

# Interobserver Agreement on Histopathological Lesions in Class III or IV Lupus Nephritis

Suzanne Wilhelmus,\* H. Terence Cook,<sup>†</sup> Laure-Hélène Noël,<sup>‡</sup> Franco Ferrario,<sup>§</sup> Ron Wolterbeek,<sup>||</sup> Jan A. Bruijn,\* and Ingeborg M. Bajema\*

## Abstract

**Background and objectives** To treat lupus nephritis effectively, proper identification of the histologic class is essential. Although the classification system for lupus nephritis is nearly 40 years old, remarkably few studies have investigated interobserver agreement. Interobserver agreement among nephropathologists was studied, particularly with respect to the recognition of class III/IV lupus nephritis lesions, and possible causes of disagreement were determined.

**Design, setting, participants, & measurements** A link to a survey containing pictures of 30 glomeruli was provided to all 360 members of the Renal Pathology Society; 34 responses were received from 12 countries (a response rate of 9.4%). The nephropathologist was asked whether glomerular lesions were present that would categorize the biopsy as class III/IV. If so, additional parameters were scored. To determine the interobserver agreement among the participants,  $\kappa$  or intraclass correlation values were calculated. The intraclass correlation or  $\kappa$ -value was also calculated for two separate levels of experience (specifically, nephropathologists who were new to the field or moderately experienced [less experienced] and nephropathologists who were highly experienced).

**Results** Intraclass correlation for the presence of a class III/IV lesion was 0.39 (poor). The  $\kappa$ /intraclass correlation values for the additional parameters were as follows: active, chronic, or both: 0.36; segmental versus global: 0.39; endocapillary proliferation: 0.46; influx of inflammatory cells: 0.32; swelling of endothelial cells: 0.46; extracapillary proliferation: 0.57; type of crescent: 0.46; and wire loops: 0.35. The highly experienced nephropathologists had significantly less interobserver variability compared with the less experienced nephropathologists ( $P=0.004$ ).

**Conclusions** There is generally poor agreement in terms of recognizing class III/IV lesions. Because experience clearly increases interobserver agreement, this agreement may be improved by training nephropathologists. These results also underscore the importance of a central review by experienced nephropathologists in clinical trials.

*Clin J Am Soc Nephrol* 10: 47–53, 2015. doi: 10.2215/CJN.03580414

## Introduction

One of the most severe manifestations of SLE is lupus nephritis, a major cause of high morbidity and mortality, because of either the disease itself or adverse effects associated with immunosuppressive therapy (1,2). In treating lupus nephritis, the histologic class plays a central role in guiding treatment decisions (3–5). The current version of the classification system was proposed in 2003 by the International Society of the Nephrology and the Renal Pathology Society (ISN/RPS) (6) and has been adopted by renal pathologists worldwide. After the introduction of this revised classification system, several studies were performed comparing the revised classification system with the 1995 World Health Organization (WHO) classification system. Each of these studies (7–9) found higher interobserver agreement using the revised system, and this was attributed to clearer definitions and fewer subclasses.

The most important decision with respect to treating lupus nephritis is whether the biopsy can be classified as class III or IV rather than class I, II, or V. On the basis of

current guidelines for treating lupus nephritis (3–5), a diagnosis of either class III or IV is an indication for initiating immunosuppressive therapy. Furthermore, in the most recent version of the classification system, allowance has been made to include lesions that are not strictly proliferative—for example, the sole presence of wire loops—in classes III and IV. Because obtaining an accurate diagnosis is essential for determining the subsequent treatment, it is imperative that renal pathologists reach consensus regarding what constitutes a lesion that would place the biopsy in class III or IV.

While performing a central review of lupus nephritis trials, we noted that, even in a selected group of highly experienced nephropathologists, there were differences of opinion with respect to what constitutes a lesion that would categorize the biopsy as class III or IV. On the basis of this experience, we sought to measure the (global) interobserver agreement regarding the recognition of classes III and IV lesions specifically. In addition, we also attempted to determine the sources of possible disagreement among nephropathologists.

Departments of  
\*Pathology and  
||Medical Statistics,  
Leiden University  
Medical Center,  
Leiden, The  
Netherlands;  
†Department of  
Medicine, Imperial  
College, London,  
United Kingdom;  
‡Department of  
Pathology, Necker  
Hospital, French  
National Institute of  
Health and Medical  
Research (INSERM)  
U845, Paris, France;  
and §Nephropathology  
Center, San Gerardo  
Hospital, Monza, Italy

## Correspondence:

Dr. Suzanne  
Wilhelmus, Leiden  
University Medical  
Center, Department of  
Pathology, L1-Q, PO  
Box 9600, 2300 RC  
Leiden, The  
Netherlands. Email:  
s.wilhelmus@lumc.nl

## Materials and Methods

### Case Selection and Survey

Three trained nephropathologists carefully selected 30 glomeruli from several randomly chosen biopsies from classes III and IV lupus nephritis cases, thus obtaining a representative sample of the various lesions that occur in lupus nephritis (except membranous lupus nephritis). All of the biopsies selected had high staining quality and suitable section thickness, and they were selected from the archives at Leiden University Medical Center. All biopsies were handled in a coded and anonymized fashion according to the Dutch National Ethical guidelines (Code for Proper Secondary Use of Human Tissue, Dutch Federation of Medical Scientific Societies). High-quality pictures of these glomeruli were included in the questionnaire as a PDF file. One half of the pictures were images of silver-stained glomeruli, and one half were images of periodic acid–Schiff-stained glomeruli, because these stains are commonly used to assess glomerular pathology. The entire membership (approximately 360 regular members) of the ISN/RPS was provided with a link to the questionnaire with the exception of the pathologists who selected the glomeruli. We received a total of 34 responses from the following countries: Australia (2), Canada (3), India (3), Italy (1), Japan (1), The Netherlands (3), Poland (1), Romania (1), Spain (1), Thailand (2), the United Kingdom (2), and the United States (14). For each picture of a glomerulus, the participants were asked to decide whether the glomerulus contained a lesion that would classify the biopsy as either class III or IV. If the answer was no, the participants were then asked whether other lesions were present, after which they could move on to the next picture; if the answer was yes (*i.e.*, the biopsy could be classified as class III/IV), they were asked to score additional parameters, which are shown in Figure 1. The participants were also encouraged to provide comments regarding each glomerulus. Lastly, the respondents were asked to indicate their level of experience as a nephropathologist (*i.e.*, new to the field, moderately experienced, or highly experienced).

In addition to this self-assessment, the participating pathologists were asked how many years they had practiced as a renal pathologist, how many native biopsies they see each year, what percentage of these biopsies was diagnosed as lupus nephritis, and how many lupus nephritis cases they evaluated in the context of research. The answers to these questions were then combined into a single value that was used to estimate the total number of lupus biopsies evaluated by each pathologist before participation in this study.

### Statistical Procedures

For each outcome parameter, the  $\kappa$ -value (10) or intra-class correlation (ICC) value (11) was calculated to measure the degree of interobserver agreement (0=no agreement, 1=perfect agreement). To calculate these values, we compared the answers given by the participants rather than comparing the answers with a gold standard. ICCs were calculated using a mixed model to estimate the variance components of the ICC. If there were more than two nominal (nonordinal) categories, an unweighted  $\kappa$ -value was calculated (in such cases, the ICC would be less appropriate, because ICC values imply quadratic weights for differences in agreement). The ICC or  $\kappa$ -value was also calculated for two separate levels of experience (specifically, nephropathologists who were new to the field/moderately experienced [less experienced] and nephropathologists who were highly experienced) and different continents. We used a sign test (exact variant) to test the null hypothesis that, within the set of parameters, the direction of the difference between less experienced and highly experienced pathologists would be random. A  $\kappa$  or ICC of <0.4, 0.4–0.6, 0.6–0.8, or >0.8 was considered to reflect poor, moderate, good, or excellent agreement, respectively (8). For the analysis, if no lesion was present that would categorize the biopsy as either class III or class IV, the other parameters were also considered absent unless otherwise specified by the responding pathologist. A separate analysis was performed on the presence of endocapillary proliferation with respect to the presence of swelling of

Does this glomerulus contain a lesion which would put the biopsy in class III or IV?		<input type="checkbox"/> Yes	<input type="checkbox"/> No
If so, please also indicate			
Active and/or chronic	<input type="checkbox"/> Active	<input type="checkbox"/> Chronic	<input type="checkbox"/> Active/chronic
Segmental or global	<input type="checkbox"/> Segmental	<input type="checkbox"/> Global	
Endocapillary proliferation	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
Influx of inflammatory cells	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
Swelling of endothelial cells	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
Extracapillary proliferation	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
If present	<input type="checkbox"/> Cellular	<input type="checkbox"/> Fibrocellular	<input type="checkbox"/> Fibrous
Wire loops	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
Fibrinoid necrosis	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
Karyorrhexis	<input type="checkbox"/> Present	<input type="checkbox"/> Absent	
Other lesions?	<input type="checkbox"/> Yes	<input type="checkbox"/> No	If yes, please describe <input type="text"/>
Comments	<input type="text"/>		

**Figure 1.** | Scoring form sent to the membership of the Renal Pathology Society. Respondents were asked to score 30 images containing one glomerulus each.

endothelial cells and influx of inflammatory cells. For this analysis, the total number of observations (calculated by multiplying the number of participants by the number of glomeruli) for endocapillary proliferation was used (*i.e.*, the total number of times that the question about endocapillary proliferation was answered with either absent or present). The answers to the (sub-) questions on swelling of endothelial cells and influx of inflammatory cells were related to that number of observations (in percentages). In addition, we provided the number of pathologists who gave a specific combination of answers at least one time.

We used the independent samples Mann–Whitney *U* test to compare the total number of biopsies with lupus nephritis evaluated by less experienced pathologists (*i.e.*, new to the field and moderately experienced pathologists) with the total number of biopsies with lupus nephritis evaluated by more experienced pathologists (highly experienced). This non-parametric test was used, because the data were not distributed normally. A *P* value <0.05 was considered statistically significant. All analyses were performed using SPSS 20.0 (IBM, Armonk, NY).

## Results

The ICC for the presence of a lesion that would classify the biopsy as class III or IV was 0.39, which is relatively poor. The  $\kappa$ /ICC values for the additional parameters were as follows: active, chronic, or both: 0.36; segmental (S) versus global (G): 0.39; endocapillary proliferation: 0.46; influx of inflammatory cells: 0.32; swelling of endothelial cells: 0.46; extracapillary proliferation: 0.57; type of crescent: 0.46; and wire loops: 0.35. Fibrinoid necrosis and karyorrhexis were excluded from the analysis, because these two parameters lacked sufficient variance in our cohort to calculate a reliable  $\kappa$  or ICC value. The highly experienced nephropathologists (*n*=19) had higher interobserver agreement for all parameters compared with the less experienced nephropathologists (*n*=15; five nephropathologists were new to the field, and 10 nephropathologists were moderately experienced; *P*=0.004) (Table 1). Before their participation in this study, the pathologists who considered themselves to be highly experienced had evaluated significantly more lupus nephritis biopsies (median=525 biopsies per pathologist) than less experienced

pathologists who considered themselves to be either new to the field or moderately experienced (median=128 biopsies per pathologist; *P*=0.002 versus highly experienced pathologists) (Supplemental Figure 1).

In view of the regional variations in the prevalence of lupus (nephritis) (12,13), we examined whether interobserver agreement differs between continents. We found that the between-continent differences were not consistent for all parameters and seemed to primarily reflect the relative proportion of highly experienced nephropathologists in each continent (Supplemental Table 1).

Our study had a relatively low response rate (9.4%). To test for a possible response bias on the basis of each participant's country of origin, we compared the distribution of participants throughout the six continents with the distribution of RPS members in general and found similar distributions (Supplemental Table 2).

We next analyzed the role, as perceived by the participants, of the swelling of endothelial cells and the influx of inflammatory cells in endocapillary proliferation. This analysis revealed that, in 23% of the times that a glomerulus was considered to have endocapillary proliferation, it did not include the presence of inflammatory cell influx. In 15% of the times that a glomerulus was designated by the participant as lacking endocapillary proliferation, inflammatory cell influx was marked as being present but apparently not considered to be representative of endocapillary proliferation. In an additional 22% of the times that a glomerulus was designated as lacking endocapillary proliferation, swelling of endothelial cells was noted by the participant (Table 2).

Figure 2 shows three example images of glomeruli. For each glomerulus, a selection of the scoring results is shown, including a selection of comments made by the responding participants.

## Discussion

The aims of our study were to measure global interobserver agreement regarding the identification of class III/IV lesions in lupus nephritis and determine possible causes of disagreement. We found that agreement among 34 pathologists from 12 different countries was poor, leaving considerable room for

**Table 1. Intraclass correlation and  $\kappa$ -values for all parameters**

Parameter	Statistical Method	Overall <sup>a</sup>	Level of Experience	
			New or Moderate <sup>a</sup> ( <i>n</i> =15)	High <sup>a</sup> ( <i>n</i> =19)
Class III/IV	Intraclass correlation	0.39	0.33	0.43
Active and/or chronic	$\kappa$	0.36	0.27	0.38
Segmental or global	$\kappa$	0.39	0.27	0.46
Endocapillary proliferation	Intraclass correlation	0.46	0.42	0.50
Influx of inflammatory cells	Intraclass correlation	0.32	0.19	0.45
Swelling of endothelial cells	Intraclass correlation	0.46	0.40	0.52
Extracapillary proliferation	Intraclass correlation	0.57	0.42	0.71
Type of crescent	$\kappa$	0.46	0.40	0.55
Wire loops	Intraclass correlation	0.35	0.27	0.41

<sup>a</sup>Excellent, >0.8; good, 0.6–0.8; moderate, 0.4–0.6; poor, <0.4.

**Table 2. Inflammatory cell influx and swelling of endothelial cells in endocapillary proliferation as scored by the participants**

Endocapillary Proliferation	Swelling of Endothelial Cells	Inflammatory Cell Influx	Observations <i>n</i> (%)	Pathologists <sup>b</sup> <i>n</i>
Present (observations; <i>n</i> =535 <sup>a</sup> )	–	–	10 (2)	5
	+	–	115 (21)	24
Absent (observations; <i>n</i> =65 <sup>a</sup> )	–	+	48 (9)	17
	+	+	362 (68)	34
	–	–	41 (63)	18
	+	–	14 (22)	4
	–	+	3 (5)	2
	+	+	7 (10)	4

<sup>a</sup>The total number of observations for endocapillary proliferation (calculated by multiplying the number of participants by the number of glomeruli) if answered with either absent or present.

<sup>b</sup>Number of participating pathologists (total *n*=34) who gave this combination of answers at least one time.

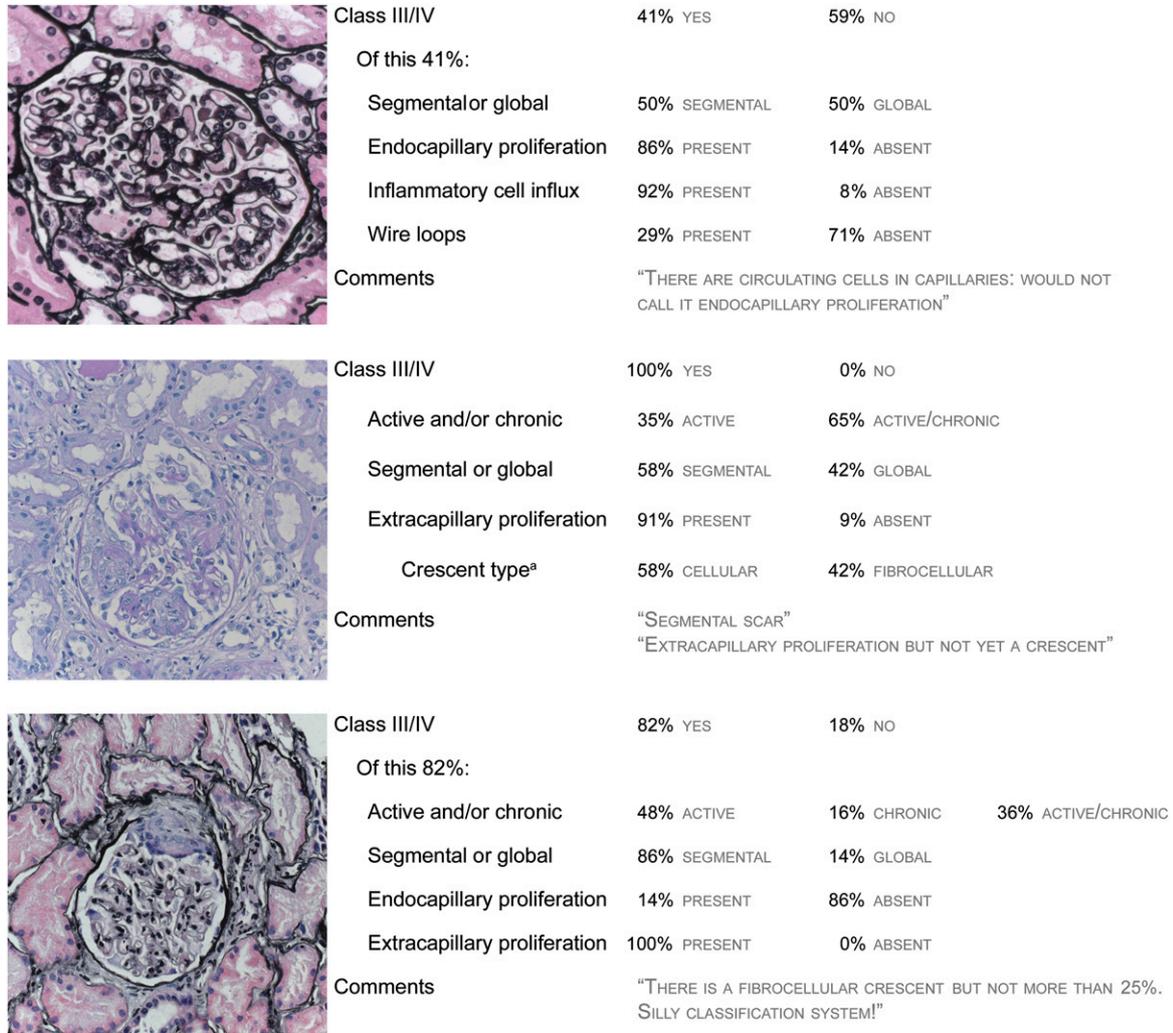
improvement. Interestingly, the responding nephrologists with more experience had higher agreement than less experienced pathologists. However, even within the subgroup of highly experienced nephrologists, the agreement was only moderate at best.

Although the histologic classification system for lupus nephritis is nearly 40 years old, remarkably few studies have been performed to investigate interobserver agreement using this classification system. The most recent studies of interobserver agreement (7–9) have been done in the aftermath of the introduction of the revised ISN/RPS classification system in 2003. In contrast to these earlier studies, we decided to focus on interobserver agreement on the lesions of class III/IV—instead of the entire classification, including classes I, II, V, and VI—because the recognition of these lesions has the highest clinical relevance. However, the results from our study can be compared with previous studies regarding the reproducibility of the activity and chronicity indices in lupus nephritis. Two studies from the 1990s examining interobserver and intraobserver agreement showed a low reproducibility of the (components of the) indices. Schwartz *et al.* (14) attributed this finding primarily to differences in interpreting the components of the indices among four participating expert pathologists. Wernick *et al.* (15) attributed the low reproducibility to a lack of experience of the pathologists in their study's nonacademic setting. Wernick *et al.* (15) also noted that the most experienced pathologist in their study had the highest intraobserver reliability; this effect of experience is consistent with our results. Moreover, as in our study, the components of the indices relating to endocapillary proliferation were less reproducible than the components relating to extracapillary proliferation.

To measure interobserver agreement regarding class III/IV lesions in lupus nephritis, we chose to use pictures of individual glomeruli rather than biopsy slides. The main advantage of this approach is that it minimizes the sources of variability that are inherent to looking at whole biopsies (*i.e.*, each pathologist looking at a different glomerulus). Our approach also allowed us to focus specifically on the reproducibility of class III/IV lesion identification and examine possible causes of variability in higher detail. However, because of the nature of our study, a drawback is that there is a

bias in the selection of the glomeruli used; however, we made every effort to include representative samples of the various lesions that occur in lupus nephritis. Also, we noted that there may have been a response bias, because only 9.4% of all ISN/RPS members responded. However, the participants in this study seemed to reflect the ISN/RPS membership reasonably well with respect to the participants' distribution among the different continents and the wide range of experience levels among our participants. Finally, our results cannot be directly extrapolated to the reproducibility of the entire ISN/RPS classification system for lupus nephritis, because classifying an image of a glomerulus differs from classifying renal biopsy specimens that contain multiple glomeruli cut at multiple levels and stained with several different stains. Because we did not study whole biopsies, we can only speculate about the effect of our results on the agreement in clinical practice. On one hand, the agreement with respect to particular features of the biopsy could be worse, because instead of having only one glomerulus on which to agree, there would be many glomeruli. Using the Banff classification system of renal transplant pathology, Furness *et al.* (16) reported that reproducibility increased when pictures were assessed rather than an entire biopsy. On the other hand, agreement with respect to class assignment may be higher: when one or a few glomeruli contain a clear lesion, the effect of other less clear lesions would be diminished. Nevertheless, this does not detract from the value of this study in revealing that there is considerable room for improvement in the identification of lesions belonging to class III/IV lupus nephritis.

Several factors could have caused interobserver variation, including a lack of objectivity, technical variability, the participating pathologist's experience, ambiguous definitions of lesions, and nonadherence to the classification methodology. In this study, a lack of objectivity and technical variability likely played only a minor role. Objectivity was ensured by blinding the participants to the clinical data. Technical variability was minimized by using sections that were cut and stained in the same laboratory and distributing the same pictures of glomeruli to all participants. However, the participating nephrologists' experience may have played an important role, because more experienced nephrologists have higher interobserver agreement than



**Figure 2.** | Examples of glomeruli with poor agreement between observers and a selection of the scoring results. Left panel shows pictures of three glomeruli (two silver-stained and one periodic acid-Schiff-stained), and right panel shows the summary responses from 34 respondents for each glomerulus. <sup>a</sup>If extracapillary proliferation is present.

less experienced nephrologists. Finally, the role of ambiguous lesion definitions and nonadherence to classification methodology are discussed below.

To characterize a lesion as belonging to class III/IV, one of the challenges is to decide whether endocapillary proliferation is present; in this respect, ICC was only 0.46. When studying how the influx of inflammatory cells and endothelial cell swelling were scored in this context, we found that, in 23% of the times that a glomerulus was perceived to have endocapillary proliferation, influx of inflammatory cells was marked as absent and that, in 9% of these times, the designation of endocapillary proliferation was solely on the basis of the influx of inflammatory cells. Moreover, in cases where participants considered endocapillary proliferation to be absent, they still marked influx of inflammatory cells to be present in 15% and endothelial cell swelling to be present in 22% of cases. Thus, it seems to be unclear how to interpret the definition of endocapillary proliferation. In the ISN/RPS 2003 classification system (6), endocapillary proliferation is defined two times. The first

definition is "endocapillary hypercellularity due to increased numbers of mesangial cells, endothelial cells, AND infiltrating monocytes, AND causing narrowing of the glomerular capillary lumina" (6), and the second definition is "endocapillary hypercellularity with OR without leukocyte infiltration AND with substantial lumen reduction" (6). These two definitions have in common that a lumen reduction is required for endocapillary proliferation. It is, however, unclear how narrowed the lumina should be, which might explain why, in 37% of the observations with perceived absence of endocapillary proliferation, influx of inflammatory cells and/or endothelial cell swelling were marked as present. However, these two definitions differ with respect to the composition of the endocapillary hypercellularity. In contrast to the first definition, the second definition states that an influx of inflammatory cells is not necessarily a part of endocapillary proliferation. It is also not clear if the influx of inflammatory cells alone is sufficient. This ambiguity seems to be reflected in the abovementioned scoring of endocapillary proliferation by the participants.

Finally, the question remains if it is even possible to reliably distinguish between the different components of endocapillary proliferation.

The agreement on the distribution of lesions within the glomeruli (*i.e.*, S or G) was poor ( $\kappa$ -value=0.39). S lesions are currently defined as “involving less than half of the glomerular tuft” (6). However, this definition does not tell the pathologist how to account for extracapillary proliferation (which occurs outside the tuft). Furthermore, wire loops that are not obviously G are difficult to incorporate in this context. Moreover, the relevance of distinguishing between S and G lesions has been a subject of debate. For example, Haring *et al.* (17) performed a meta-analysis and found no difference in clinical outcome between patients with class IV-S lesions and patients with class IV-G lesions. However, other groups suggest that a biologic difference exists between IV-S and IV-G and argue that the distinction should remain in the classification system (18). If this distinction remains, more explicit definitions should be devised to make the distinction both reliable and reproducible.

Finally, extracapillary proliferation and the designation of lesions as active or chronic (or both) caused confusion among the respondents. First, although the classification system states that extracapillary proliferation should occupy at least one quarter of the glomerular capsular circumference to qualify as extracapillary proliferation, only a few of the responding pathologists used this criterion in their scoring (even in the highly experienced pathologists group). Second, although fibrocellular crescents are designated as active lesions, many respondents seemed to interpret them as chronic or active/chronic lesions. Third, although double contours are not listed as chronic lesions in the classification system, some of the respondents apparently perceived them as such.

Although the introduction of the 2003 ISN/RPS classification system significantly improved interobserver agreement relative to the 1995 WHO system, our results indicate that there is still considerable room for improvement in the identification of lesions (in individual glomeruli) belonging to class III/IV lupus nephritis. Improving interobserver agreement—particularly with respect to the presence of class III/IV lesions—has high clinical relevance, because correctly identifying the histologic class plays an essential role in deciding whether to initiate immunosuppressive therapy when treating patients with lupus nephritis. The observation that highly experienced pathologists have higher agreement than less experienced pathologists suggests that agreement can be improved—at least in part—by educating nephropathologists. Moreover, as discussed above, agreement might be improved by revising and clarifying some of the definitions in the current classification system. Finally, our results underscore the need for a central review of biopsies in clinical trials by a minimum of two experienced nephropathologists.

#### Acknowledgments

We thank the Renal Pathology Society, particularly H.K. Singh, and all of the participating members for their contribution to this study. We also thank M. Mengel for reviewing the manuscript.

The results presented in this paper have been presented in abstract form at the Annual American Society of Nephrology Kidney Week, Atlanta, Georgia, November 5-10, 2013.

#### Disclosures

None.

#### References

- Bernatsky S, Boivin JF, Joseph L, Manzi S, Ginzler E, Gladman DD, Urowitz M, Fortin PR, Petri M, Barr S, Gordon C, Bae SC, Isenberg D, Zoma A, Aranow C, Dooley MA, Nived O, Sturfelt G, Steinsson K, Alarcón G, Senécal JL, Zummer M, Hanly J, Ensworth S, Pope J, Edworthy S, Rahman A, Sibley J, El-Gabalawy H, McCarthy T, St Pierre Y, Clarke A, Ramsey-Goldman R: Mortality in systemic lupus erythematosus. *Arthritis Rheum* 54: 2550–2557, 2006
- Bono L, Cameron JS, Hicks JA: The very long-term prognosis and complications of lupus nephritis and its treatment. *QJM* 92: 211–218, 1999
- Bertsias GK, Tektonidou M, Amoura Z, Aringer M, Bajema I, Berden JHM, Boletis J, Cervera R, Dörner T, Doria A, Ferrario F, Floege J, Houssiau FA, Ioannidis JPA, Isenberg DA, Kallenberg CGM, Lightstone L, Marks SD, Martini A, Moroni G, Neumann I, Praga M, Schneider M, Starra A, Tesar V, Vasconcelos C, van Vollenhoven RF, Zakharova H, Haubitz M, Gordon C, Jayne D, Boumpas DT; European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association: Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis. *Ann Rheum Dis* 71: 1771–1782, 2012
- Hahn BH, McMahon MA, Wilkinson A, Wallace WD, Daikh DI, Fitzgerald JD, Karpouzas GA, Merrill JT, Wallace DJ, Yazdany J, Ramsey-Goldman R, Singh K, Khalighi M, Choi SI, Gogia M, Kafaja S, Kamgar M, Lau C, Martin WJ, Parikh S, Peng J, Rastogi A, Chen W, Grossman JM; American College of Rheumatology: American College of Rheumatology guidelines for screening, treatment, and management of lupus nephritis. *Arthritis Care Res (Hoboken)* 64: 797–808, 2012
- Kidney Disease: Improving Global Outcomes (KDIGO) Glomerulonephritis Work Group: KDIGO clinical practice guideline for glomerulonephritis. *Kidney Int Suppl* 2: 139–274, 2012
- Weening JJ, D’Agati VD, Schwartz MM, Seshan SV, Alpers CE, Appel GB, Balow JE, Bruijn JA, Cook T, Ferrario F, Fogo AB, Ginzler EM, Hebert L, Hill G, Hill P, Jennette JC, Kong NC, Lesavre P, Lockshin M, Looi LM, Makino H, Moura LA, Nagata M: The classification of glomerulonephritis in systemic lupus erythematosus revisited. *J Am Soc Nephrol* 15: 241–250, 2004
- Furness PN, Taub N: Interobserver reproducibility and application of the ISN/RPS classification of lupus nephritis—a UK-wide study. *Am J Surg Pathol* 30: 1030–1035, 2006
- Grootsholten C, Bajema IM, Florquin S, Steenbergen EJ, Peutz-Kootstra CJ, Goldschmeding R, Bijl M, Hagen EC, van Houwelingen HC, Derksen RHW, Berden JHM: Interobserver agreement of scoring of histopathological characteristics and classification of lupus nephritis. *Nephrol Dial Transplant* 23: 223–230, 2008
- Yokoyama H, Wada T, Hara A, Yamahana J, Nakaya I, Kobayashi M, Kitagawa K, Kokubo S, Iwata Y, Yoshimoto K, Shimizu K, Sakai N, Furuichi K; Kanazawa Study Group for Renal Diseases and Hypertension: The outcome and a new ISN/RPS 2003 classification of lupus nephritis in Japanese. *Kidney Int* 66: 2382–2388, 2004
- Silcocks PB: Measuring repeatability and validity of histological diagnosis—a brief review with some practical examples. *J Clin Pathol* 36: 1269–1275, 1983
- Fleiss JL, Cohen J: Equivalence of weighted kappa and intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 33: 613–619, 1973
- Pons-Estel GJ, Alarcón GS, Scofield L, Reinlib L, Cooper GS: Understanding the epidemiology and progression of systemic lupus erythematosus. *Semin Arthritis Rheum* 39: 257–268, 2010
- Feldman CH, Hiraki LT, Liu J, Fischer MA, Solomon DH, Alarcón GS, Winkelmayr WC, Costenbader KH: Epidemiology and sociodemographics of systemic lupus erythematosus and lupus nephritis among US adults with Medicaid coverage, 2000–2004. *Arthritis Rheum* 65: 753–763, 2013

14. Schwartz MM, Lan SP, Bernstein J, Hill GS, Holley K, Lewis EJ; Lupus Nephritis Collaborative Study Group: Irreproducibility of the activity and chronicity indices limits their utility in the management of lupus nephritis. *Am J Kidney Dis* 21: 374–377, 1993
15. Wernick RM, Smith DL, Houghton DC, Phillips DS, Booth JL, Runckel DN, Johnson DS, Brown KK, Gaboury CL: Reliability of histologic scoring for lupus nephritis: A community-based evaluation. *Ann Intern Med* 119: 805–811, 1993
16. Furness PN, Taub N, Assmann KJ, Banfi G, Cosyns JP, Dorman AM, Hill CM, Kapper SK, Waldherr R, Laurinavicius A, Marcussen N, Martins AP, Nogueira M, Regele H, Seron D, Carrera M, Sund S, Taskinen EI, Paavonen T, Tihomirova T, Rosenthal R: International variation in histologic grading is large, and persistent feedback does not improve reproducibility. *Am J Surg Pathol* 27: 805–810, 2003
17. Haring CM, Rietveld A, van den Brand JA, Berden JH: Segmental and global subclasses of class IV lupus nephritis have similar renal outcomes. *J Am Soc Nephrol* 23: 149–154, 2012
18. Hill GS, Delahousse M, Nochy D, Bariéty J: Class IV-S versus class IV-G lupus nephritis: Clinical and morphologic differences suggesting different pathogenesis. *Kidney Int* 68: 2288–2297, 2005

**Received:** April 8, 2014 **Accepted:** September 22, 2014

Published online ahead of print. Publication date available at [www.cjasn.org](http://www.cjasn.org).

This article contains supplemental material online at <http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.03580414/-/DCSupplemental>.

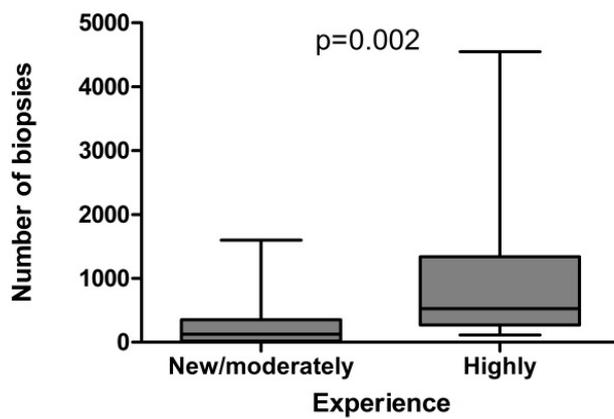
## ONLINE SUPPLEMENTARY DATA

### INTEROBSERVER AGREEMENT ON HISTOPATHOLOGICAL LESIONS IN CLASS III OR IV LUPUS

#### NEPHRITIS

Suzanne Wilhelmus, MD, H. Terence Cook, MD, PhD, Laure-Hélène Noël, MD, Franco Ferrario, MD, PhD, Ron Wolterbeek, MD, Jan A. Bruijn, MD, PhD and Ingeborg M. Bajema, MD, PhD

**Figure S1.** The total number of biopsies with lupus nephritis that were evaluated by pathologists prior to participation in this study, grouped by experience.



**Table S1.** Kappa and ICC values for all parameters, analyzed per continent

Parameter	Statistical method	Overall <sup>a</sup>	Continent			
			Europe <sup>b</sup> (n=9)	Asia <sup>c</sup> (n=7)	Australia <sup>d</sup> (n=2)	North America <sup>e</sup> (n=16)
<b>Class III/IV</b>	ICC	0.39	0.32	0.36	0.61	0.42
<b>Active and/or chronic</b>	Kappa	0.36	0.34	0.25	0.51	0.36
<b>Segmental or global</b>	Kappa	0.37	0.33	0.28	0.49	0.38
<b>Endocapillary proliferation</b>	ICC	0.46	0.45	0.41	0.73	0.47
<b>Influx of inflammatory cells</b>	ICC	0.32	0.43	0.20	0.12	0.31
<b>Swelling of endothelial cells</b>	ICC	0.46	0.51	0.37	0.84	0.46
<b>Extracapillary proliferation</b>	ICC	0.57	0.51	0.38	0.68	0.68
<b>Type of crescent</b>	Kappa	0.46	0.46	0.34	0.61	0.60
<b>Wire loops</b>	ICC	0.35	0.39	0.33	0.57	0.34

ICC, intraclass correlation

<sup>a</sup> >0.8: excellent; 0.6–0.8: good; 0.4–0.6: moderate; <0.4: poor

<sup>b</sup> Europe: 2 new, 2 moderately experienced, and 5 highly experienced nephrologists.

<sup>c</sup> Asia: 6 moderately experienced and 1 highly experienced nephrologist.

<sup>d</sup> Australia: 1 moderately experienced and 1 highly experienced nephrologist. Comparison with other continents was not due to the small number of pathologists from this continent.

<sup>e</sup> North America: 3 new, 1 moderately experienced, and 12 highly experienced nephrologists.

**Table S2.** Distribution of the study participants and the entire RPS membership by continent

	Our study (%)	RPS membership (%)
<b>Europe</b>	26.5	15.2
<b>Asia</b>	20.6	16.3
<b>Australia</b>	5.9	1.6
<b>North America</b>	47.1	62.6
<b>South America</b>	0.0	3.5
<b>Africa</b>	0.0	0.8