriminate between the 4 level of response. All remaining measures (serum creatinine, GFR, C_4, parent's evaluations, SLICC DI renal subscore and physician's global assessment of patient's overall disease damage) were not able to discriminate among the 4 levels of response of the SLICC renal index (*p*-values not significant).

Agreement between multi-dimensional and renal specific criteria

The Cohen's kappa level of agreement in the evaluation of response between the SLICC renal index used as the gold standard, and the original version of the PRINTO/ACR juvenile SLE criteria was in the moderate range being 0.45 (95% CI: 0.31–0.59). When we modified the PRINTO/ACR criteria (Table I) by adding the contingency that proteinuria could not worsen, the agreement rose to 0.53 (95% CI: 0.39–0.66) and to 0.57 (95% CI: 0.44–0.71) when we also substitute the total score of the ECLAM with the renal sub-score.

Discussion

Using a data-driven approach in a large paediatric dataset, we compared the statistical properties of renal-specific and multi-dimensional measures and criteria in the evaluation of response to

Table III. Clinical and laboratory data at baseline, after 6, 12 and 24-month. The SLICC renal score is based on 24-hour proteinuria and urine sediment (see methods for details). Data are reported as median (1st; 3rd Q) for continuous variables or numbers (%) for categorical measures. The SRM are reported only for continuous measures. Blood pressure values were standardized according to gender, age and height (39) and reported as z-scores and mean (SD) percentiles values.

Measures (min-max) ↑ higher worse; ↓ lower worse Mean (SD) or n (%)	Baseline median (1st; 3rd Q) n=205	Month 6 median (1st; 3rd Q) n=204	Month 6 SRM	Month 12 median (1st; 3rd Q) n=138	Month 12 SRM	Month 24 median (1st; 3rd Q) n=129	Month 24 SRM
Renal measures							
SLICC renal score (score 0-15)↑	8 (6; 11)	3 (0; 5)*	1.13	3 (0; 3)*	1.59	0 (0; 3)*	1.37
**24-hour proteinuria (g/24 h)*↑*	1.5 (0.8; 3.3)	0.3 (0.1; 1)*	0.62	0.2 (0; 0.5)*	0.89	0.1 (0; 0.5)*	0.76
>5 RBC/hpf↑ n (%)	155 (75.6%)	72 (35.5%)*		28 (20.4%)*		22 (17.3%)*	
>5 WBC/hpf↑ n (%)	71 (34.8%)	36 (17.7%)*		6 (4.4%)*		13 (10.3%)*	
>1 cell cast↑ n (%)	117 (57.1%)	35 (17.2%)*		28 (20.4%)*		10 (7.9%)*	
Active sediment (>5 RBC/hpf and >5 WBC or >1 cell cast) ↑ n (%)	135 (66.2%)	45 (22.2%)*		15 (11%)*		13 (10.3%)*	
Serum creatinine (0.6-1.2 mg/dl) ↑	0.8 (0.6; 1)	0.7 (0.6; 0.9)§	0.19	0.7 (0.6; 0.9)	0.05	0.8 (0.6; 0.9)§	0.24
GFR (Schwartz equation)↓	110 (87; 134)	121 (101; 138)§	0.23	119 (97; 135)	0.17	114 (92; 134)	0.16
C_3 (0.7-1.6 g/L) ↓	0.4 (0.1; 0.7)	0.8 (0.6; 1.1)*	0.84	0.9 (0.7; 1.1)*	0.95	0.9 (0.7; 1.1)*	1.07
C_4 (0.2-0.4 g/L) ↓	0.2 (0.1; 0.2)	0.2 (0.2; 0.3)*	0.49	0.2 (0.2; 0.3)*	0.50	0.2 (0.2; 0.3)#	0.44
SLEDAI renal sub-score (0-16 score) ↑	12 (8; 12)	0 (0; 8)*	1.13	0 (0; 4)*	1.47	0 (0; 4)*	1.42
SLAM renal sub-score (0-3 score) ↑	2 (2; 3)	1 (0; 2)*	0.84	0 (0; 1)*	1.12	0 (0; 1)*	1.18
ECLAM renal sub-score (0-2 score) ↑	2 (0.5; 2)	0.5 (0; 0.5)*	0.90	0 (0; 0.5)*	1.14	0 (0; 0.5)*	1.06
Z-score systolic blood pressure	0.7 (-0.2; 1.8)	0.3 (-0.3; 1.1)*	0.29	-0.1 (-0.7; 0.9)*	0.37	0.1 (-0.7; 0.8)*	0.38
Percentile systolic blood pressure	76.7 (42.8; 96)	63.3 (37.1-87.4)**	0.2	47 (25.6-82.8)*	0.36	53.8 (23.2; 77.6)*	0.34
Z-score systolic blood pressure	1 (0.1; 1.8)	0.7 (0.2; 1.4)**	0.19	0.7 (0.1; 1.3)**	0.26	0.5 (0; 1.2)*	0.37
Percentile systolic blood pressure	83.9 (56; 96.5)	76 (58.5; 92.2)	0.07	75.7 (54.5; 90.7)	0.17	68.9 (51.3; 88.1)**	0.3
SLICC DI renal sub-score (0-5 score) ↑	0 (0; 0)	0 (0; 0)	0.13	0 (0; 0)	0.10	0 (0; 0)	0.09
Multi-dimensional measures							
SLEDAI total score (0-105 score) ↑	22 (16; 28.5)	6 (2; 10)*	1.46	2.5 (0; 6)*	1.72	2 (0; 6)*	1.59
SLAM total score (0-90 score) ↑	17 (11; 22)	4 (2; 6)*	1.36	2 (0; 5)*	1.63	2 (0; 4)*	1.69
**ECLAM total score(0-10 score)*↑	7 (5; 9)	2 (1; 3)*	1.38	2 (0; 3)*	1.66	1 (0; 2)*	1.68
**Physician's global assessment of the patient's overall disease activity (0-10 cm)*↑	6.4 (5; 8.2)	1 (0.3; 2.9)*	1.44	0.4 (0; 2)*	1.63	0.2 (0; 1.4)	1.68
**Parent's global assessment of the overall child's well-being (0-10 cm)*↑	4.6 (2.2; 7.2)	0.4 (0; 1.9)*	1.00	0.1 (0; 1.4)*	0.91	0 (0; 1.2)*	1.04
**CHQ Physical summary score (PhS) (40-60 score)*↓	27.6 (10.8; 43)	49.2 (39.8; 53.9)*	1.04	50.3 (41.7; 54.5)*	1.02	51.6 (44.7; 54.2)*	1.05
SLICC DI total score (0-49) ↑	0 (0; 1)	0 (0; 1)	0.09	0 (0; 1)	0.18	0 (0; 1)	0.09
Physician's global assessment of the patient's overall disease damage (0-10 cm) (0-10 cm) ↑	0 (0; 1.2)	0 (0; 0.7)§	0.25	0 (0; 0.4)	0.23	0 (0; 0.6)	0.17

*Measures in italics are in included in the PRINTO core set for the evaluation of response in juvenile SLE (20). SRM value <0.5 are considered small, ≥0.5 <0.8 moderate, and values ≥0.8 representing large effect (46, 47).

P-values refer to Wilcoxon's test or Chi-square with Bonferroni's correction. All comparisons were made with respect to baseline; *p<0.0001; #p<0.001: §p<0.05.

therapy in children with juvenile SLE nephritis. Multi-dimensional indexes, like the PRINTO/ACR criteria, which includes the evaluation of 24-hour proteinuria, showed a moderate correlation with organ specific criteria, like the SLICC renal index.

When we applied to the PRINTO juvenile SLE dataset the criteria published in the literature (1, 2, 2-6, 6-14, 16, 17, 19-21) the response rate varied greatly, from 17.6% to 98.6%, with criteria based on serum creatinine doubling, or ESRD identifying almost all patients as improved since very few children had elevated serum creatinine, or ESRD even at 2-year follow-up.

All measures specific for the evaluation of renal disease, showed a statistical significant change at 6, 12 and 24 months from baseline. More specifically, the largest responsiveness at 6-month was shown by several renal specific measures (e.g. the SLICC renal score, the 3 renal sub-scores of the SLEDAI, SLAM and ECLAM, C_3) but also by the other 4 domains included in the PRINTO juvenile core set (physician's subjective estimation of the level of disease activity, global disease activity scoring either with SLEDAI, ECLAM or SLAM, parent's global assessment of the overall patient's well-being, and HRQOL). In addition, an increased responsiveness was obtained for data analysed at 12 and 24-month. Similar results were obtained when we evaluated the ability of renal specific and multi-dimensional measures, to discriminate between the patients who improved or not based on the SLICC renal index level of response. Damage levels, as measured by the renal sub-score and total score of the SLICC DI and the physician's global assessment of patient's overall disease damage, showed on the contrary poor responsiveness reflecting the relatively short disease duration of this pediatric cohort with renal involvement.

Until now, organ-specific criteria like the SLICC renal index, have been

Table IV. Relationships between median (1st; 3rd Q) absolute changes or frequencies for renal and multi-dimensional measures and the SLICC renal index at 6-month follow-up. Patients were divided into the 4 mutually exclusive categories of the SLICC renal index as improved, partially improved, stable or worsened (see methods). Sample equal to 203 patients.

	SLICC renal index Improvement n=82 (40%)	SLICC renal index Partial Improvement n=89 (44%)	SLICC renal index Stable n=21 (10%)	SLICC renal index Worsening n=11 (6%)	*p*-value
Renal measures					
SLICC renal score (score 0-15)↑	-7 (-9; -5)	-5 (-6; -3)	0 (0; 0)	5 (2; 8)	<0.0001
**24-hour proteinuria (g/24 h)* ↑	-1 (-2.3; -0.6)	-1.1 (-2.1; -0.5)	-0.1 (-1.7; 0.02)	1.9 (0.9; 2.3)	<0.0001
>5 RBC/hpf↑ improved/positive patients n (%)	56/56 (100%)	30/75 (40.0%)	0/13 (0%)	3/7 (42.9%)	<0.0001
>5 WBC/hpf↑ improved/positive patients n (%)	16/16 (100%)	23/44 (52.3%)	0/3 (0%)	3/5 (60.0%)	<0.0001
>1 cell cast↑ improved/positive patients n (%)	35/36 (97.2%)	45/63 (71.4%)	3/10 (30%)	3/7 (42.9%)	<0.0001
Active sediment (>5 RBC/hpf and >5 WBC or >1 cell cast) no. n (%)↑ improved/positive patients (%)	39/40 (97.5%)	47/73 (64.4%)	3/10 (30%)	3/8 (37.5%)	<0.0001
Serum creatinine (0.6-1.2 mg/dl) ↑	-0.03 (-0.2; 0.1)	-0.1 (-0.2; 0.04)	0 (-0.1; 0.1)	-0.03 (-0.4; 0.2)	0.59
GFR (Schwartz equation)↓	5 (-14; 33)	10.7 (-5.6; 34.8)	0 (-9.6; 13.8)	4.6 (-24.9; 47.5)	0.43
C_3 (0.7-1.6 g/L) ↓	0.4 (0.1; 0.7)	0.4 (0.04; 0.7)	0.3 (0.01; 0.8)	0.01 (-0.2; 0.2)	0.016
C_4 (0.2-0.4 g/L) ↓	0.04 (0; 0.1)	0.03 (0; 0.1)	0.02 (0; 0.1)	0.01 (-0.01; 0.1)	0.64
SLEDAI renal sub-score (0-16 score) ↑	-8 (-12; -4)	-4 (-8; -4)	0 (0; 0)	0 (-4; 8)	<0.0001
SLAM renal sub-score (0-3 score) ↑	-2 (-2; -1)	-1 (-1; 0)	0 (-1; 0)	1 (0; 1)	<0.0001
ECLAM renal sub-score (0-2 score) ↑	-0.5 (-2; -0.5)	-1.5 (-1.5; 0)	0 (-1.5; 0)	0 (0; 0.5)	<0.0001
SLICC DI renal sub-score (0-5 score) ↑	0 (0; 0)	0 (0; 0)	0 (0; 0)	0 (0; 1)	0.18
Multi-dimensional measures					
SLEDAI total score (0-105 score) ↑	-16 (-24; -10)	-16 (-23; -10)	-4 (-10; 0)	-8 (-18; -1)	<0.0001
SLAM total score (0-90 score) ↑	-12 (-17; -7)	-12 (-18; -6)	-3 (-7; -1)	-7 (-18; -1)	0.0002
**ECLAM total score(0-10 score)* ↑	-5 (-7; -3)	-5 (-7; -2)	-2 (-4; 0)	0 (-3; 0)	<0.0001
**Physician's global assessment of the patient's overall disease activity (0-10 cm)* ↑	-5.2 (-7.2; -2.4)	-4.5 (-6.6; -2.8)	-1 (-3.7; -0.5)	-1.2 (-6; -0.3)	0.0001
**Parent's global assessment of the overall child's well-being (0-10 cm)* ↑	-2.2 (-5.1; -0.1)	-3.9 (-6; -0.8)	-0.6 (-4.5; 0.2)	-3.7 (-6.3; -0.7)	0.11
**CHQ Physical summary score (PhS) (40-60 score)* ↓	17.5 (4.3; 34)	17.6 (4.2; 34.8)	5.8 (0.8; 17.1)	12.2 (0.7; 28.5)	0.15
SLICC DI total score (0-49) ↑	0 (0; 0)	0 (-1; 0)	0 (0; 0)	1 (0; 2)	0.0005
Physician's global assessment of the patient's overall disease damage (0-10 cm) (0-10 cm) ↑	0 (-0.1; 0)	0 (-0.2; 0)	0 (-1.2; 0.05)	1 (0; 2)	0.63

p-values refer to the Kruskal-Wallis for continuous measures and Fisher's exact test for categorical measures.

used in most SLE nephritis clinical trials as opposed to multi-dimensional endpoints like the one proposed by PRINTO. There are advantages and disadvantages in each approach. In particular the use of measures related to single-organ involvement may provide more meaningful information to the trial, but this focus may limit the information on the clinical status of the patients as a whole. On the other hand, the use of multi-dimensional measures of SLE activity as a whole would dilute information related to a particular organ by contributions from other systems. When we evaluated directly the agreement between the SLICC renal index, and the PRINTO/ACR criteria, the first as prototype for organ specific and the second for multi-dimensional criteria, we found a moderate correlation (0.57). This result is in line with the observation by the SLICC group who showed a kappa coefficient, between the plurality physician rating and the calculated score obtained using other multi-dimensional indexes like the RIFLE (0.50), the original BILAG (0.14), and the BILAG 2004 (0.23).

Overall, our findings confirm the importance to consider multi-dimensional measures in addition to renal specific variables, to assess response to therapy in children with juvenile SLE and renal involvement at baseline. The PRINTO core set and the related PRINTO/ACR criteria have their strength on the evidence-based selection and validation performed on a very large sample of patients assessed in a prospective fashion. The present analysis confirms that both can be applied for patients with SLE as a whole irrespective of their disease phenotype, as well as in the subgroup of patients with SLE in which renal disease was the predominant feature. Future studies should however assess specifically whether the PRINTO core set can indeed be applied to the remaining JSLE disease subtypes.

Our study should be viewed in light of certain limitations, which include the lack of some measures (serum albumin, accuracy of urine sediment, random spot urinary protein:creatinine ratio, RIFLE and BILAG were not available) and the fact that the definition of lupus nephritis was not guided by renal biopsy class (53). In addition, since proteinuria may reflect renal damage in the absence of renal activity, a careful assessment of nephritis activity parameters as well as the potential role of angiotensin converting enzyme inhibitors in decreasing proteinuria, (54) should be done to ensure that the patient has active renal disease.

In summary, we propose to incorporate multi-dimensional response measure,

in addition to renal specific variables, in future clinical trials in patients with juvenile SLE with renal involvement. In the absence of available therapeutic trial data in juvenile SLE, both multi-dimensional criteria and renal specific criteria deserves further validation in future controlled studies to examine their discriminant validity in detecting a therapeutic response greater than placebo or the active comparator, and to assess whether further refinements of the currently available instruments are required.

References

1. ILLEI GG, YARBORO CH, KUROIWA T *et al*.: Long-term effects of combination treatment with fludarabine and low-dose pulse cyclophosphamide in patients with lupus nephritis. *Rheumatology* (Oxford) 2007; 46: 952-6.
2. GROOTSCHOLTEN C, BAJEMA IM, FLORQUIN S *et al*.: Treatment with cyclophosphamide delays the progression of chronic lesions more effectively than does treatment with azathioprine plus methylprednisolone in patients with proliferative lupus nephritis. *Arthritis Rheum* 2007; 56: 924-37.
3. BACA V, CATALAN T, VILLASIS-KEEVER M, RAMON G, MORALES AM, RODRIGUEZ-LEYVA F: Effect of low-dose cyclosporine A in the treatment of refractory proteinuria in childhood-onset lupus nephritis. *Lupus* 2006; 15: 490-5.
4. GINZLER EM, DOOLEY MA, ARANOW C *et al*.: Mycophenolate mofetil or intravenous cyclophosphamide for lupus nephritis. *N Engl J Med* 2005; 353: 2219-28.
5. CHAN TM, TSE KC, TANG CS, MOK MY, LI FK: Long-term study of mycophenolate mofetil as continuous induction and maintenance treatment for diffuse proliferative lupus nephritis. *J Am Soc Nephrol* 2005; 16: 1076-84.
6. CONTRERAS G, PARDO V, LECLERCQ B *et al*.: Sequential therapies for proliferative lupus nephritis. *N Engl J Med* 2004; 350: 971-80.
7. ILLEI GG, AUSTIN HA, III, CRANE M *et al*.: Combination therapy with pulse cyclophosphamide plus pulse methylprednisolone improves long-term renal outcome without adding toxicity in patients with lupus nephritis. *Ann Intern Med* 2001; 135: 248-57.
8. CHAN TM, LI FK, TANG CSO *et al*.: Efficacy of mycophenolate mofetil in patients with diffuse proliferative lupus nephritis. *N Engl J Med* 2000; 343: 1156-62.
9. GOURLEY MF, AUSTIN HA, III, SCOTT D *et al*.: Methylprednisolone and cyclophosphamide, alone or in combination, in patients with lupus nephritis. A randomized, controlled trial. *Ann Intern Med* 1996; 125: 549-57.
10. HOUSSIAU FA, VASCONCELOS C, D'CRUZ D *et al*.: Immunosuppressive therapy in lupus nephritis: the Euro-Lupus Nephritis Trial, a randomized trial of low-dose versus high-dose intravenous cyclophosphamide. *Arthritis Rheum* 2002; 46: 2121-31.
11. HOUSSIAU FA, VASCONCELOS C, D'CRUZ D *et al*.: Early response to immunosuppressive therapy predicts good renal outcome in lupus nephritis lessons from long-term followup of patients in the Euro-Lupus Nephritis Trial - Lessons from long-term followup of patients in the Euro-Lupus Nephritis Trial. *Arthritis Rheum* 2004; 50: 3934-40.
12. CLARK WF, CATTRAN DC, BALFE JW, WILLIAMS W, LINDSAY RM, LINTON AL: Chronic plasma exchange in systemic lupus erythematosus nephritis. *Proc Eur Dial Transplant Assoc* 1983; 20: 629-35.
13. BOUMPAS DT, AUSTIN III HA, VAUGHN EM *et al*.: Controlled trial of pulse methylprednisolone versus two regimens of pulse cyclophosphamide in severe lupus nephritis. *Lancet* 1992; 340: 741-5.
14. RENAL DISEASE SUBCOMMITTEE OF THE AMERICAN COLLEGE OF RHEUMATOLOGY AD HOC COMMITTEE ON SYSTEMIC LUPUS ERYTHEMATOSUS RESPONSE CRITERIA: The American College of Rheumatology response criteria for proliferative and membranous renal disease in systemic lupus erythematosus clinical trials. *Arthritis Rheum* 2006; 54: 421-32.
15. MOSCA M, BOMBARDIERI S: Disease-specific quality indicators, guidelines, and outcome measures in systemic lupus erythematosus (SLE). *Clin Exp Rheumatol* 2007; 25 (Suppl. 47): S107-S13.
16. PETRI M, KASITANON N, LEE SS *et al*.: Systemic lupus international collaborating clinics renal activity/response exercise: development of a renal activity score and renal response index. *Arthritis Rheum* 2008; 58: 1784-8.
17. PETRI M, KASITANON N, SINGH S *et al*.: Systemic lupus international collaborating clinics renal activity/response exercise: comparison of agreement in rating renal response. *Arthritis Rheum* 2008; 58: 1789-95.
18. RUPERTO N, MARTINI A: International research networks in pediatric rheumatology: the PRINTO perspective. *Curr Opin Rheumatol* 2004; 16: 566-70.
19. RUPERTO N, RAVELLI A, MURRAY KJ *et al*.: Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis. *Rheumatology* (Oxford) 2003; 42: 1452-9.
20. RUPERTO N, RAVELLI A, CUTTICA R *et al*.: The Pediatric Rheumatology International Trials Organization criteria for the evaluation of response to therapy in juvenile systemic lupus erythematosus: Prospective validation of the disease activity core set. *Arthritis Rheum* 2005; 52: 2854-64.
21. RUPERTO N, RAVELLI A, OLIVEIRA S *et al*.: The Pediatric Rheumatology International Trials Organization/American College of Rheumatology provisional criteria for the evaluation of response to therapy in juvenile systemic lupus erythematosus. Prospective validation of the definition of improvement. *Arthritis Rheum* 2006; 55: 355-63.
22. BRUNNER HI, GLADMAN DD, IBANEZ D, UROWITZ MD, SILVERMAN ED: Difference in disease features between childhood-onset and adult-onset systemic lupus erythematosus. *Arthritis Rheum* 2008; 58: 556-62.
23. TUCKER LB, URIBE AG, FERNANDEZ M *et al*.: Adolescent onset of lupus results in more aggressive disease and worse outcomes: results of a nested matched case-control study within LUMINA, a multiethnic US cohort (LUMINA LVII). *Lupus* 2008; 17: 314-22.
24. KLEIN A, CIMAZ R, QUARTIER P *et al*.: Causes of death in pediatric systemic lupus erythematosus. *Clin Exp Rheumatol* 2009; 27: 538-9.
25. FELSON DT, ANDERSON JJ, BOERS M *et al*.: American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 727-35.
26. GIANNINI EH, RUPERTO N, RAVELLI A, LOVELL DJ, FELSON DT, MARTINI A: Preliminary definition of improvement in juvenile arthritis. *Arthritis Rheum* 1997; 40: 1202-9.
27. RUPERTO N, RAVELLI A, FALCINI F *et al*.: Performance of the preliminary definition of improvement in juvenile chronic arthritis patients treated with methotrexate. *Ann Rheum Dis* 1998; 57: 38-41.
28. RIDER LG, GIANNINI EH, BRUNNER HI *et al*.: International consensus on preliminary definitions of improvement in adult and juvenile myositis. *Arthritis Rheum* 2004; 50: 2281-90.
29. RUPERTO N, RAVELLI A, PISTORIO A *et al*.: The provisional Pediatric Rheumatology International Trial Organization/American College of Rheumatology/European League Against Rheumatism disease activity core set for the evaluation of response to therapy in juvenile dermatomyositis: a prospective validation study. *Arthritis Rheum* 2008; 59: 4-13.
30. TAN EM, COHEN AS, FRIES JF *et al*.: The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982; 25: 1271-7.
31. HOCHBERG MC: Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997; 40: 1725.
32. LANE C, BROWN M, DUNSMUIR W, KELLY J, MANGOS G: Can spot urine protein/creatinine ratio replace 24 h urine protein in usual clinical nephrology? *Nephrology* (Carlton) 2006; 11: 245-9.
33. LEUNG YY, SZETO CC, TAM LS *et al*.: Urine protein-to-creatinine ratio in an untimed urine collection is a reliable measure of proteinuria in lupus nephritis. *Rheumatology* (Oxford) 2007; 46: 649-52.
34. 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.
35. 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.
36. VITALI C, BENCIVELLI W, ISENBERG DA *et al*.: Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research. II. Identification of the variables indicative of disease activity and their use in the development of an activity

score. *Clin Exp Rheumatol* 1992; 10: 541-7.

37. VITALI C, BENCIVELLI W, MOSCA M, CARRAI P, SERENI M, BOMBARDIERI S: Development of a clinical chart to compute different disease activity indices for systemic lupus erythematosus. *J Rheumatol* 1999; 26: 498-501.
38. 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.
39. The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents: *Pediatrics* 2004; 114 (2 Suppl. 4th Report): 555-76.
40. LEVEY AS, CORESH J, BALK E *et al*.: National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med* 2003; 139: 137-47.
41. RUPERTO N, RAVELLI A, PISTORIO A *et al*.: Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. *Clin Exp Rheumatol* 2001; 19: S1-S9.
42. LANDGRAF JM, ABETZ L, WARE JE: The CHQ User's Manual. First Edition. ed., Boston, MA, USA, The Health Institute, New England Medical Center, 1996.
43. GLADMAN DD, GOLDSMITH CH, UROWITZ MB *et al*.: The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index for Systemic Lupus Erythematosus International Comparison. *J Rheumatol* 2000; 27: 373-6.
44. LIANG MH, FOSSEL AH, LARSON MG: Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990; 28: 632-42.
45. MOSTELLER F, TUKEY JW: Data analysis and regression. Reading, MA ed., Addison Wesley, 1977: 133.
46. LIANG MH, LARSON MG, CULLEN KE, SCHWARTZ JA: Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum* 1985; 28: 542-7.
47. LIANG MH: Evaluating measurement responsiveness. *J Rheumatol* 1995; 22: 1191-2.
48. COHEN J: A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37-46.
49. COHEN J: Statistical Power Analysis for the Behavioral Sciences. New York, Academic Press, 1977.
50. LANDIS JR, KOCH GC: The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74.
51. AUSTIN HA, KLIPPEL JH, BALOW JE *et al*.: Therapy of lupus nephritis. Controlled trial of prednisone and cytotoxic drugs. *N Engl J Med* 1986; 314: 614-9.
52. YEE CS, GORDON C, DOSTAL C *et al*.: EULAR randomised controlled trial of pulse cyclophosphamide and methylprednisolone versus continuous cyclophosphamide and prednisolone followed by azathioprine and prednisolone in lupus nephritis. *Ann Rheum Dis* 2004; 63: 525-9.
53. WEENING JJ, D'AGATI VD, SCHWARTZ MM *et al*.: The classification of glomerulonephritis in systemic lupus erythematosus revisited. *Kidney Int* 2004; 65: 521-30.
54. BOUMPAS DT, BALOW JE: Outcome criteria for lupus nephritis trials: a critical overview. *Lupus* 1998; 7: 622-9.
55. UROWITZ MB, IBANEZ D, ALI Y, GLADMAN DD: Outcomes in patients with active lupus nephritis requiring immunosuppressives who never received cyclophosphamide. *J Rheumatol* 2007; 34: 1491-6.