Original Article

**Title:** High collagen I gene expression as an independent predictor of long term adverse renal outcome in lupus nephritis

**Authors:** Chiraporn Tachaudomdach, Surasak Kantachuvesiri, Suwikran Wongpraphairot, Suchin Worawichawong, Pleumjit Tankee, Koset Pinpradap, Suda Riengrojpitak, Chagriya Kitiyakara

**Authors affiliations**
1. Chiraporn Tachaudomdach MSc
   Molecular Medicine Graduate Program, Faculty of Science, Mahidol University, Bangkok, Thailand
   Email: tachaudomdach.c@hotmail.com
2. Surasak Kantachuvesiri M.D.
   Department of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.
   Email: raskh@mahidol.ac.th
3. Suwikran Wongpraphairot M.D.
   Department of Medicine, Faculty of Medicine, Prince of Songkla University, Songkla, Thailand.
   Email: suwikranw@yahoo.com
4. Suchin Worawichawong M.D.
   Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand
   Email: schinw@yahoo.com
5. Pleumjit Tankee M.D.
   Department of Medicine, Vachira Phuket Hospital, Phuket, Thailand. Email: pleumjit@yahoo.com
6. Koset Pinpradap BSc
   Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. Email: ragpp@mahidol.ac.th
7. Suda Riengrojpitak PhD
   Department of Pathobiology, Faculty of Science, Mahidol University, Bangkok, Thailand
   Email: scsrr@mahidol.ac.th
8. Chagriya Kitiyakara M.D.
   Department of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.
   Email: kitiyakc@yahoo.com

**Running head:** Extracellular matrix genes in LN

**Word Count:** Abstract 296, text 3497

**Corresponding author:**
Chagriya Kitiyakara, M.B.B.S. (Lond), F.R.C.P. (UK)
Renal Division, Department of Medicine
Ramathibodi Hospital, Mahidol University
270 Rama 6 Rd, Bangkok 10400, Thailand
Tel: 662+ 201-1400, Fax: 662+ 201-1715
Email: kitiyakc@yahoo.com
**ABSTRACT**

The deposition of extracellular matrix (ECM) leading to fibrosis is a major pathogenic mechanism for progressive decline in renal function in patients with lupus nephritis (LN). Currently available clinic-pathological features cannot predict renal outcome consistently. This study examined whether increases in the expression of fibrogenic genes can predict long term renal outcome in LN. Renal gene expression levels of transforming growth factor beta 1 (TGFβ1), and collagen I (COL1) were studied by real-time multiplex quantitative PCR in a prospective cohort of patients with LN (n=39). After a long follow-up (median 43.9 months), **Renal failure** (50% reduction in GFR or dialysis) developed in 13 subjects. The expression levels of renal fibrogenic genes were increased compared to controls without LN. ECM deposition by Sirius red staining correlated with TGF β1 and COL1 expression. By multivariate analysis which included traditional clinic-pathological factors, high renal COL1 expression (HR = 4.7, P = 0.030) and low baseline GFR (HR= 5.2, p= 0.023) were independent predictors of **renal failure**. TGFβ1 tended to show increased risk by univariate analysis, but this was not significant. In conclusion, high renal COL1 expression is a strong predictor of long term adverse renal outcome independent of baseline GFR and could be a useful marker that may allow clinicians to target patients at high risk of progression to more aggressive therapy, and minimize treatment related side-effects in low risk patients.

**Keywords:** cytokines; extracellular matrix; fibrosis; gene expression; kidney; transforming growth factor
Introduction

Patients with lupus nephritis (LN) have increased risk of progressive renal failure requiring renal replacement therapy. Treatment of LN includes corticosteroids and immunosuppressive agents, which can lead to significantly increased risks of morbidity and mortality (1). LN encompasses many forms of glomerulonephritis with inflammatory and fibrotic components (2). A renal biopsy is typically used to obtain prognostic information and to serve as a guide to therapy. The World Health Organization (WHO) and later RPS/ISN have proposed morphologic classifications for LN (2). While these classification systems are valuable, individual patients even within the same class may show large variability in long term outcomes. Semi-quantitative activity and chronicity indices have been used to augment prognostic information from the WHO classification system, but the values of these indices remain controversial (3, 4).

In LN, active inflammation may be followed by progressive decline in kidney function due to the development of fibrosis. Interstitial fibrosis has been associated with poor clinical outcome in many renal diseases including LN (5-8). The deposition of extracellular matrix (ECM) proteins such as collagen I (COL 1) is a key component of fibrosis. A number of cytokines including transforming growth factor-β1 (TGFβ1) may mediate fibrosis by stimulating COL1 synthesis as well as through other mechanisms (5, 6).

The presence of interstitial fibrosis often implies poor outcome, but for such processes to be demonstrable on routine histology, the process is already largely irreversible (9). Genes involved in matrix expansion and fibrosis has been shown to correlate with fibrosis in many experimental models including LN (5, 7). Inhibition of these fibrogenic pathways may be a potent therapeutic means for treating fibrosis in various kidney disease models (10). Currently, data for expression of fibrogenic genes on response to treatment and long-term outcome in LN in human subjects are limited. Renal fibrogenic genes may predict renal
outcome and serve as a guide to therapy in LN. Furthermore, it could pave a way for therapeutic interventions aimed at decreasing the fibrogenic proteins.

In this study, we proposed that the degree of fibrosis associated in LN is associated with increased expression of renal fibrogenic genes: TGFβ1 and COL1 and that increased fibrogenic genes expression are associated with adverse long term renal outcomes.

**MATERIAL AND METHODS**

**Patients’ baseline data and management**

Patients with a formal diagnosis of SLE undergoing a kidney biopsy at Ramathibodi Hospital for clinical indications between 2002 and 2004 were included in this study. In addition to the core of tissue sample obtained for diagnosis, an addition core of tissue was obtained, snapped frozen, and stored in liquid nitrogen. Control subjects with preserved renal function (serum creatinine < 1.2 mg/dl) were recruited from patients undergoing nephrectomy for renal cell carcinoma. A tissue core was obtained from the renal cortex of non-involved pole of the kidney, snapped frozen and stored in liquid nitrogen. This study was approved by the Ethical Committee of Ramathibodi Hospital, Faculty of Medicine. All participants gave written informed consent.

Routine clinical characteristics for each patient were recorded at baseline and at each follow-up. Glomerular filtration rate (GFR) was estimated using CKD-EPI formula (11). Patients were managed at the physicians’ discretion. In general, patients received antihypertensive therapy if their blood pressure exceeded 130/80 mmHg, and received immunosuppressive regimens according to the WHO category of glomerulonephritis.

The kidney biopsy tissues were fixed in Zenker, and evaluated for histology and immunofluorescence, and classified according to the RPS/ISN criteria (12) by a specialist nephropathologist, who did not have knowledge of the clinical outcome. In addition, the activity index (AI) and chronicity index (CI) were also assessed (12, 13). The patients were
also divided into categories: high or low AI (AI >7 or ≤7) and high or low CI (CI >or ≤ 3),
based on a previous study showing these values to be predictive of adverse outcome (14).

**Quantification of established fibrosis with picro-Sirius red stain**

The extent of fibrosis was evaluated in renal biopsy samples by computerized quantification of picro-sirius stained sections using a protocol modified from previous studies (15, 16). Six µm thick kidney tissues were prepared on glass slides, deparafinized, and hydrated in ethanol series. After the Zenker fixative was removed by 1% iodine in 80% ethanol and 5% sodium thiosulfate, slides were stained in 0.1% picro-sirius red overnight, placed in 0.01N hydrochloric acid, dehydrated through graded ethanol, placed in xylene, and cover slipped in permount.

The kidney biopsy was imaged (x40 magnification) in normal light, and double polarized light field with a digital camera (DS camera Control Unit DS-L2, Nikon). Ten images of the kidney cortex were randomly photographed for each subject. The microscope light source and the condenser were used at fixed settings in all images. Under normal light, extracellular matrix material is stained red, whereas under polarized light, fibrillar collagen is shown as bright areas.

Image analysis was analyzed by software (Image J1.30; National Institutes of Health, Bethesda, MD, USA). Color images were converted to gray scale RGB stack, then differentiated into three gray scale images representing red, green and blue. For analysis, the green-gray scale images were used in bright field picture, and the red-gray scale images were used for polarized light picture. The total area, intraluminal area, interstitial area (non-polarized), polarized positive areas were measured by threshold graph. Collagenous matrix index (CMI) and fibrillar collagen index (FCI) was calculated from the mean of the 10 photographs for each subject as follows:

\[
\text{The total cortical area} = \text{Total area} - \text{intraluminal area.}
\]
Collagenous matrix index (CMI) = (interstitial area/ total cortical area) x 100

Fibrillary collagen index (FCI) = (positive polarized light area/ total cortical area) x 100

**RNA extraction and cDNA synthesis**

Total RNA was isolated from kidney tissue by silica gel-based membrane spin technology with DNase I treatment (RNeasy micro kit: Qiagen, Chatsworth, CA). cDNA was synthesized by iScripts cDNA synthesis Kit using MMLV-derived reverse transcriptase pre-mixed with an RNase inhibitor (Biorad, USA) using a blend of oligo (dT) and random hexamer primers.

**Quantification of renal fibrogenic gene expression**

Messenger RNA expression was quantified by real-time quantitative polymerase chain reaction (RT-PCR) using Fast Start Universal Probe Master kit (Applied Biosystems, Foster City, CA). cDNA from kidney cortex was amplified using iScripts cDNA synthesis Kit (Biorad, U.S.A.) in a 96 well plate. Multiplex quantitative PCR was performed using target and housekeeping gene in the same well (17). By this method, the threshold cycle (Ct) of target genes in each sample was normalized with housekeeping to account for individual sample variation in RNA amount. Ct is defined as the cycle number at which the fluorescence intensity generated by the tracer dye, which is released from the probe during DNA amplification, reaches a fixed limit of detection. Ct was determined at the exponential phase, and is inversely related to the initial mRNA amount.

Primer sequences for target genes: TGFβ1, and COL1 are shown in Table 1. The mRNA expression of target genes was calculated by using the ΔCt procedure with VIC-TAMRA labeled GAPDH (Applied Biosystem, U.S.A.) used as housekeeping gene for normalization among samples where ΔCt values were calculated from ΔCt = Ct target gene – Ct
The primers and probes were tested without reverse transcriptase to ensure that they do not amplify genomic DNA. A standard curve was generated for each gene using pooled cDNA. Conditions of the PCR reaction were optimized so that the amplification efficiency of the target genes and the endogenous reference gene were comparable across 3 log dilutions of pooled cDNA (18). The corresponding ΔCt was plotted against the log concentrations of template cDNA. The slope of the linear regression of 0.1 indicated that the amplification efficiencies were comparable (data not shown).

The optimized conditions were: TGFβ1 or COL 1, 250 nM primers and 250 nM probes, with 1µl GAPDH. The PCR conditions were: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles at 95°C for 0.15 min, and 60°C for 1 min.

The mean Ct of triplicates was used. Mean CV % for the Ct were 0.01 – 0.02%. The relative levels of gene expression in LN patients were evaluated using the comparative cycle (Ct) method (18). The expression of target genes in LN tissues were expressed as fold change of the mean value of control subjects (n=3), which were arbitrary set as 1 as follows:

Relative expression to control tissues = $2^{-\Delta\Delta\text{Ct}}$

Where $\Delta\text{Ct} = \text{Ct target gene} - \text{Ct housekeeping gene}$; $\Delta\Delta\text{Ct} = \Delta\text{Ct LN tissues} - \Delta\text{Ct normal tissues}$

**Outcome definitions**

The two outcomes were i) *Remission*, defined as *complete remission* urine protein:creatinine <0.3g/g or *partial remission*, defined as urine protein:creatinine reduction by 50% or less than 1g/g without a reduction of eGFR by 50%, and ii) *Renal failure*, defined as a reduction of GFR by 50% ($\Delta50\%\text{GFR}$) or *ESRD* (eGFR<15 ml/min/1.73m$^2$ or dialysis for more than 3 months).

**6. Statistical analysis**
Data were summarized by mean ± SD, and by percentages. Gene expression was analyzed continuously, and also dichotomized into high (top 20%) and low to moderate (lower 80%). Correlations were calculated by Spearman’s rank coefficient. Student t test or nonparametric tests were used as appropriate to compare differences between groups for continuous data. Comparisons of categorical data were performed with the chi-square test. Survival methods were used to evaluate the prognostic value of the relative gene expression levels in the prediction of the outcomes in the presence of censoring by log-rank statistics and the results were presented by Kaplan–Meier plots (19). Predictive modeling was performed with Cox proportional hazards models (20). Multivariate models, a step-wise backward elimination procedure was used based on a likelihood ratio test with $P > 0.10$ for removal and $P < 0.05$ for entry of variables. Statistical significance was defined as a $P < 0.05$, two-tailed. All statistical analysis was conducted by software package SPSS version 15.0 (SPSS, Inc., Chicago, IL, USA).

RESULTS

Baseline clinicopathological characteristics

Thirty-nine patients had adequate kidney cortex tissues (>10 glomeruli) with optimal quality RNA (ratio of the absorbance at 260 and 280 nm ranging from 1.9 – 2.1). Nearly all were women (n=38). The age of the patients at the time of biopsy was 31.4 ± 10.6 years (Table 2). Time of disease onset to kidney biopsy was 4.6 ± 4.9 years. Most (87%) had received prior treatment with either prednisolone or immunosuppressive agents. Systolic BP was 130.3 ± 13.5 mmHg. Diastolic BP was 80.5 ± 10.2 mmHg. 53.9% had hypertension (BP>140/90), serum creatinine level was 1.7 ± 2.3 mg/dl. eGFR at the time of biopsy was 89.2 ± 39.2 ml/min. Eight patients (20.5%) had low GFR (GFR < 60 ml/min/1.73 m²), and 31 patients (79.5%) had high GFR (GFR ≥ 60 ml/min/1.73 m²). Proteinuria was 2.5 ± 4.1
g/24 h and 76.9 % had nephrotic range proteinuria. Urine rbc (Urbc) was 2.7 ± 2.3/hpf ,and 12.8 % had Urbc > 10/hpf.

One patient had mixed V + II classes, and the patient was designated as class II. There were 7 (17.9%) patients with class II, 3 (7.7%) with class III, 20 (51.3%) with class IV, and 9 (23.1%) with class V. AI was 4.2 ± 5.3. CI was 2.0 ± 1.7. 33 % had high AI (AI >7) and 33% had high CI (CI >3). Mean CMI was 7.4 ± 3.8. Mean FCI was 3.1 ± 1.1

Treatment and outcomes

The median follow up was 43.9 months   (IQR: 14.4 - 66.0 months). The majority of patients received intravenous cyclophosphamide according to NIH regimen (n=27). Other treatment protocols were: corticosteroid alone (n=5), corticosteroids + azathioprine (n=5), and mycophenolate mofetil + corticosteroids (n= 2). Twenty- seven (69.2%) patients were treated with ACEI. Two (5.1%) were treated with ARB. Ten (25.6%) did not receive ACEI, ARB, or other antihypertensive drugs.

There were 17 (43.6%) patients with remission, of whom 5 had complete remission, and 12 had partial remission. Relapse was observed in 12 cases (30.7%) Total of 13 had renal failure (9 had decreased 50% eGFR, 4 patients reached ESRD), and 16 (41.0%) patients died (Table 2).

Relative gene expression levels of COL1 and TGFβ1

The expression level of COL1 and TGFβ1 in LN were expressed as folds change of the controls. The mean folds increase for COL1 was higher compared to TGFβ1 (8.0 ± 1.6 vs 2.8 ± 1.8, p <0.001). The expression of COL 1 tended to correlate with TGFβ1 (R = +0.268, p = 0.090).

Gene expression, and baseline clinic-pathological characteristics
There were significant correlations between baseline clinico-pathological characteristics, and renal fibrosis gene expression, but these were not consistent (Table 3). COL 1 was positively correlated with proteinuria, hematuria, and CMI. The levels of TGFβ1 expression correlated positively with FCI.

The levels of gene expression were also compared between the following categories: high vs low eGFR, nephrotic range proteinuria (3.5g/d) vs sub-nephrotic, high hematuria (≥10 rbc/hpf) vs low (<10 rbc/hpf), high AI (≥7) vs low AI (<7), high CI (≥3) vs low CI (CI < 3). CMI and FCI were also divided into high or low at the median or top 20% vs lower 80%. There were no significant differences in COL1 or TGFβ1 gene expression between any of these categories (data not shown).

We also analyzed the gene expression levels according to the ISN/RPS class of the renal histopathology. Because of low sample numbers, we combined classes III and IV together. We found, the expression of TGFβ1 in class II was lower compared to class V (P < 0.05), but the expression of COL1 was not significantly different between classes (data not shown).

**Fibrosis gene and remission**

We analyzed the predictive capability of high gene expression (top 20%) on achieving remission compared to those with lower gene expression (lower 80%) using survival analysis. Neither High TGF nor Collagen predicted remission or complete remission compared to those with low gene expression values (p=NS).

**Fibrosis genes and renal failure**

We analyzed the predictive capability of high gene expression on development of renal failure. Patients with high gene expression (top 20%) were compared to those with lower gene expression (lower 80%) using survival analysis. Figure 1a showed that the cumulative risk of renal failure was significantly higher in those with high collagen (≥8 fold).
compared to those with low to moderate collagen (< 8 fold). A similar tendency is shown for high TGFβ1 (≥ 7 fold) compared to low to moderate TGF (< 7 fold). (figure 1b).

Next we evaluated the independent predictive value of high COL 1 gene in predicting renal failure using multivariate analysis along with traditional clinico-pathological predictors. GFR was less than 60 ml/min/1.73 m^2 was used to define moderate CKD according to Kidney Disease: Improving Global Outcomes initiative (KDIGO) (21). By univariate analysis, high COL1, low GFR, high CMI, high SBP tended to be predictive of renal failure (p<0.1), and were included in the multivariate analysis. By multivariate analysis, only COL1, and low baseline GFR remained independent predictors for renal failure with HR of 4.1, and 5.2 respectively (table 4). High TGFβ1 tended to increase risk of renal failure with HR of 1.7 by univariate analysis, but this was not significant. Neither genes were significantly predictive of renal failure when they were analyzed as continuous parameters.

**Discussion**

This study found that the fibrogenic genes were increased in LN with the gene expression for COL 1 showing higher relative increase than TGFβ1. The expression of these genes correlated with ECM deposition based on Sirius red staining, but the correlations were not consistent with other clinico-pathological parameters. The most interesting finding in this study is the fact that very high renal COL 1 gene expression was an independent predictor of long term adverse renal outcome.

A number of investigators have evaluated the predictive values of different clinico-pathologic variables as markers of outcome in LN. The WHO categories are useful for guiding immunosuppression when applied to a large group as a whole, but, because of the large variations within these groups, do not consistently identify which patients will develop
progressive renal disease (22). The Activity and Chronicity Indices, derived from composite scores of various histological parameters (13), were found to be predictive in some studies (23, 24) but not by others (9, 25). The role of tubulointerstitial damage in predicting outcome has been noted previously for many types of kidney diseases including LN (5, 7). Esdaile et al (26) showed that an index composed tubular atrophy, tubulointerstitial fibrosis was the best predictor of outcome. On the other hand, Hill et al (9) did not find that interstitial changes in the initial biopsy was predictive of outcome demonstrating that routine histopathological studies cannot consistently identify those likely to have adverse outcome (27).

In kidney diseases, the amount of ECM deposition depend on the balance between factors promoting accumulation and factors affecting ECM degradation (5, 7). The findings of high expression of TGFβ1 and COL1 in this study support the role of increased matrix synthesis in LN. Collagen I is normally present in relatively small amounts in the adult kidney (28). Accumulation of these molecules has been reported in a variety of chronic human kidney diseases (29) and CAN (30). Collagen I mRNA were overexpressed in lupus MRL/lpr mice compared to controls (27). In our study, collagen I mRNA expression correlated with the degree of proteinuria, and Sirius red staining, which is known to stain many types of collagens as well as other matrix components. A recent PCR array analysis of archival kidney biopsies of LN patients has identified COL1 as the gene, which has one of the highest correlation with chronicity scores in a cross-sectional analysis (31). We extended these findings and found that very high increase in COL1 mRNA was predictive of poor renal function. The survival analysis pointed to a predictive value of COL1 mRNA only when patient groups were dichotomized to compare those with markedly increased expression with groups with milder increases. This might suggest that there may not a clear linear relationship between COL1 gene expression at baseline, and poor outcome, rather that only in those individuals with large increases in collagen synthesis would the outcome be poor.
Numerous experimental studies support the role of TGF β1 as an important mediator of collagen synthesis and fibrosis in response to renal injury (5-7). Upregulated expression of TGF β1 is a feature of virtually all human and experimental models of renal fibrosis including murine lupus models (27). In a previous study, increased TGF β1 immunostaining was associated with adverse outcome in human LN (32). In our study, LN patients with high TGF gene expression tended to have poorer long term survival compared to those with lower expression, although this was not significant. Larger numbers of subjects would probably have demonstrated significantly adverse renal outcomes among those with high TGF β1 gene expression.

The interactions between collagen, TGF β1, and fibrosis are complex (5, 6). In this study, the expressions of the TGF β1 and COL1 genes were broadly correlated, but the coefficients of correlations were not high. Factors besides TGF β1 can affect collagen synthesis (5, 7). Furthermore, the biological action of TGF β1 is dependent on a number of other proteins, which regulate its ability to stimulate ECM synthesis. Other factors such as prior therapy could affect the transcription levels of genes to different extents for each individual. Such factors could explain the inconsistent correlations between TGF β1 and COL1, and between genes and clinical parameters.

The major limitations to this study include the rather small sample size. Nonetheless, to our knowledge this is the largest study to evaluate the long term effects of collagen I in LN. This is a mixed group of subjects with LN. A large proportion of patients were already on prior therapy. Most patients received cyclophosphamide according to the NIH regimen, but medication dose adjustments were individualized. Given these constraints, however, it is perhaps more striking that high COL1 mRNA expression is the strongest independent predictor of long term outcome. COL1 synthesis therefore may be a critical determinant of renal fibrosis in LN in real world clinical settings. Therefore, it is possible that in an
unselected group of subjects with LN, very high renal collagen I gene expression could be capable of identifying patients who have such a strong shift towards renal fibrosis even among those who received prior immunosuppressive therapy.

**Conclusions**

Treatment of patients with LN can lead to life-threatening complications. Current clinico-pathological parameters do not identify individuals at risk of progressive renal failure accurately. In this study, we found that high collagen expression is an independent predictor of long term renal outcome such that patients with COL1 gene expression 8 folds above normal have 5 folds increased risk of progressive renal failure. If these findings can be confirmed in other studies, COL1 gene expression could provide an extremely useful marker of poor disease outcome to target such patients to more aggressive immunosuppression or specific anti-fibrotic therapy. This may decrease rates of renal failure in those with highest risks, and minimize treatment related side-effects in low risk patients.
ACKNOWLEDGMENTS

CT was supported by grants from the Strategic Scholarships for Frontier Research Network for the Join Ph.D. Program Thai Doctoral degree, and the Office of the Higher Education Commission. CK was supported by grants from the Office of the Higher Education Commission, and Thailand Research Fund (RMU4880048), Mahidol University, and Ramathibodi Hospital. SK was supported by the Research Fund of “Genetic studies in transgenic animal models of SLE” (no. 02011854-0005) Mahidol University. This paper was taken from a dissertation submitted by Chiraporn Tachaudomdach to fulfill requirements for the Ph.D. degree from Mahidol University in 2011.
REFERENCES

<table>
<thead>
<tr>
<th>Genes</th>
<th>Accession No</th>
<th>Primer (forward/reverse)</th>
<th>Product length (bp)</th>
</tr>
</thead>
</table>
| TGFβ1 | NM_000660.4  | F: 5’–CCAGCATTCTGCAAAGCTC  
|       |              | R: 5’–GTCAATGTACAGCTGCCGCA  
|       |              | ACACCAACTATTGCTTCAGCTCCACGGA (Tamra) | 100 |
| COL1  | NM_000088.3  | F: 5’–CCTCAA GGCTCCAACGAG  
|       |              | R: 5’–TCAATCAGTCTTTGCCCCA  
|       |              | ATGGCTGCACGAGTCACACCGGA (Tamra) | 117 |
Table 2. Baseline clinical data (n=39)

<table>
<thead>
<tr>
<th>Clinical Characteristics</th>
<th>Mean ± SD or percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at initial biopsy (years)</td>
<td>31.4 ± 10.6</td>
</tr>
<tr>
<td>Time from disease onset to Biopsy (years)</td>
<td>4.6 ± 4.9</td>
</tr>
<tr>
<td>Women (%)</td>
<td>97.4 %</td>
</tr>
<tr>
<td>Prior immunosuppressive agent (%)</td>
<td>87.2 %</td>
</tr>
<tr>
<td>Systolic BP (mmHg)</td>
<td>130.3 ± 13.5</td>
</tr>
<tr>
<td>Diastolic BP (mmHg)</td>
<td>80.5 ± 10.2</td>
</tr>
<tr>
<td>Serum albumin (g/l)</td>
<td>29.7 ±7.4</td>
</tr>
<tr>
<td>Creatinine (mg/dl)</td>
<td>1.7 ± 2.3</td>
</tr>
<tr>
<td>eGFR (ml/min/1.73 m^2)</td>
<td>89.2 ± 39.2</td>
</tr>
<tr>
<td>Urine protein (g/day)</td>
<td>2.5 ± 4.1</td>
</tr>
<tr>
<td>Hypertensive (BP&gt;140/90) (%)</td>
<td>53.9 %</td>
</tr>
<tr>
<td>Low GFR (eGFR&lt; 60 ml/min/1.73 m^2) (%)</td>
<td>20.5 %</td>
</tr>
<tr>
<td>Nephrotic range proteinuria, &gt;3g (%)</td>
<td>76.9 %</td>
</tr>
<tr>
<td>Urine rbc (&gt; 10/hpf) (%)</td>
<td>12.8 %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kidney Biopsy</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPS/ISN Class: II/III/IV/V</td>
<td>7/ 3/20/9/2 (17.9/ 7.7/ 51.3/23.1)</td>
</tr>
<tr>
<td>Activity index</td>
<td>4.2 ± 5.3</td>
</tr>
<tr>
<td>Chronicity index</td>
<td>2.0 ± 1.7</td>
</tr>
<tr>
<td>Collagen Matrix Index (CMI)</td>
<td>7.4 ± 3.8</td>
</tr>
<tr>
<td>Fibrillary Collagen Index (FCI)</td>
<td>3.1 ± 1.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All remission: complete/partial</td>
<td>17: 5/12 (43.6: 12.8/30.7%)</td>
</tr>
<tr>
<td>Decreased 50% eGFR</td>
<td>9 (23.1 %)</td>
</tr>
<tr>
<td>Dialysis</td>
<td>4 (10.3 %)</td>
</tr>
</tbody>
</table>

BP blood pressure; GFR glomerular filtration rate; RPS/ISN Renal Pathological society/International society of Nephrology.
Table 3. Relationships between genes expression and clinical parameters at baseline

<table>
<thead>
<tr>
<th>Genes</th>
<th>AI</th>
<th>CI</th>
<th>CMI</th>
<th>FCI</th>
<th>GFR</th>
<th>Proteinuria</th>
<th>Hematuria</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL 1</td>
<td>R=-0.12</td>
<td>R=-0.20</td>
<td><strong>R=+0.37</strong></td>
<td>R=-0.07</td>
<td>R=-0.16</td>
<td><strong>R=+0.45</strong></td>
<td><strong>R=+0.45</strong></td>
</tr>
<tr>
<td></td>
<td>P=0.46</td>
<td>P=0.21</td>
<td><strong>P=0.04</strong></td>
<td>P=0.64</td>
<td>P=0.38</td>
<td><strong>P=0.011</strong></td>
<td><strong>P=0.011</strong></td>
</tr>
<tr>
<td>TGFβ1</td>
<td>R=-0.17</td>
<td>R=-0.04</td>
<td>R=0.23</td>
<td><strong>R=0.32</strong></td>
<td>R=-0.11</td>
<td>R=+0.16</td>
<td>R=0.01</td>
</tr>
<tr>
<td></td>
<td>P=0.28</td>
<td>P=0.82</td>
<td>P=0.16</td>
<td><strong>P=0.04</strong></td>
<td>P=0.49</td>
<td>P=0.29</td>
<td>P=0.92</td>
</tr>
</tbody>
</table>

AI: Activity Index, CI: Chronicity Index, CMI: Collagen Matrix Index, FCI: Fibrillary Collagen Index, GFR: Glomerular filtration rate, COL1: Collagen Type I, TGFβ1: Transforming growth factor beta 1. **Bold** p<0.05
<table>
<thead>
<tr>
<th>Predictor</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard</td>
<td>Hazard</td>
</tr>
<tr>
<td></td>
<td>Ratios</td>
<td>Ratios</td>
</tr>
<tr>
<td>COL 1</td>
<td>4.69</td>
<td>4.11</td>
</tr>
<tr>
<td>(fold ≥8 vs &lt;8)</td>
<td>1.74</td>
<td>NS</td>
</tr>
<tr>
<td>TGFβ1 (fold ≥7 vs &lt;7)</td>
<td>1.22</td>
<td>NS</td>
</tr>
<tr>
<td>RPS/ISN Class (II+V vs III+IV)</td>
<td>1.42</td>
<td>NS</td>
</tr>
<tr>
<td>Activity Index (AI ≥7 vs &lt;7)</td>
<td>1.43</td>
<td>NS</td>
</tr>
<tr>
<td>Chronicity Index (CI ≥3 vs &lt;3)</td>
<td>3.19</td>
<td>1.25</td>
</tr>
<tr>
<td>Collagen Matrix Index (CMI ≥0.08 vs &lt;0.08)</td>
<td>1.65</td>
<td>NS</td>
</tr>
<tr>
<td>Fibrillary Collagen Index (FCI ≥0.04 vs &lt;0.04)</td>
<td>0.91</td>
<td>NS</td>
</tr>
<tr>
<td>Age at nephritis onset (Age ≥ 25 vs &lt; 25)</td>
<td>2.01</td>
<td>NS</td>
</tr>
<tr>
<td>Time from onset to Biopsy (Time ≥ 1 yr vs &lt; 1 yr)</td>
<td>2.94</td>
<td>1.94</td>
</tr>
<tr>
<td>Systolic BP (SBP ≥ 140 vs &lt; 140)</td>
<td>0.00</td>
<td>NS</td>
</tr>
<tr>
<td>Diastolic BP (DBP ≥ 90 vs &lt; 90)</td>
<td>5.35</td>
<td>5.15</td>
</tr>
<tr>
<td>Glomerular filtration (eGFR&lt; 60 vs ≥60 ml/min/1.73 m²)</td>
<td>0.03</td>
<td>NS</td>
</tr>
<tr>
<td>Proteinuria (Nephrotic vs sub-nephrotic)</td>
<td>1.07</td>
<td>NS</td>
</tr>
<tr>
<td>Hematuria (≥ 10rbc vs &lt;10 rbc/hpf)</td>
<td>0.28</td>
<td>NS</td>
</tr>
<tr>
<td>Immunosupressive vs steroid alone</td>
<td>0.20</td>
<td>NS</td>
</tr>
<tr>
<td>ACEI or ARB (Yes vs No)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ACEI angiotensin converting enzyme inhibitors, ARB angiotensin receptor blockers, COL1: Collagen Type I, TGFβ1: Transforming growth factor beta 1. Parameters with p<0.1 by univariate analysis were included in the multivariate analysis. NS P>0.1 by univariate. Bold shows significant parameter by multivariate analysis.
**Legend for figures**

Figure 1. Cumulative risks of renal failure in a) Patients with high (≥ 8 folds) COL1 gene expression versus low to moderate (< 8 folds) COL1 gene expression;  P = 0.030 and b) Patients with high (≥ 7 folds) TGFβ1 gene expression versus patients with low to moderate (<7) TGFβ1 gene expression;  P = 0.187
P=0.187

- High TGF beta 1
  (>=7 folds)

- Low-Moderate
  TGF beta 1
  (<7 folds)