Differential Diagnosis of Lupus and Primary Membranous Nephropathies by IgG Subclass Analysis

Young Soo Song,* Kyueng-Whan Min,* Ju Han Kim,† Gheun-Ho Kim,‡ and Moon Hyang Park*

Summary

Background and objectives Previous studies showed that the accuracy of IgG subclasses (ISs) in differentiating membranous lupus nephritis (MLN) from primary membranous nephropathy (PMN) is <80%. This study hypothesized that diagnostic accuracy of ISs would be increased if renal compartment measurements and decision tree analysis are applied.

Design, setting, participants, & measurements Renal biopsy specimens from 41 patients with MLN and 59 patients with PMN between October 2004 and March 2010 were examined, and immunofluorescence staining against IgG1, IgG2, IgG3, and IgG4 as well as C3, C1q, and C4 was evaluated in five different renal compartments (glomerular capillary walls, mesangium, tubules, interstitium, and blood vessels). From IS data, a decision tree to differentiate MLN from PMN was produced (IS decision tree) and its accuracy was compared with that of previous studies. Diagnostic accuracy of the IS decision tree was also compared with that of the complement decision tree as a reference.

Results The demographic information and patterns of IS deposition were similar to those of previous studies. The IS decision tree had, as decision markers, IgG1 in the mesangium and IgG2 and IgG4 along the glomerular capillary wall. The IS decision tree showed higher accuracy (88%) than that of previous studies (<80%) and also that of the complement decision tree (81%).

Conclusions Accuracy of ISs was increased due to the study methods, but the same methodology was less effective using complement measurements. Appropriate data analysis may enhance diagnostic value, but the analysis alone cannot achieve the ideal diagnostic value.


Introduction

Human IgG protein consists of four distinct subclasses (IgG1, IgG2, IgG3, and IgG4) made from each of the four different constant regions of the Ig heavy chain on human chromosome 14 (1,2). These IgG subclasses are different in structure and biologic properties, differently expressed depending on type, degree, and onset of inflammatory response (3–5), and are also associated with several human diseases, including immunodeficiencies and allergic and autoimmune diseases (6–10).

The different patterns of expression of IgG subclasses in lupus nephritis (LN), including membranous LN (MLN) and primary membranous nephropathy (PMN), have been studied. By scoring the glomerular immunofluorescence staining intensities and serum concentration of each member of the IgG subclasses, researchers found that IgG1, IgG2, and IgG3 tended to be highly expressed in LN, whereas IgG1 and IgG4 tended to be highly expressed in PMN (11–14). These findings, however, do not directly indicate that IgG subclasses can be used as a marker differentiating MLN from PMN without considering the diagnostic accuracy. The diagnostic accuracy of IgG subclasses from the data of Imai et al. (11) is 80% if the interactions between the markers are not considered. Improved diagnostic accuracy using IgG subclasses can be obtained by applying better methodologies.

First, diagnostic accuracy can be improved if the immunofluorescence staining intensities of each of the IgG subclasses are measured in five different tissue compartments of the kidney parenchyma: glomerular capillary walls (GCWs), mesangium, tubules (tubular basement membrane [TBM]), interstitium, and blood vessels. In previous studies, it was not considered and only the overall glomerular immunofluorescence intensity was scored (11–14). Different levels of immune complex deposits in the different compartments of the kidney reflect the different mechanisms of action with regard to renal injuries (15).

Second, more accurate decision rules to differentiate MLN from PMN can be obtained by applying data-driven classification models. One of these methods is decision tree analysis, which is a classification tool that uses a tree-like model of decisions on the multivariate data (16,17). The diagnostic accuracy of IgG subclasses may be increased using decision tree analysis.
Considering those backgrounds, it would be necessary to re-evaluate the diagnostic accuracy of IgG subclasses after applying better methodologies. We therefore hypothesized that diagnostic accuracy of IgG subclasses in differentiating MLN from PMN would be increased if renal compartment measurements and decision tree analysis are applied. Furthermore, the significance of our methodologies would be further clarified if a reference marker is introduced for comparisons with IgG subclasses. Thus, our additional hypothesis is that even with the introduction of the better methodologies, not all markers are powerful as IgG subclasses. We selected complement components (C1q, C3, and C4) as a reference because they are conventional markers used for differential diagnosis in immunofluorescence. Then, as a reference, we can compare the diagnostic accuracies between IgG subclasses and complement components after the same methodologies were applied. Overall, the effects and the significances of selection of the methodologies were investigated in this research.

Materials and Methods

The overall research scheme was designed in three parts. First, experimental data for IgG subclasses and complements for each renal compartment were generated. Second, the patterns of data were explored using various visualization methods such as heatmaps and scattergrams for more subtle pattern recognition. Finally, using decision tree analysis, we tested whether the accuracy of IgG subclasses was improved compared with the previous study and whether diagnostic accuracy of IgG subclasses was higher than that of complements.

Experimental Procedures for IgG Subclass and Complement Profiles

Patient Selection. From October 2004 to March 2010, there were 1299 patients with renal biopsies performed at Hanyang University Hospital. This study included 41 patients with MLN among 183 patients with LN and 59 patients with PMN. Ethnically, all of the patients were northeast Asians. MLN was diagnosed when the histologic findings were suitable to class V LN according to the International Society of Nephrology/Renal Pathology Society 2003 classification of LN (18) and antinuclear antibodies were detected at the time of onset or during the progress of the disease. PMN was diagnosed when the histologic findings were suitable to the diagnostic criteria of PMN and no secondary etiologies attributable to the occurrence of MN such as SLE, hepatitis B, or malignancy were found at the onset or during the progress of the disease. Clinical findings of each of these patients were collected from medical records and pathology consultation sheets.

Specimen Preparation and Staining. Renal needle biopsy specimens were divided and processed for light microscopy, immunofluorescence, and electron microscopy. For light microscopy, renal tissue was fixed in Duboscq–Brasil solution and embedded in paraffin. Serial sections were cut at 2 μm and stained with hematoxylin and eosin, periodic acid–Schiff, Masson’s trichrome, and methylene blue. For immunofluorescence, fragments of the biopsy were snap-frozen and cut with a cryostat at 3–4 μm and stained with fluoresceinated antiserum monospecific for IgG, IgM, IgA, C3, C1q, C4, k and λ, fibrinogen and albumin (1:30; DakoCytomation, Denmark), IgG1, IgG2, IgG3, and IgG4 (1:30; Sigma-Aldrich Inc, St. Louis, MO). Immunofluorescence sections were examined with an Olympus BX51 microscope. For electron microscopy, the tissue was diced into 1-mm cube fragments, fixed in chilled 2.5% glutaraldehyde in a cacodylate buffer, washed in buffer, postfixed in 1% osmium tetroxide, dehydrated in graduated concentrations of alcohol, and finally embedded in Epon. Ultrathin sections were stained with lead citrate and examined with a Hitachi H-7600 electron microscope at 80 kV.

Renal IgG Subclasses and Complements. Immunofluorescence staining for IgG subclasses and complements was evaluated according to the fluorescence intensity at the following five different renal compartments: the peripheral GCW, mesangium, TBM, interstitium, and blood vessel wall. The intensity of immunofluorescence staining was scored on a scale of 0–3+ as follows: 0, negative; 1, weak; 2, moderate; and 3, strong (positive for each biopsy specimen). All slides were evaluated in a blinded fashion by two pathologists (M.H.P. and Y.S.S.) and discordant patients were reviewed to achieve a consensus.

Visualization and Descriptive Statistical Analyses

To explore the patterns of distribution of IgG subclasses in MLN and PMN, descriptive statistical methods were used and data were displayed using various visualization methods, including a heatmap (19) and a scattergram. In making a heatmap, a dendrogram was accompanied. The dendrogram shows grouping of patients at various scales according to the similarity of IgG subclass profile. The IgG subclass profile of each patient was represented as a 20-dimensional vector, with each immunofluorescence intensity as a value. The similarity between two patients is determined by the distance between the corresponding vectors; thus, each of the patients is grouped with the most similar case. Each of these two groups is then grouped according to the similarity between groups. This process was repeated until one big group is made. The mean immunofluorescence intensities of the 20 markers (4 subclasses × 5 renal compartments) were compared with the t test, between MLN and PMN (R2.14.1; R Development Core Team, Vienna). We investigated sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of each diagnostic criteria generated by changing the range of the intensities of each marker (Supplemental Appendix 1). Accuracy for each diagnostic criterion is defined as the proportion of correctly classified patients over the total cases.

Decision Tree Analysis

Two decision trees using the C5.0 algorithm (Clementine 12.0; SPSS, CA) (16,17) were made each using data from IgG subclasses and complements. Mean diagnostic accuracies were also calculated by 10-fold cross-validation. Improvement of diagnostic accuracy of IgG subclasses by the introduction of decision tree analysis was investigated. Then diagnostic accuracies between two decision trees were compared with test marker dependency of the decision tree analysis.
The diagnostic accuracies of the previous study were not explicitly described, but obtaining those values from data published by Imai et al. were straightforward (11). Because immunofluorescence intensity of each marker was semiquantitatively measured from 0 to 4, we could exhaustively generate all of the possible diagnostic criteria of one disease by changing the range of intensities of each marker, making the diagnostic criteria of the other disease automatically determined. For each diagnostic criterion, diagnostic accuracies were calculated by calculating the proportion of accurately classified patients over the total cases. All of the diagnostic accuracies were ≥80%.

Ten-fold cross-validation is a statistical estimation using the current data set on how accurately a predictive model will perform in an independent data set (16). The original data are randomly partitioned into 10 parts. Nine subsamples are designated as training sets, which will make a model, and the remaining subsample is designated as a test set, which will test the model. Another nine subsamples and the remaining subsample then become a training set and a test set, respectively. These processes are repeated until all of the 10 subsamples are used as a test set. For each of these steps, the performance of each model can be measured and we can obtain the distribution of them. The other K-fold cross-validation can be applied to this problem, but 10-fold cross-validation is the most commonly used method.

**Results**

**IgG Subclass and Complement Profiles**

**Patient Demographics and Clinical Findings.** The patient demographics and clinical findings from the MLN and PMN patients were investigated to confirm the representativeness of the selected patients (Table 1). The patients with MLN were predominantly female, whereas patients with PMN were predominantly male. The patients with MLN manifested the disease at a younger age than those with PMN. The levels of 24-hour urine protein were higher in patients with PMN than those with MLN.

**Immunofluorescence Staining for IgG Subclasses.** Morphologic impression of immunofluorescence staining for IgG subclasses in the patients with MLN and PMN were very similar to the descriptions in previous studies. In summary, in MLN patients, immunofluorescence staining for IgG subclasses showed that IgG1, IgG2, and IgG3 tended to be strongly deposited in the subepithelial portion along the GCW with focal small deposits in the mesangium and a few unusual weak IgG4 deposits (Figure 1, A–D). In the patients with PMN, positive immunofluorescence staining for IgG1 and IgG4 tended to be intense in the

<table>
<thead>
<tr>
<th>Table 1. Patient demographics and clinical findings at renal biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MLN (n=41)</strong></td>
</tr>
<tr>
<td>Sex (male/female)</td>
</tr>
<tr>
<td>Age (yr)</td>
</tr>
<tr>
<td>Serum creatinine concentration (mg/dl)</td>
</tr>
<tr>
<td>Creatinine clearance (ml/min)</td>
</tr>
<tr>
<td>24-h urine protein (mg/dl)</td>
</tr>
</tbody>
</table>

Data are presented as median (interquartile range). MLN, membranous lupus nephritis; PMN, primary membranous nephropathy.

![Figure 1.](https://example.com/figure1.png) | Immunofluorescence staining for IgG subclasses. A representative biopsy case of lupus nephritis, class V (A–D) and primary membranous nephropathy (E–H). Note the positive immunofluorescence staining for IgG1, IgG2, and IgG3 in the subepithelial portion along the glomerular capillary wall and in the mesangium (A–C) and the complete negative for IgG4 in the mesangium (D). A case of primary membranous nephropathy (E–H) shows strong immunofluorescence staining for IgG1 and IgG4 in the subepithelial portion along the glomerular capillary wall but negative in the mesangium.
subepithelial portion along the GCW with absent mesangial deposition (Figure 1, E–H). Deposition of these markers in extravascular locations such as TBM, interstitium, and blood vessels was rarely found in patients with class V LN (MLN) compared with class IV patients. The full IgG subclass and complement profiles can be downloaded in Supplemental Appendix 2.

Visualization and Descriptive Statistical Analyses
Visualization with Heatmaps and Scattergrams. Patterns of data could be more easily identified through heatmaps (Figure 2) and scattergrams (Figure 3 and Supplemental Appendix 3). Identified patterns of IgG subclass deposition can be summarized as follows (Figures 2A and 3). First, mesangial deposition was absent or very rare in PMN. Second, IgG1 deposition in the subepithelial portion along the GCW was a common feature of both diseases. Third, IgG2 and IgG3 were usually deposited in patients with MLN in the mesangium and/or GCW. Fourth, IgG4 was usually deposited along the GCW in PMN cases, but not in the cases of MLN. Finally, deposition of any IgG subclasses in the TBM interstitium or vessel wall was rare. There was a cluster of PMN with higher intensities of IgG4 deposition along the GCW. The other patients with PMN showed weak intensities for all markers overall.

Figure 2. Heatmap for immunofluorescence intensities on membranous lupus nephritis (LN) (41 patients) and primary membranous nephropathy (PMN) (59 patients) with a dendrogram (left side of the heatmap) and a color key and histogram (bottom of the heatmap) from IgG subclasses (A) and complements (B). The x-axis of the heatmap represents a list of markers and y-axis represents a list of patients, each of which is either LN or PMN. The patients are reordered and grouped according to the similarity of patients as shown in the dendrogram. Each cell represents an immunofluorescence intensity of the corresponding marker and patient, whose value can be 0, 1, 2, or 3 and its value is color coded from red through white as shown in the color key. The histogram shows counts of the cell values in the heatmap. Labels of rows are either LN or PMN, which is represented by the color bar along the label (LN is red, PMN is blue). The heatmap shows distinctive patterns of the overall distribution of the IgG subclasses for the patients with LN and PMN in the glomerular capillary wall (G), mesangium (M), tubular basement membrane (T), interstitium (I), and vessel (V). LN patients tend to show strong immunofluorescence staining for IgG1, IgG2, and IgG3 along the glomerular capillary wall and/or in the mesangium, whereas PMN patients tend to show strong immunofluorescence staining for IgG1 and IgG4 along the glomerular capillary wall. PMN patients are largely grouped into two.
When complement profiles were investigated in the same way, C1q tended to be expressed more in the patients with MLN along the GCW and/or in the mesangium (Figure 2B and Supplemental Appendix 3).

Statistical Analyses of the Immunofluorescence Intensity of the IgG Subclasses and Complements. When the immunofluorescence staining intensities of the IgG subclasses between MLN and PMN were compared (Table 2), the mean intensities of IgG1, IgG2, and IgG3 were significantly higher in the MLN patients than for PMN both in the GCW and mesangium. When the immunofluorescence intensities of complement components were compared between MLN and PMN (Table 2), the mean intensities of C3, C4, and C1q were also higher along the GCW and in the mesangium in MLN than in PMN. Among complement components, C3 was mostly stronger along the GCW as well as in the mesangium in MLN. In PMN, C3 was strongly positive along the GCW, but negligible in the mesangium. Mean intensities of these markers in TBM, interstitium, and blood vessels were very low.

The diagnostic performances of the individual markers including IgG subclass and complement deposits in the diagnosis of MLN differentiated from PMN were compared (Supplemental Appendix 4). Among these markers, positive (immunofluorescence intensities 1, 2, 3) expression of IgG1 in the mesangium, IgG2 along the GCW and mesangium, C3 in the mesangium, and C1q along the GCW and in the mesangium, and negative (immunofluorescence intensities 0) expression of IgG4 along the GCW showed notable accuracies in the differential diagnosis of MLN from PMN. The highest diagnostic accuracy was 86% found in IgG1 in the mesangium, which is higher than the highest diagnostic accuracy in the previous studies (80%). These findings suggest that IF intensities should be measured at each renal compartment to increase the diagnostic power in differentiating MLN from PMN.

Decision Tree Analyses

Decision tree analysis was performed in two datasets (IgG subclass profile, complement profile) (Figure 4, A and
B). The decision tree using IgG subclass profile was computationally produced by the C5.0 algorithm and the decision rules were set as follows (the decision rules are equivalent to that shown in Figure 4A). A patient is diagnosed with MLN if (1) IgG1 is positive in the mesangium or (2) IgG1 is negative in the mesangium, and IgG2 is positive and IgG4 is negative along the GCW. A patient is diagnosed with PMN if (3) IgG1 is negative in the mesangium and IgG2 is negative or (4) IgG1 is negative in the mesangium, and IgG2 and IgG4 are positive along the GCW.

After 10-fold cross-validation, the mean accuracies were estimated to be 88% with 3.3% of SEM, which is higher than the highest diagnostic accuracy of IgG subclasses in the previous study (80%) (11). The decision tree using complements showed lower estimated accuracy (81.0±3.8%) than that of the decision tree using IgG subclasses, which were statistically significant ($P=0.0004$) (Table 3). When the combined IgG subclass and complement profile was used to make a decision tree, the accuracy (92.0±3.3) was more increased than that of IgG subclasses alone, but it was statistically insignificant (Table 3 and Supplemental Appendix 4).

Discussion
In this study, we proved that previously described diagnostic accuracy of IgG subclasses in differentiating MLN from PMN can be increased by applying immunofluorescence intensity measurement for each renal compartment and decision tree analysis. This result signifies that the diagnostic value of IgG subclasses has been underestimated and that appropriate methodologies can recover their real diagnostic value. We also demonstrated...
that this enhancement might be dependent on the marker by showing that the diagnostic accuracy of IgG subclasses was higher than that of complements even when the same methodologies were applied.

Selection of appropriate markers and analytical methods resulted in these significant results. Previous research showing the distinct patterns of IgG subclasses between MLN and PMN (11–14) led us to select these markers. Decision tree analysis was selected because complex patterns can be easily simplified in this approach, the results are easy to understand, and it is computationally feasible. Because markers and analytical methods were not thoroughly investigated, we cannot exclude the possibility that there might be better combinations of markers and analytical methods than those presented in this study. For this, further research will be required.

The decision tree results of the IgG subclasses are consistent with the pathogenesis of MLN and PMN. First, immune deposits can be found only along the GCW in PMN; however, they can be found in some amount in the mesangium as well in MLN. In the decision trees, the node that was positive for markers in the mesangium was diagnosed as MLN, whereas a TH2 response is predominant in PMN and these responses are mediated by IgG subclasses (20, 21). IgG2 and IgG3 were known to be produced in TH1 response and IgG4 in TH2 response (20, 21). In the decision trees, nodes positive for TH1-related isotypes or negative for TH2-related isotypes were diagnosed as MLN, whereas nodes positive for TH2-related isotypes or negative for TH1-related isotypes were diagnosed as PMN. The consistency of the decision tree results with the current understanding of the pathogenesis of MLN and PMN in relation to IgG subtype and localization of immune deposits support our conclusions.

Most LN patients other than class V (MLN) showed similar patterns of IgG subclass profiles as shown in MLN,
exhibiting strong positive immunofluorescence staining for IgG1, IgG2, and IgG3 and negative for IgG4 in the subendothelial portion along the GCW and in the mesangium as well. Deposition of these markers in extraglomerular locations such as TBM, interstitium, and blood vessels was rarely found in patients with MLN compared with class IV patients. No extraglomerular IgG subclass deposits were seen in PMN patients.

Our decision tree of IgG subclasses can be applicable in certain clinical situations. Patients with early stage MLN were found to be antinuclear antibody–negative in 35% of cases (22). There are no definitive diagnostic procedures for these patients. In this background, the usefulness of IgG subclass deposition can be tested in further studies. The possibility of ethnic dependence of IgG subclass deposition should be considered. In African-American and Caucasian patients with LN, interstitial inflammation was more common in African Americans than Caucasians (23). In this study, IgG staining was present in the TBM in 62% of patients and in the vascular walls in 48% of patients compared with 20% and 12% of patients with MLN in our Korean patients, respectively. These results should prompt research on IgG subclass profiles in common renal diseases among different ethnic groups.

In conclusion, we showed that the diagnostic value of IgG subclass deposition in LN and PMN is enhanced if immunofluorescence intensity is measured separately in each renal compartment and decision tree analysis is applied. The immunofluorescence intensities for IgG1 in the mesangium and IgG2 and IgG4 along the GCW were found to be markers effectively differentiating MLN from PMN. Furthermore, this enhanced accuracy could not be achieved by the use of complements although the same methodologies were applied. These results imply that appropriate selection of analysis methods may be as important as the appropriate selection of markers. Our
methods have a wide range of applications in nephrology and renal pathology. Further study should be undertaken to find better combinations of markers and analysis methods to aid in clinical decision making.

Acknowledgments

We thank Dr. James Robinson Hackney for critical review of this article.

Disclosures

None.

References

3. Hammarstro¨m L, Smith CI: IgG subclasses in bacterial infections. Monogr Allergy 19: 122–133, 1986

Table 3. Comparison of the performances of decision tree analysis among IgG subclasses, complements, and both in the differential diagnosis between membranous lupus nephritis and primary membranous nephropathy

<table>
<thead>
<tr>
<th>Marker</th>
<th>SEN (%)</th>
<th>SPC (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>ACC (%)</th>
<th>ACC after Cross-Validation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgG subclasses</td>
<td>92.7</td>
<td>93.2</td>
<td>90.5</td>
<td>94.8</td>
<td>93.0</td>
<td>88.0±3.3</td>
</tr>
<tr>
<td>Complements</td>
<td>85.4</td>
<td>88.1</td>
<td>83.3</td>
<td>89.7</td>
<td>87.0</td>
<td>81.0±3.8</td>
</tr>
<tr>
<td>IgG subclasses + complements</td>
<td>95.1</td>
<td>98.3</td>
<td>97.5</td>
<td>96.7</td>
<td>97.0</td>
<td>92.0±3.3</td>
</tr>
</tbody>
</table>

P<0.0004 for difference between IgG subclasses and complements and P<0.001 for difference between complements and IgG subclasses + complements. SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; ACC, accuracy.