NEPHROLOGY - ORIGINAL PAPER

Semiquantitative and semi-automated morphometric evaluation of chronic lesions in renal biopsies

Daniel Abensur Athanazio · Gloria Maria Maranhão Sweet ·
Carlos Alberto Silva · Washington Luis Conrado dos-Santos

Abstract Chronic lesions in renal biopsies are a well recognized prognostic factor for renal diseases, including lupus nephritis. The methods used for assessment of chronic lesions are, however, largely based on semiquantitative evaluation and may lead to poor reproducibility. Interobserver variation is particularly important in lupus nephritis, in which acute and chronic lesions may occur simultaneously. In this study we tested the reproducibility of chronic lesion assessment performed by three pathologists, two with specific training in renal pathology, using 20 renal biopsies and a standard semiquantitative method. In a second experiment, we evaluated the reproducibility of chronic lesion assessment in 33 biopsies of lupus nephritis by the two nephropathologists. The semiquantitative estimated values were compared with those from a previously proposed morphometric method for quantification of chronic lesions in renal biopsies. Although correlations were observed among the estimated values, there was a wide range of variation when semiquantitative methods were used. In particular, activity and chronicity indices of lupus nephritis were poorly reproducible. In contrast, use of a morphometric score, although not eliminating interobserver variability, led to better reproducibility of estimated values than that obtained with semiquantitative methods.

Keywords Lupus nephritis · Kidney · Renal biopsy · Reproducibility of results · Interstitial fibrosis

D. A. Athanazio · G. M. M. Sweet · C. A. Silva ·
W. L. C. dos-Santos
Gonçalo Moniz Research Center, Oswaldo Cruz
Foundation, Salvador, Brasil

D. A. Athanazio
Departament of Biointeraction, Federal University
of Bahia, Salvador, Brasil

D. A. Athanazio (✉)
Departamento de Biointeração—ICS—UFBA,
Av. Reitor Miguel Calmon s/no—Campus do Canela,
40.110-100 Salvador, Bahia, Brasil
e-mail: daa@ufba.br; daniel_fiocruz@hotmail.com

C. A. Silva
Departament of Anatomic Pathology and Legal Medicine,
Federal University of Bahia, Salvador, Brasil

Introduction

The extent of chronic lesions in kidney biopsies is the most reliable predictor of chronic renal failure [1–3]. Most surgical pathologists use subjective grading systems routinely instead of computer-based objective analysis. Although subjective analysis is easier to use, digital measurements are potentially more reliable and reproducible [4]. A well known example of this situation is the use of activity and chronicity indices in lupus nephritis [5].

Lupus nephritis is characterized by considerable variability in morphological presentation among

different patients, different biopsies from the same patient, and within the same biopsy [6]. Moreover, renal morphology may have poor correlation with clinical presentation [6], and the extent to which some lesions are associated with underlying immune responses in systemic lupus erythematosus is not clear [7]. Hence, concerns about how pathologists should report predictors of disease outcome, for example chronic damage, are more justified than in other diseases. To deal with the broad spectrum of lesions present in lupus nephritis, activity and chronicity indices were proposed by the National Institute of Health (NIH, United States) [8] and are widely used [6]. Evaluations of these NIH-proposed systems have led to inconsistent results, ranging from a strong power for predicting treatment response (activity index) [8, 9] and prognosis (chronicity index) [8–12] to limited application, because of their poor reproducibility, among pathologists [5, 13]. Stratification of chronicity index into risk groups (0–1, 2–3 or >4) further complicates this scenario, because small differences between observers may result in the inclusion of the same patient in different risk groups. Such observations have hampered the use of the NIH indices for decision making on management of individual cases [5]. The classification of lupus nephritis has been recently reviewed [14] and still emphasizes the distinction of predominant active or chronic lesions in each class. However, the issue on how to score chronic lesions still persists.

More recently, a semi-automated morphometric score based on the ratio of chronic lesions to total cortical area was shown to be a reliable method for predicting long-term progression of renal diseases [15], including lupus nephritis [16]. To the best of our knowledge, this system has not yet been tested in other reference centers for lupus nephritis. Thus, the objectives of this study were:

1 to evaluate intra and interobserver variability of the recently proposed morphometric score among nephropathologists without mutual education (i.e. trained in different institutions);
2 to compare assessment of chronic lesions by semiquantitative and morphometric methods in renal biopsies with diverse diagnoses;
3 to evaluate agreement of semiquantitative assessment of individual lesions typically associated with lupus nephritis;

4 to test agreement of the NIH's activity and chronicity indices in lupus nephritis biopsies between nephropathologists; and
5 to compare assessment of chronic lesions by semiquantitative and morphometric methods in lupus nephritis biopsies.

Subjects and methods

Renal biopsy specimens

We used samples of renal biopsies from 464 patients referred to the Gonçalo Moniz Research Center, Oswaldo Cruz Foundation (CPqGM-FIOCRUZ), by community hospitals from Salvador, Brazil, between 1998 and 2007: The first sample comprised 20 randomly selected biopsies with any conclusive diagnosis of nephropathy and available measurements of serum urea and creatinine levels performed at the time of the biopsy. This group sample was selected for comparative study of cortical interstitial fibrosis assessment by three pathologists using semiquantitative analysis, and for evaluation of the relationship between the semiquantitative analysis performed by the pathologists, the semi-automated morphometric score of chronic damage proposed by Howie and colleagues [15, 16], and the extent of glomerulosclerosis. Each slide was evaluated by semiquantitative assessment of cortical interstitial fibrosis and the morphometric score. All methods of chronic damage assessment were correlated with serum urea and creatinine levels.

The second group included all renal biopsies with the diagnosis of lupus nephritis received from January 2002 to August 2007. This group (33 biopsies) was used to compare the semiquantitative assessment of lesions commonly seen in lupus nephritis, the activity and chronicity indices scored by two pathologists, and how these measurements correlated with the semi-automated morphometric index of chronic damage proposed by Howie and colleagues [15, 16]. Each slide was evaluated by semiquantitative assessment of cortical interstitial fibrosis, activity and chronicity indices, grading of individual lesions, and the morphometric score (see below). The methods of damage assessment were correlated with serum urea and creatinine levels.

Histopathological evaluation

Three pathologists reviewed the slides. Two (nephropathologist 1, NP1 and nephropathologist 2, NP2) had received specific training in renal pathology at different institutions for at least six months and were now working together at the same institution (CPqGM-FIOCRUZ). The third (general pathologist, GP) had no specific previous training in renal pathology.

Semiquantitative analysis

For the first sample, pathologists were asked to evaluate cortical interstitial fibrosis, reporting it as a percentage of the represented cortical tissue area in sirius red-stained sections. For comparison purposes, these measurements were stratified in categories from 0 to 3 based on the Banff criteria of cortical interstitial fibrosis for evaluation of renal allograft rejection: $ci0$ (interstitial fibrosis in up to 5% of cortical area), $ci1$ (6–25%), $ci2$ (26–50%), and $ci3$ (>50%) [17]. For the second sample, pathologists were also asked to evaluate the extent of each individual lesion found in the biopsies, using a subjective semiquantitative grading (0, absent; 1, mild; 2, moderate; 3, severe) and to score biopsy specimens using the NIH activity index [8]. Briefly, each biopsy specimen was assessed for ten components, six of which were included in the activity index or in the chronicity index. Each individual component scored 0 (normal), 1 (slight abnormality), 2 (moderate abnormality), or 3 (severe abnormality). In calculating the activity index, cellular/endothelial proliferation, leukocyte/neutrophil glomerular infiltration, hyaline thrombi or wire-loop lesions, inflammatory mononuclear interstitial infiltrates, fibrinoid necrosis, and cellular crescents were evaluated.

Fibrinoid necrosis and cellular crescents are weighted by a factor 2. In calculating the chronicity index, all lesions (glomerular sclerosis, fibrous crescents, interstitial fibrosis, and tubular atrophy) are weighted by a factor 1. The maximum activity index is 24 points and the maximum chronicity index is 12 points. Pathologists scored each case by examining slides routinely stained by use of the hematoxylin–eosin, periodic acid–Schiff, silver methamine, sirius red, and Azan techniques.

Semi-automated morphometric score for chronic damage

An interactive image-analysis system (ImagePro-Plus, Media Cybernetics) was used to measure the extent of chronic damage for biopsy specimens. Digital images were obtained from hematoxylin–eosin-stained sections using a microscope linked to a computer. On each image the total cortical area and the areas of chronic damage were measured (Fig. 1). Chronic damage was defined as globally sclerosed glomeruli, atrophic tubules, interstitial fibrosis, thickened vessels, and tubular cysts [14, 15]. In the papers by Howie and colleagues [13, 14], silver methamine staining was used. In this study, each observer had all routine stain sections available (hematoxylin–eosin, periodic acid–Schiff, silver methamine, sirius red, and Azan), however, they were asked to use hematoxylin–eosin-stained sections for the drawing procedure (demarcation of chronic lesions in cortical areas, as illustrated in Fig. 1).

Statistical analysis

Concordance between categorical results of the tests was evaluated using the kappa coefficient (calculated according to Landis and Koch) [18]. The Spearman correlation test was used for comparison of numeric variables. The Kendall correlation test was used for comparison of numeric ordinal variables. Bland and Altman plots were used to further evaluate the reproducibility of results from numeric variables [19].

Results

Morphometric score of chronic damage: intraobserver and interobserver reproducibility

Each nephropathologist evaluated the morphometric score of chronic damage in ten randomly selected lupus nephritis biopsies. The intraobsever variation of NP1 had a mean difference of -0.2 between the first and second observations (CI: -7.0 to -7.4) (Fig. 2) and there was a strong correlation between both estimated values (Spearman $r = 0.96$; $P < 0.0001$). The intraobsever variation of NP2 had a mean difference of -1.7 between the first and second
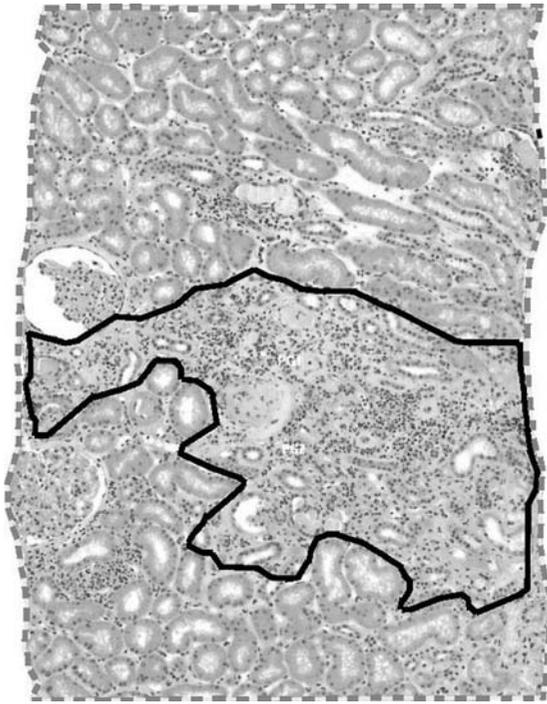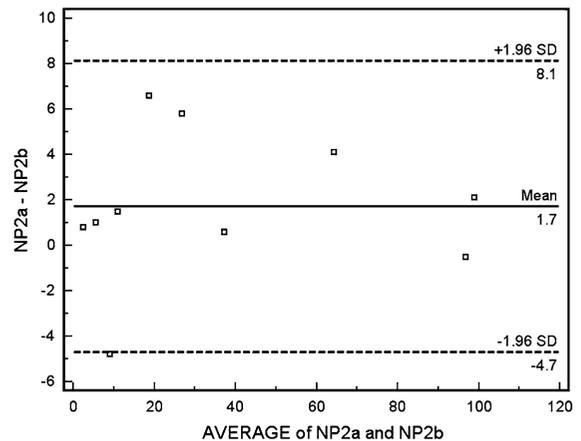
Fig. 1 An example of how the morphometric score of chronic lesions is measured as described in the section Semi-automated morphometric score for chronic damage. The cortical area is outlined (*solid line*) and a central area containing interstitial fibrosis, tubular atrophy, and one glomerulus of global sclerosis is marked (*dashed line*)
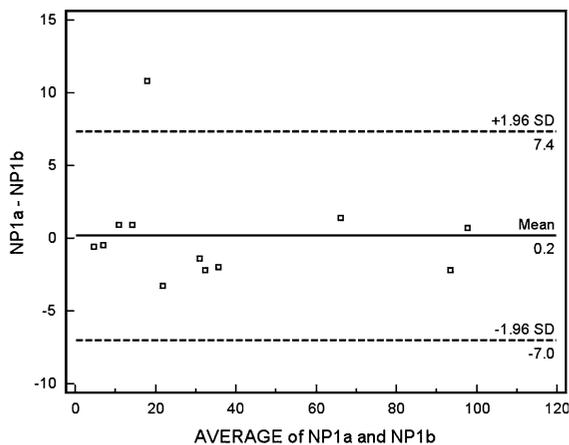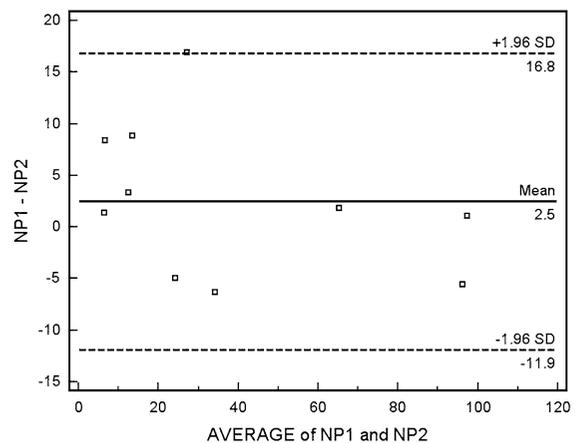

Fig. 2 Bland and Altman plot. Agreement of the morphometric scores of chronic lesions for nephropathologist 1 (*NP1a* denotes the first observation and *NP1b* the second observation)

Fig. 3 Bland and Altman plot. Agreement of the morphometric scores of chronic lesions for nephropathologist 2 (*NP2a* denotes the first observation and *NP2b* the second observation)

Fig. 4 Bland and Altman plot. Agreement between nephropathologists for the morphometric score of chronic lesions (*NP1* denotes nephropathologist 1 and *NP2* denotes nephropathologist 2)

interobsever variability was compared, the mean difference between NP1 and NP2 was 2.5 (CI: -11.9 to 16.8) (Fig. 4) maintaining a strong correlation (Spearman $r = 0.93$; $P < 0.0001$).

Comparison of semiquantitative and semi-automated indices in 20 renal biopsies with diverse diagnoses of renal disease

In the first sample of 20 renal biopsies the diagnoses were: focal segmental glomerulosclerosis in four,

observations (CI: -4.7 to 8.1) (Fig. 3) and there was also a strong correlation between both estimated values (Spearman $r = 0.99$; $P < 0.0001$). When

lupus nephritis in three (all class IV), sclerosing glomerulonephritis in three, diffuse proliferative glomerulonephritis in three, membranoproliferative glomerulonephritis in two, and interstitial nephritis in two. The remaining three cases had diagnoses of crescentic glomerulonephritis, mesangial glomerulonephritis, and sclerosing tubulointerstitial nephritis. The biopsies had 14.7 ± 7.3 glomeruli, 4.7 ± 5.7 of which with global sclerosis. The mean interstitial cortical fibrosis evaluated by the three pathologists was $32.9 \pm 30.6\%$ and the semi-automated morphometric score of chronic damage was $36.9 \pm 30.7\%$. The patients had serum creatinine levels of 2.3 ± 1.7 mg/dl at the time of the biopsy.

When we tested the correlation between the morphometric score of chronic damage and specific morphologic changes in renal biopsies (interstitial fibrosis and glomerulosclerosis) or clinical markers of renal dysfunction (increased serum urea and creatinine levels), there were positive correlations for all four tests (Table 1). There was a moderate correlation between serum creatinine levels and the morphometric score of chronic damage. A strong correlation was observed between the morphometric score and interstitial fibrosis and a strong correlation for glomerulosclerosis. These strong correlations were expected, because these lesions were integral parts of the morphometric analysis. Interestingly, however, the glomerular lesions correlated better with the index than did interstitial fibrosis, even if the interstitium in the renal cortex predominated over total glomerular area and had a higher weight in the final calculation of the morphometric score.

When the assessment of interstitial fibrosis was compared between pathologists, the mean differences (and 95% confidence intervals) between pathologists were 28.4% (−14.7 to 13.7) for NP1 and NP2, 45.7% (−22.6 to 23.1) for NP1 and GP, and 44.3% (−21.9 to 22.4) for NP2 and GP (Bland and Altman plot

figures not shown). The Spearman correlation coefficients were: rs = 0.96 $P < 0.0001$ for NP1 and NP2, rs = 0.89 $P < 0.0001$ for NP1 and GP, and rs = 0.91 $P < 0.0001$ for NP2 and GP.

The agreement coefficient (*kappa*) among the three pathologists for interstitial fibrosis after stratification by Banff criteria was 0.76 (CI: 0.60–0.92). The agreement for each class was $k = 0.64$ (CI: 0.39–0.89) for Banff 0, $k = 0.72$ (CI: 0.47–0.98) for Banff 1, $k = 0.68$ (CI: 0.42–0.93) for Banff 2 and $k = 0.92$ (CI: 0.67–1.00) for Banff 3. All these agreements were statistically significant ($P < 0.001$). When testing the agreement between only pairs of observers, the concordance among nephropathologists (NP1 and NP2) was higher ($k = 0.86$ CI:0.59–1.00 $P < 0.001$) than between NP1 and GP ($k = 0.71$ CI: 0.44–0.98 $P < 0.001$) and NP2 and GP ($k = 0.72$ CI: 0.45–0.99 $P < 0.001$). All these agreement coefficients were statistically significant ($P < 0.001$).

Agreement of semiquantitative assessment of lupus nephritis-associated lesions

There was overall poor agreement of semiquantitative estimated values of lupus nephritis lesions by different pathologists. The lowest agreement was observed for mesangial hypercellularity (which could be explained by chance) and the highest agreement was for interstitial nephritis. Agreement improved within the lowest class of each individual lesion (Table 2).

Agreement of semiquantitative assessment of lupus nephritis using the NIH activity and chronicity indices

When subdividing both activity and chronicity indices into three categories (0–5, 6–10, and >10 for activity index and 0–4, 5–8, and >8 for chronicity

Table 1 Correlation between semi-automated morphometric score of chronic damage and morphologic and clinical features in 20 biopsies of different diagnoses of renal diseases

Spearman correlation	rs	P
Score of chronic damage × serum creatinine levels	0.7489	0.0001
Score of chronic damage × serum urea levels	0.6235	0.0033
Score of chronic damage × interstitial fibrosis[a]	0.8791	0.0000
Score of chronic damage × glomerulosclerosis rate	0.8993	0.0000

[a] Mean of semiquantitative assessment by three observers
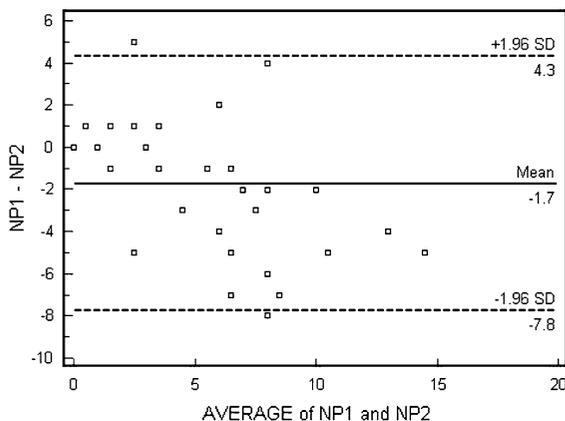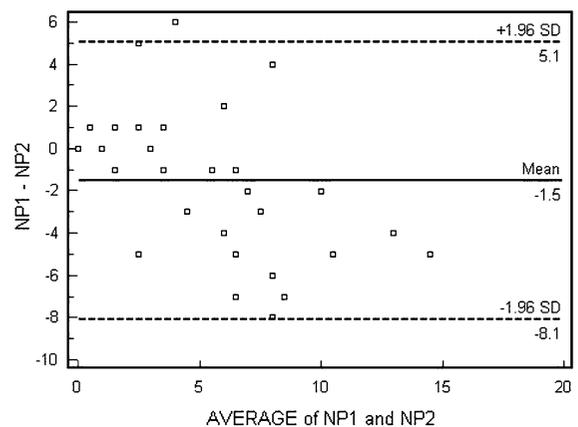
Table 2 Agreement of semiquantitative assessment of lupus nephritis-associated lesions

	Kappa (all classes)	P	Kappa (for lowest class)	P
Mesangial hypercellularity	0.14	0.20	0.40 (class 1)	0.01
Endocapillary proliferation	0.21	0.06	0.55 (class 0)	0.001
Neutrophils	0.42	0.001	0.88 (class 0)	<0.001
Fibrous and fibrocellular crescents	0.40	0.001	0.52 (class 0)	0.002
Interstitial nephritis	0.48	<0.001	0.55 (class 0)	0.001
Tubular atrophy	0.12	0.05	0.55 (class 0)	0.001
Interstitial fibrosis	0.52	<0.001	0.91 (class 0)	<0.001

index), there was only moderate agreement between the nephropathologists: $k = 0.49$ ($P < 0.0001$) and $k = 0.40$ ($P = 0.0001$) for activity and chronicity indices, respectively. The Kendall coefficient showed a strong correlation for activity index ($r = 0.62$ $P < 0.0001$) and for chronicity index ($r = 0.82$ $P < 0.0001$) between observers. The Bland and Altman plots, however, showed a substantially wide range of variation between pathologists. The mean difference between NP1 and NP2 was -1.7 (CI: -7.8 to 4.7) for activity index and -1.5 (CI: -8.1 to 5.1) for chronicity index (Figs. 5, 6).

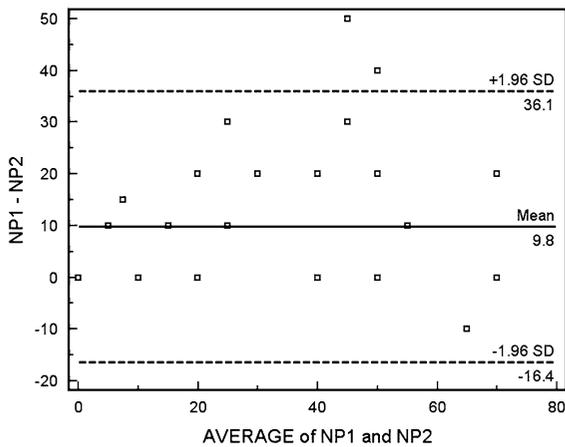## Comparison of semiquantitative and semi-automated methods for 33 renal biopsies with diagnosis of lupus nephritis

Estimated values of interstitial fibrosis by nephropathologists showed a strong correlation (Spearman $r = 0.89$ $P < 0.0001$). Bland and Altman plot

Fig. 6 Bland and Altman plot. Agreement between nephropathologists for the NIH chronicity index (*NP1* denotes nephropathologist 1 and *NP2* denotes nephropathologist 2)

analysis showed a mean difference of 9.8%, ranging from -16.4 to 36.1, for a 95% confidence interval (Fig. 7).

Only the estimated values of NIH activity index did not correlate with increased serum creatinine levels. The glomerulosclerosis, interstitial fibrosis (measured by both nephropathologists), NIH chronicity index (measured by both nephropathologists), and the morphometric chronicity score (measured by NP1) significantly correlated with serum creatinine levels. Such correlation was lowest for glomeruloesclerosis ratio (rs = 0.50 $P < 0.007$) and highest for the morphometric chronicity score (rs = 0.73 $P < 0.0001$) (Table 3).

The morphometric score of chronic lesions correlated significantly with all specific measurements of chronic lesions. It was, however, lowest for glomerulosclerosis ratio (rs = 0.69 $P < 0.0001$) and highest for NIH chronicity index (rs = 0.84 $P < 0.0001$) and interstitial fibrosis (rs = 0.86 $P < 0.0001$) (Table 4).

Fig. 5 Bland and Altman plot. Agreement between nephropathologists for the NIH activity index (*NP1* denotes nephropathologist 1 and *NP2* denotes nephropathologist 2)

Fig. 7 Bland and Altman plot. Agreement between nephro-pathologists for interstitial fibrosis in lupus nephritis biopsies (*NP1* denotes nephropathologist 1 and *NP2* denotes nephropathologist 2)

Table 3 Spearman correlation between semiquantitative or morphometric measurements and serum creatinine levels in lupus nephritis

Correlation with serum creatinine	rs	P
Glomerulosclerosis ratio	0.50	0.007
NIH activity index (NP1)	0.25	0.19
NIH activity index (NP2)	0.23	0.22
NIH chronicity index (NP1)	0.56	0.002
NIH chronicity index (NP2)	0.62	<0.001
Interstitial fibrosis (NP1)	0.64	<0.001
Interstitial fibrosis (NP2)	0.61	<0.001
Morphometric score of chronic lesions (NP1)	0.73	<0.0001

Table 4 Spearman correlation between the morphometric score of chronic cortical damage and semiquantitative methods in lupus nephritis

	rs	P
Glomerulosclerosis ratio	0.69	<0.0001
Mean NIH chronicity index[a]	0.84	<0.0001
Mean interstitial fibrosis[a]	0.86	<0.0001

[a] Mean of both observers

Comparisons for both NIH chronicity index and interstitial fibrosis were performed using the values estimated by NP1, who also determined the morphometric chronicity score for all cases.

Discussion

The prognosis of lupus nephritis has markedly improved in the last decades because of better management of the disease by potent immunosuppressive drugs. On the other hand, failure to achieve remission may range from 3 to 50% in some series of patients receiving cyclophosphamide [20, 21]. More recently, different studies have suggested higher rates of remission with mycophenolate mofetil [22], which is now starting to be used also in nonlupus glomerulonephritis [23]. The rates of treatment failure, however, ranges from 3 to 40% in patients treated with mycophenolate mofetil [22]. Such variations of outcome emphasize the importance of reliable predictors of prognosis in lupus nephritis. Most proposed prognostic factors based on renal biopsy findings have the drawback of poor or uncertain reproducibility. Howie and colleagues reported a powerful morphometric indicator of prognosis in lupus nephritis [16] but the reproducibility of this method has not been evaluated in other nephropathology reference centers.

The interobserver variability in renal biopsy interpretation of lupus nephritis associated lesions was expected. It has been suggested that good reproducibility of NIH activity and chronicity indices in some reports could be explained by mutual education, because observers were usually trained in the same institution [5]. Unacceptably high intra and interobserver variability has been reported in some studies [5, 13], including the field application of these indices by pathologists of different institutions (community hospitals), thus excluding the bias of mutual education [5]. Our results provide further evidence that categorization of these indices leads to considerable variation between observers and should not be used in clinical protocols or clinical practice to make decisions for individual cases [5]. A recent work reported good agreement for activity index and a moderate agreement for chronicity index among three nephropathologists but, again, mutual education (as inferred by consensus meeting) could explain their better performance [24].

We observed strong correlations between the rate of glomerulosclerosis and the morphometric score of chronic damage in the 20 renal biopsies with various diagnoses and in biopsies from patients with lupus nephritis. On the other hand, the morphometric score increases the weight of tubulointerstitial chronic

lesions over glomerular changes. For comparison, glomerular chronic damage has the weight of 50% of the NIH chronicity index, because two of the four criteria are glomerular sclerosis and fibrous crescents. This is important, because tubulointerstitial chronic lesions are increasingly recognized as important and underestimated prognostic markers in lupus nephritis [14, 25].

As described by Bland and Altman, we confirmed that even high correlation coefficients are not sufficient to evaluate the reproducibility of results [19]. When plotting numeric variables evaluated by two nephropathologists, we observed a wide range variation (Figs. 2–7). Based on comparisons made in the present study, it seems that the NIH activity and chronicity indices are the least reproducible morphological grading systems. For activity index (variation ranges from 0 to 24), the confidence interval for the difference among two nephropathologists ranged from −7.8 to 4.3, which means that the distribution of differences comprises 50% of the scale used (Fig. 5). For the chronicity index, the variation of differences comprises the whole scale used of 0–12 points (Fig. 6).

The use of a scale from 0 to 100, not based on the semiquantification of different individual lesions, has the potential to minimize such variations. For cortical interstitial fibrosis in lupus nephritis, however, such an approach was shown to be inefficient, because the confidence interval between measurements ranged from −16.4 to 36.1 comprising 52.5% of the entire scale (Fig. 7). Interestingly, the reproducibility of interstitial fibrosis assessment was higher among the 20 biopsies of diverse renal disease. The confidence intervals of differences between observers were 28.4% for NP1 and NP2, 45.7% for NP1 and GP, and 44.3% for NP2 and GP. This was also apparent from higher correlation coefficients in the 20 biopsies of various diagnoses between nephropathologists. The concordance among pathologists in biopsies of various diagnoses ($k = 0.75$) was higher than that observed in lupus nephritis ($k = 0.52$) for stratification of interstitial fibrosis according to the Banff criteria. Such variation may be explained by the common coexistence of active and chronic lesions in lupus nephritis and the discrimination of both may result in an additional source of variation in estimated values compared with when only lesions with the same nature are observed in one biopsy specimen.

The other measurement of chronic damage that could be evaluated on a 0–100% scale was the morphometric score. In this case, the confidence interval of difference between two pathologists was 28.7 (−11.9 to 16.8) (Fig. 4). Importantly, the mean interobserver difference was 2.5 compared with 9.8 achieved for the estimated values for interstitial fibrosis. Hence, the morphometric score was the most reproducible method tested in our study.

The data in Table 2 deserve further comments. The concordance of both nephropathologists was low for the semiquantitative result of individual lesions (grades 0–3). There was higher concordance for the lowest observed class. This implies that there are two main causes of variation of results for biopsy samples. Pathologists will judge differently the presence or absence of some lesions (and this is not directly related to the method of assessment) and may differently quantify the detected lesions. Automated morphometric analysis solves the problem in the second case. At least this excludes the variations of semiquantitative methods and tends to minimize variations if both pathologists are classifying the lesion in the same way. Routine renal biopsies are already expensive and time-consuming, because of the requirement of immunofluorescence, electron microscopy, and multiple stains. The use of a computer-based morphometric score would add cost and time to this routine. On the other hand, the current routine would seem to be useless if the observations are poorly reproducible. Thus, use of such a new tool to improve the way pathologists report chronic lesions, making it more reliable and understandable to clinicians may justify the added cost in terms of software, hardware, and pathologist time.

Conclusions

Although the tested semi-automated morphometric evaluation of renal biopsies does not solve the problem of interobserver variation, it is more reproducible than current widely used semiquantitative approaches.

References

1. Risdon RA, Sloper JC, De Wardener HE (1968) Relationship between renal function and histological changes

found in renal-biopsy specimens from patients with persistent glomerular nephritis. Lancet 2:363–366. doi: 10.1016/S0140-6736(68)90589-8

2. Striker GE, Schainuck LI, Cutler RE, Benditt EP (1970) Structural-functional correlations in renal disease. I. A method for assaying and classifying histopathologic changes in renal disease. Hum Pathol 1:615–630. doi: 10.1016/S0046-8177(70)80060-0

3. Bohle A, Mackensen-Haen S, von Gise H et al (1990) The consequences of tubulo-interstitial changes for renal function in glomerulopathies. A morphometric and cytological analysis. Pathol Res Pract 186:135–144

4. Furness PN (1997) The use of digital images in pathology. J Pathol 183:253–263. doi:10.1002/(SICI)1096-9896 (199711)183:3<253::AID-PATH927>3.0.CO;2-P

5. Wernick RM, Smith DL, Houghton DC et al (1993) Reliability of histologic scoring for lupus nephritis: a community-based evaluation. Ann Intern Med 119:805–811

6. Cameron JS (2003) Lupus nephritis. In: Johnson RJ, Feehally J (eds) Comprehensive clinical nephrology, 2nd edn. Mosby, London, pp 357–371

7. Liapis H, Tsokos GC (2007) Pathology and immunology of lupus glomerulonephritis: can we bridge the two? Int Urol Nephrol 39:223–231. doi:10.1007/s11255-006-9170-x

8. Austin HAIII, Muenz LR, Joyce KM et al (1984) Diffuse proliferative lupus nephritis: identification of specific pathologic features affecting renal outcome. Kidney Int 25:689–695. doi:10.1038/ki.1984.75

9. Rush PJ, Baumal R, Shore A et al (1986) Correlation of renal histology with outcome in children with lupus nephritis. Kidney Int 29:1066–1071. doi:10.1038/ki.1986.108

10. Banfi G, Mazzucco G, Barbiano di Belgiojoso G et al (1985) Morphological parameters in lupus nephritis: their relevance for classification and relationship with clinical and histological findings and outcome. Q J Med 55:153–168

11. Esdaile JM, Levinton C, Federgreen W et al (1989) The clinical and renal biopsy predictors of long-term outcome in lupus nephritis: a study of 87 patients and review of the literature. Q J Med 72:779–833

12. Nossent HC, Henzen-Logmans SC, Vroom TM et al (1990) Contribution of renal biopsy data in predicting outcome in lupus nephritis. Analysis of 116 patients. Arthritis Rheum 33:970–977. doi:10.1002/art.1780330708

13. Schwartz MM, Lan SP, Bernstein J et al (1993) Irreproducibility of the activity and chronicity indices limits their utility in the management of lupus nephritis. Lupus Nephritis Collaborative Study Group. Am J Kidney Dis 21:374–377

14. Weening JJ, D'Agati VD, Schwartz MM et al (2004) The classification of glomerulonephritis in systemic lupus erythematosus revisited. Kidney Int 65:521–530. doi:10.1111/j.1523-1755.2004.00443.x

15. Howie AJ, Ferreira MA, Adu D (2001) Prognostic value of simple measurement of chronic damage in renal biopsy specimens. Nephrol Dial Transplant 16:1163–1169. doi:10.1093/ndt/16.6.1163

16. Howie AJ, Turhan N, Adu D (2003) Powerful morphometric indicator of prognosis in lupus nephritis. QJM 96:411–420. doi:10.1093/qjmed/hcg074

17. Solez K, Axelsen RA, Benediktsson H et al (1993) International standardization of criteria for the histologic diagnosis of renal allograft rejection: the Banff working classification of kidney transplant pathology. Kidney Int 44:411–422. doi:10.1038/ki.1993.259

18. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174. doi:10.2307/2529310

19. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310

20. Sabry A, Abo-Zenah H, Medhat T et al (2008) A comparative study of two intensified pulse cyclophosphamide remission-inducing regimens for diffuse proliferative lupus nephritis: an Egyptian experience. Int Urol Nephrol. doi:10.1007/s11255-007-9325-4

21. Mok CC, Wong RW, Lai KN (2003) Treatment of severe proliferative lupus nephritis: the current state. Ann Rheum Dis 62:799–804. doi:10.1136/ard.62.9.799

22. Walsh M, James M, Jayne D et al (2007) Mycophenolate mofetil for induction therapy of lupus nephritis: a systematic review and meta-analysis. Clin J Am Soc Nephrol 2:968–975. doi:10.2215/CJN.01200307

23. Paydas S, Kurt C, Taskapan H et al (2008) The effect of mycophenolate mofetil on primary and secondary treatment of primary glomerulonephritis and lupus nephritis. Int Urol Nephrol. doi:10.1007/s11255-008-9454-4

24. Grootscholten C, Bajema IM, Florquin S et al (2008) Interobserver agreement of scoring of histopathological characteristics and classification of lupus nephritis. Nephrol Dial Transplant 23:223–230. doi:10.1093/ndt/gfm555

25. Ordonez NG (2004) Urinary tract—non neoplastic diseases. In: Houston M (ed) Rosai and Ackerman's surgical pathology, 9th edn. Mosby Elsevier, St Louis, pp 1166–1205