OPEN

Mesangial Expansion by Morphometry at 5 y After Kidney Transplantation: Incidence, Risk Factors, and Association With Graft Loss

Aleksandar Denic, MD, PhD,[1] Alessia Buglioni, MD,[2] Sandor Turkevi-Nagy, MD,[3] Mateo Velasquez Mejia, MD,[4] Byron H. Smith, PhD,[5] Walter D. Park, BS,[4] Rashmi Subramani, MBBS,[1] Aleksandra Kukla, MD,[1] Tayyab S. Diwan, MD,[4] Joseph P. Grande, MD,[2] and Mark D. Stegall, MD,[4]

Background. Mesangial expansion (ME) is an understudied histologic lesion in renal allografts. The current Banff *mm* score is not reproducible and may miss important ME features. The study aimed to improve the quantification of ME using morphometry, assess changes over time, and determine its association with allograft loss. **Methods.** We studied ME in 1-y and 5-y surveillance biopsies in 835 kidney transplants performed between January 2000 and December 2013. ME was assessed using the Banff *mm* score by a central pathologist and by morphometry. We derived 3 different morphometric measures: (1) %ME*mm* (%glomeruli with ME in ≥2 lobules, like Banff *mm*); (2) $\%ME_{any}$ (%glomeruli with any ME lesion); and (3) %ME area (sum of all ME areas/all glomerular tuft areas). Unadjusted and adjusted Cox models assessed the risk of death-censored allograft loss. **Results.** From 1- to 5-y biopsies, the mean Banff *mm* score increased from 0.18 to 0.34, whereas %ME*mm* increased from 2.5% to 13.3%. Banff *mm* score had modest correlations with morphometric ME measures. Moderate-severe %ME*mm* was present in 20.1% of 5-y biopsies, whereas only 6.6% of Banff *mm* scores were. In general, higher ME on both 1- and 5-y biopsies was associated with a deceased donor, older recipient age, recipient diabetes/obesity (present in >50% of severely affected biopsies), higher hemoglobin A1c at 5 y posttransplant, and recurrent kidney disease. Higher ME on 5-y biopsies was associated with delayed graft function. A higher Banff *mm* score at 1-y biopsy and morphometry ME measures at 5-y biopsy were associated with rejection during the first year posttransplant. Morphometric ME measures were associated with allograft loss independent of Banff scores and all clinical characteristics, including kidney function and recurrent disease. The model with $\%ME_{any}$ had the highest c-statistic (0.872). **Conclusions.** Banff *mm* score underestimates the pervasiveness of ME in 5-y biopsies. ME is common and associated with alloimmune and nonalloimmune causes of graft loss.

(*Transplantation Direct* 2024;10: e1652; doi: 10.1097/TXD.0000000000001652.)

Received 20 February 2024. Revision received .
Accepted 6 March 2024.
[1] *Division of Nephrology and Hypertension, Mayo Clinic, Rochester, MN.*
[2] *Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN.*
[3] *Department of Pathology, Albert Szent-Györgyi Health Centre, University of Szeged, Szeged, Hungary.*
[4] *Departments of Surgery and Immunology, Mayo Clinic, Rochester, MN.*
[5] *Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN.*
A.D. was supported by the Robert W. Fulk Career Development Award Fund in Nephrology Research Honoring Dr. Fernando Fervenza.
The authors declare no conflicts of interest.
A.D., W.D.P., and M.D.S. designed the study. A.D., J.P.G., A.B., S.T.-N., B.H.S., W.D.P., M.V.M., and R.S. collected or provided the data. A.D. and B.H.S. analyzed the data. A.D. wrote the original draft. All authors contributed to revisions and approved the final version of the article.
Supplemental digital content (SDC) is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.transplantationdirect.com).
Correspondence: Mark D. Stegall, Departments of Surgery and Immunology, Mayo Clinic, 200 1st St SW, Rochester, MN 55905. (stegall.mark@mayo.edu).

ISSN: 2373-8731
DOI: 10.1097/TXD.0000000000001652

Improving the long-term survival of renal allografts requires a better understanding of the cause of late allograft loss. Although alloimmunity and recurrent kidney disease are clearly recognized causes of late allograft loss,[1,2] recent studies suggest that by 5 y after kidney transplantation, many grafts develop a progressive increase in chronic lesions, such as glomerular sclerosis and mesangial expansion (ME), that might represent nonalloimmune mechanisms of late graft loss. For example, in surveillance biopsies, moderate-to-severe ME scored by Banff or Tervaert classification was common in patients with both diabetes and obesity and correlated with death-censored allograft loss.[3] Similarly, a Japanese study of 89 patients with 7-y surveillance biopsies showed that the presence of any ME by Banff scoring predicted glomerular filtration rate (GFR) decline.[4]

ME is characterized by the aberrant proliferation of mesangial cells and excess production of matrix proteins that has been linked to diabetic nephropathy[5,6] and kidney functional decline.[7] The Banff *mm* score is defined as the percentage of nonsclerosed glomeruli with at least moderate mesangial matrix expansion.[8] However, because ME can be focal or pervasive, several ways exist to assess its severity.

Recently, we measured ME using morphometry and demonstrated that both Banff *mm* score and morphometric percentage of ME (%ME) area predicted death-censored

allograft loss (DCGL) beyond 5 y.[9] Additionally, Banff *mm* score suffers from poor interobserver reproducibility with kappa values ranging from 0.19 to 0.37 when pathologists used glass slides.[10] Conversely, we showed that intraclass correlation for %ME area between 2 trained morphometry readers was 79.2%.[9]

The goals of this study were to (1) use computer-assisted morphometry to measure ME in several ways, (2) assess the relative prevalence of ME at 1 and 5 y after kidney transplantation using surveillance biopsies,[11] (3) identify risk factors for ME, and (4) compare the correlation between the various approaches to measuring ME including Banff *mm* score with DCGL.

MATERIALS AND METHODS

Study Population

This study was approved by the Mayo Clinic Institutional Review Board (approval number 17-002391). The clinical and research activities being reported are consistent with the Principles of the Declaration of Istanbul as outlined in the "Declaration of Istanbul on Organ Trafficking and Transplant Tourism." This study included 835 patients from a previously described cohort who had received a solitary conventional ABO and HLA compatible (negative cytotoxic crossmatch) kidney transplant at Mayo Clinic, Rochester, MN, between 2000 and 2013; and had an adequate 5-y surveillance kidney biopsy (at least 4 glomeruli).[12]

Donor and Recipient Clinical Characteristics

For the living donors, we obtained age and body mass index (BMI) from the medical evaluation before kidney donation and for the deceased donors at the time of organ procurement. For the recipients, baseline clinical characteristics were obtained from the medical records at the time of transplantation. These characteristics included age, race, BMI, presence of pretransplant diabetes, presence of pretransplant hypertension, delayed graft function (DGF), and the number of HLA class 1 mismatches. At 1 y posttransplant, we collected biopsy data (n = 769) and determined whether any episodes of acute rejection had occurred since the transplant. Clinical characteristics collected at 5 y were estimated GFR, proteinuria, and hemoglobin A1c (HbA1c). Because of missing data, a window of HbA1c values has been expanded to 90 d before or after the 5-y biopsy. From the 1- and 5-y pathology reports, we collected available Banff scores (*mm*, interstitial fibrosis [*ci*], tubular atrophy [*ct*], transplant glomerulopathy [*cg*], glomerulitis [*g*], arteriolar hyalinosis [*ah*], arterial fibrointimal thickening [*cv*], peritubular capillaritis [*ptc*], inflammation [*i*], tubulitis [*t*], intimal arteritis [*v*], and total inflammation [*ti*]) that were recorded according to Banff classification at the time of biopsy.

ME by Pathologist

Per Banff, ME was scored as follows: *mm0*—No more than mild mesangial matrix increase in any glomerulus; *mm1*—at least moderate mesangial matrix increase in up to 25% of nonsclerotic glomeruli; *mm2*—at least moderate mesangial matrix increase in 26% to 50% of nonsclerotic glomeruli; and *mm3*—at least moderate mesangial matrix increase in >50% of nonsclerotic glomeruli.[8] Because the Banff *mm* score was missing in many 1 and 5-y biopsies, a central pathologist (J.P.G.) reread Banff *mm* scores in all biopsies, masked to clinical characteristics and outcomes.

Kidney Biopsy Morphometry

Periodic acid Schiff–stained slides were scanned into high-resolution digital images at 20× (0.504 µm per pixels) magnification (Aperio AT2 system scanner; Leica Microsystems Inc, Buffalo Grove, IL; https://www.leicabiosystems.com/us/digital-pathology/). Using ImageScope software (version 12.4.3.7009 Aperio), an experienced morphometrist (A.D.) magnified digital images onto a large computer tablet and manually traced all glomeruli (Figure 1A). Then, within each glomerulus, if present, each discrete focus of ME (Figure 1B–D) was traced.[9] Severely ischemic, segmentally sclerosed, or globally sclerosed glomeruli were not used to grade ME (Figure 1E). The ImageScope has a built-in tool that records the area and count measurements of all annotations. We required that an area of ME had to be at least 200 µm^2 to be scored (slightly larger than the Banff requirement that ME should be the size of at least 2 mesangial cells). We generated 3 morphometric ME measures: (1) %ME$_{mm}$—this metric was designed to parallel the Banff *mm* score by requiring at least 2 glomerular lobules to be affected by ME, and it represents the percentage of nonsclerosed glomeruli that met this criterion; (2) %ME$_{any}$—this metric represents the percentage of nonsclerosed glomeruli with any ME foci (even if a single ME focus was present) and was designed to assess the pervasiveness of ME; (3) %ME area—the sum of all of the total area of ME divided by the sum of the total area of all nonsclerosed glomeruli in the biopsy.[9] This metric was designed to assess the severity of ME in the biopsy.

Statistical Analyses

Weighted Cohen's kappa test was used to assess the interobserver reproducibility between the original Banff *mm* scores from the pathology reports and *mm* rescores (central pathologist) in 552 5-y biopsies. A subset of 45 5-y biopsies with various degrees of ME was used to assess the reproducibility of 3 morphometric ME measures between 2 investigators masked to each other's measures (A.D. and M.V.M.). The reproducibility was assessed using the pairwise, 2-way intraclass correlation coefficient (ie, the proportion of variation that occurs not because of measurement error). The Wilcoxon rank-sum test was used to compare the Banff *mm* score and morphometric ME measures by the allograft status and the presence/absence of pretransplant diabetes. Spearman's correlations compared the Banff *mm* score and morphometric ME measures, and Banff *mm* score and 3 morphometric ME measures with donor and recipient clinical characteristics at both 1 and 5 y. Because of the high correlation between pretransplant diabetes and HbA1c, we assessed their correlations with the Banff *mm* score and morphometric ME measures after adjusting for each other using partial Spearman correlations.[13] Cox proportional hazards models assessed the risk of death-censored allograft loss for Banff *mm* score and 3 morphometric ME measures at 5-y biopsy. In all Cox models, we used a time to an event that started from the date of the 5-y biopsy and stopped at 5-y following 5-y biopsy. Censoring was performed at the last follow-up (obtained from the medical record) or patient's death. We used unadjusted and several multivariable-adjusted models in all patients. One multivariable model adjusted for Banff scores that were significant

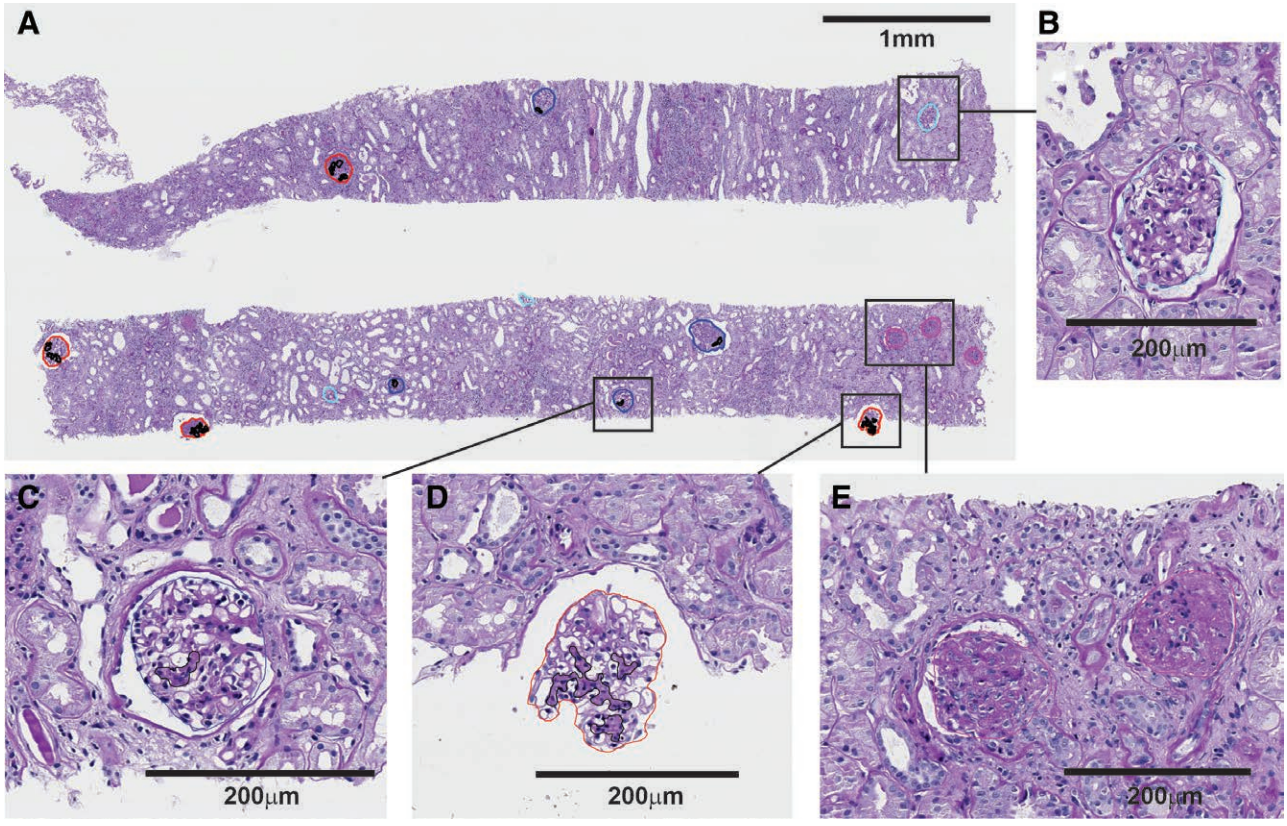

FIGURE 1. Representative images of a 5-y surveillance biopsy that was scored to have Banff *mm*=2, and how different measures of ME can be calculated. Morphometrist traced every ME lesion with an area of at least 200 μm^2 (all black areas). A, An example of a PAS-stained section with 2 cores used in morphometric analysis of mesangial expansion. Magnifications show (B) a glomerulus (traced in cyan) that has no more than a mild ME, which was not traced, (C) a glomerulus (traced in blue) that has a single focus with ME and at least 3 mesangial cells in that area, (D) a glomerulus (traced in red) that has 3 foci with ME, and (E) a severely ischemic and a globally sclerosed glomeruli that were not used in this analysis. Percent morphometric ME measures can be obtained by dividing the number of affected glomeruli by morphometry with the total number of all assessed glomeruli: (1) count only glomeruli with at least 2 affected lobules (in red traces), mimicking Banff *mm* score; or (2) count all affected glomeruli (in red and blue traces). From all ME areas and all 11 glomerular tuft areas, we can calculate the %ME area. ME, mesangial expansion; PAS, periodic acid Schiff.

in unadjusted analyses. In 3 other models, we added 1 morphometric ME measure at a time. These analyses were then repeated in a subset of 412 patients with available C4d scores. Five additional sensitivity unadjusted Cox models were performed: (1) the first model was performed in a subset of 698 biopsies with at least 7 glomeruli on the biopsy, (2) the second model was performed in a subset of 137 biopsies with <7 glomeruli, (3) the third model was performed in a subset of 791 biopsies with no chronic glomerulopathy (*cg* score = 0), (4) the fourth model was performed in a subset of 753 biopsies with no glomerulitis (*g* score = 0), and (5) the fifth model was performed in a subset of 779 biopsies with no diagnosis of recurrent disease. For each of the multivariable-adjusted Cox models, we calculated the c-statistic and a mean 10-fold cross-validated c-statistic. We also performed a series of multivariable Cox models where the Banff *mm* score and each of the 3 morphometric ME measures were adjusted for baseline donor and recipient characteristics, and then for recipient kidney function and HbA1c at 5 y. We have tested the assumption of proportional hazards for Cox models using the Schoenfeld residuals. We have also performed Fine-Gray modeling of subdistribution hazards to account for the competing risks of death with a functioning graft and graft failure. For Kaplan-Meier survival plots, the risk of death-censored allograft loss was assessed by the 4 grades of a Banff *mm* score and 4 grades of morphometric ME measures. Both $\%ME_{mm}$ and $\%ME_{any}$ were converted into 4 grades using thresholds for Banff *mm* score (0: 0, 1: ≤25%, 2: 26%–50%, 3: >50% of affected glomeruli). Finally, the %ME area was converted into a 0–3 score using a priori chosen thresholds (0: 0%, 1: ≤5%, 6%–10%, 3: >10%). All statistical analyses were performed using BlueSky Statistics software version 7.40 (BlueSky Statistics LLC, Chicago, IL) and R (RStudio) version 4.1.2.

RESULTS

Study Population

Of the 1465 solitary kidney transplants performed in the study period (2000–2013), 835 met the criteria for analysis. We excluded 497 patients who did not return for a follow-up visit or who had graft failure/death before 5 y, and 133 patients with missing or inadequate 5-y surveillance biopsy.[12] Among 835 patients, a subset of 769 also had a prior 1-y surveillance biopsy. The donors were on average 43.6 y old and had a mean BMI of 27.7 kg/m^2 (Table 1). At the time of transplantation, the recipients were on average 52.7 y old, 92% were White, 82% received a living donor transplant, 27.5% had pretransplant diabetes, and 97% had hypertension. Posttransplant, 4.6% had a DGF, and 10.8% had a rejection episode (acute T cell–mediated, including borderline,

TABLE 1.

Clinical and biopsy characteristics of the study population

Donors' characteristics	Value
Donor's age, y	43.6 (12.8)
Donor's BMI, kg/m^2	27.7 (4.9)
Living donor, n (%)	684 (81.9)
Recipients' characteristics	
Recipient's age, y	52.7 (13.0)
White race, n (%)	766 (91.7)
BMI at transplant, kg/m^2	28.5 (6.0)
Pretransplant diabetes, n (%)	230 (27.5)
Pretransplant hypertension, n (%)	810 (97.0)
No. of HLA mismatches[a]	3.1 (1.8)
Delayed graft function, n (%)	38 (4.6)
Posttransplant characteristics, n (%)	
Any rejection in the first year	90 (10.8)
Borderline	26 (3.1)
Acute T cell–mediated rejection	58 (6.9)
Antibody-mediated rejection	2 (0.2)
Mixed or other rejection	4 (0.5)
HbA1c,[b]	6.4 (1.5)
Characteristics at 5 y	
eGFR at 5 y,[c] mL/min/1.73 m^2	52.9 (17.3)
24-h urine protein at 5 y,[d] mg	108 (63–211)
Recurrent disease at 5-y biopsy	74 (8.9%)
Outcomes	
Overall follow-up time, y	12.5 (10.0–15.4)
Graft failure, n (%)	58 (6.9%)
Mesangial expansion measures at 1-y biopsy	
Banff *mm* score	0.18 (0.45)
%ME$_{mm}$	2.5 (7.3)
%ME$_{any}$	7.4 (12.2)
%ME area	0.3 (0.7)
Mesangial expansion measures at 5-y biopsy	
Banff *mm* score	0.34 (0.63)
%ME$_{mm}$	13.3 (20.7)
%ME$_{any}$	20.4 (26.2)
%ME area	2.1 (3.6)

Data shown as mean (SD), n (%), or median (IQR) for skewed data.
[a]Data missing in 1 patient.
[b]Data missing in 197 patients.
[c]Data missing in 2 patients.
[d]Data missing in 62 patients.
BMI, body mass index; eGFR, estimated glomerular filtration rate; IQR, interquartile range; %ME, percentage of mesangial expansion.

antibody-mediated, mixed rejection) during the first year posttransplant. The mean estimated GFR at 5 y was 52.9 mL/min/1.73 m^2, and the median 24-h urine protein was 108 mg. Overall median follow-up time was 12.5 y; however, to study the early effects of ME on the outcome, we truncated follow-up time at 5 y after the 5-y biopsy. Within this time frame, there were 58 allograft losses after a median of 2.4 y (interquartile range, 1.4–3.9).

Reproducibility

The Banff *mm* score on 5-y biopsy from the pathology reports was available in 552 (66.1%) cases. The kappa statistic between these original Banff *mm* scores and *mm* re-reads was 0.32. In a subset of 45 cases with various degrees of ME, the intraclass correlations between two morphometrists were 85.4% for %ME$_{mm}$, 84.5% for %ME$_{any}$, and 83.5% for %ME area.

ME by Pathology and Morphometry

ME by both pathology and morphometry was minimal on 1-y biopsy and increased from 1- to 5-y biopsies (Table 1). By Banff scoring (central pathologist), most 5-y biopsies had mild or no *mm*: 614 (73.5%) *mm* = 0, and 166 (19.9%) *mm* = 1. Only 55 (6.6%) biopsies had moderate-to-severe Banff *mm* score.

The Banff *mm* score and 3 morphometry ME measures on 5-y biopsy had a modest correlation with each other (r_s = 0.45–0.47, $P < 0.0001$ for all; Figure 2A–C). Another way to compare Banff *mm* scores to morphometric %ME$_{mm}$ is to convert the continuous %ME$_{mm}$ measure into the Banff categories. This showed that 168 (20.1%) of biopsies had moderate-to-severe scores in contrast to the 55 (6.6%) scored by the pathologist.

The central pathologist had a general agreement with morphometry when morphometric measures were stratified into 4 grades (kappa statistic was 0.30). For example, 84% of the time, they both called a case none-to-mild ME, and 70% of the time, they both scored a case as moderate-to-severe ME (**Table S1, SDC,** http://links.lww.com/TXD/A661). Many of the discrepancies were those close to the categorical cutoffs and those with mild ME by morphometry. Mild ME was common with 464 (55.6%) of biopsies having at least 1 focus of ME (%ME$_{any}$ >0). However, severe lesions were less common. Only 38 (4.6%) of biopsies had >10% of the glomerular area affected by ME. Morphometry allowed us to measure both the severity of ME (increasing total area of %ME area) and the pervasiveness (%ME$_{any}$). Importantly, the correlation between these 3 continuous measurements was very high (r_s = 0.94–0.98, $P < 0.0001$ for both; Figure 2D and E). However, the shape of the scatterplot suggests that ME was not always evenly distributed among glomeruli. At higher levels of %ME$_{any}$, %ME area varied a lot, suggesting different severity of ME at the same level of pervasiveness. A similar high correlation was found between %ME area and %ME$_{mm}$.

Baseline and Posttransplant Clinical Characteristics and ME

Overall, the Banff *mm* score and morphometric ME measures on both 1- and 5-y biopsies similarly associated with a deceased donor and higher recipient BMI (Table 2). Compared with the Banff *mm* score, morphometric ME measures on both 1- and 5-y biopsies had stronger associations with older recipient age, pretransplant diabetes, and HbA1c. All ME measures on 1- and 5-y biopsies by Banff and morphometry associated with recurrent kidney disease. All ME measures on 5-y biopsy by Banff and morphometry were associated with DGF. Only the Banff *mm* score on 1-y biopsy and %ME$_{any}$ and %ME on 5-y biopsy were associated with rejection episodes in the first year.

We found a high correlation between pretransplant diabetes and HbA1c (r_s = 0.51, $P < 0.0001$), and therefore, we assessed their correlations with ME measures after adjusting for each other. After adjusting for HbA1c, pretransplant diabetes was no longer associated with Banff *mm* score (r_s = 0.07, P = 0.07) but still strongly associated with all 3 morphometric ME measures (r_s = 0.27, $P < 0.0001$ for all). Conversely, after adjusting for pretransplant diabetes, HbA1c is no longer associated with Banff *mm* score or 3 morphometric ME measures ($P > 0.05$ for all). Finally, we compared the 28 patients with diabetes at transplant who

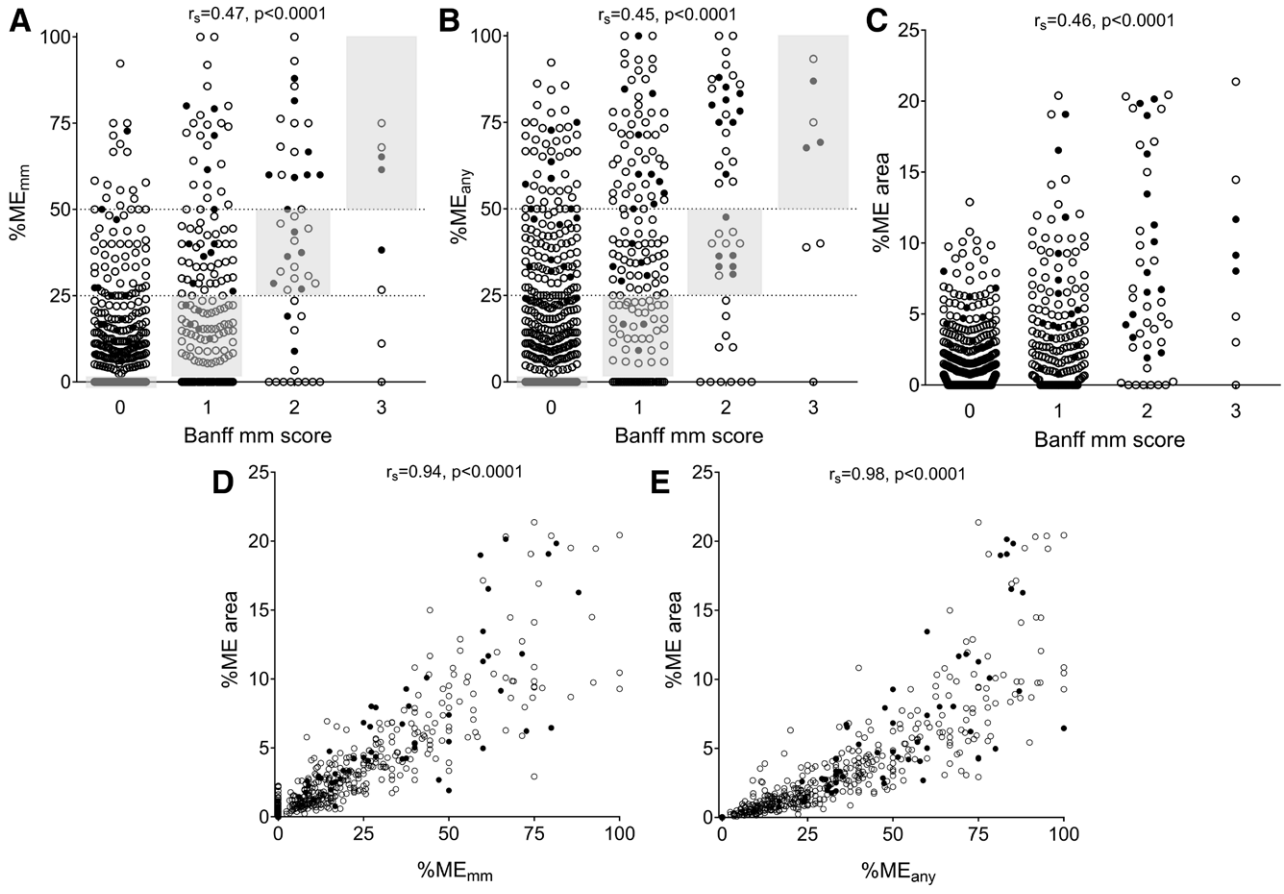

FIGURE 2. Banff *mm* scores show a significant but modest correlation with 3 morphometric ME measures: (A) mm score vs %ME$_{mm}$, (B) mm score vs %ME$_{any}$, (C) *mm* score vs %ME area, (D) %ME area vs ME$_{mm}$, and (E) %ME area vs ME$_{any}$. The full circles represent patients with graft loss, and the empty circles represent patients with still functioning grafts. Gray-shaded areas and dotted lines represent the ranges for the Banff *mm* score. %ME, percentage of mesangial expansion.

had HbA1c levels >9% around the time of the 5-y biopsy to the 202 patients with diabetes with lower HbA1c levels. The higher HbA1c group had a higher %ME$_{any}$ (43.5% versus 31.1%, $P = 0.04$), suggesting that worsening diabetic control correlated with the increasing prevalence of ME. Interestingly, pretransplant hypertension was the only characteristic associated with 5-y Banff *mm* score but not with morphometry ME measures.

Banff *mm* Score, and Morphometry Measures of ME by Graft Status and Pretransplant DM

Allografts that failed had higher mean Banff *mm* scores but similar %ME$_{mm}$, %ME$_{any}$, and %ME area on 1-y biopsy and had a higher mean Banff *mm* scores, %ME$_{mm}$, %ME$_{any}$, and %ME area on 5-y biopsy (**Table S2, SDC,** http://links.lww.com/TXD/A661). Recipients with pretransplant diabetes had similar Banff *mm* scores on 1-y biopsy, higher Banff *mm* scores on 5-y biopsy, and higher %ME$_{mm}$, %ME$_{any}$, and %ME area at both 1- and 5-y biopsies.

Graft and Patient Survival

Banff *mm* score and all morphometric ME measures predicted allograft loss in unadjusted and analyses adjusted for other Banff scores (Table 3). Results did not substantively change when limited to 412 biopsies with available C4d (**Table S3, SDC,** http://links.lww.com/TXD/A661). However, in all 3 models where we added one of each morphometric ME measure, the Banff *mm* score was no longer associated with graft loss. Of the 4 multivariable models, the model with %ME$_{any}$ is most optimally associated with allograft loss as evidenced by the highest c-statistic (0.872). When stratified by scores, grafts with the worst Banff *mm* or morphometric ME scores had worse graft survival (Figure 3) with graft failure rates of 20%–30% within the 5 y following the 5-y biopsy.

The associations between Banff *mm* score and morphometric ME measures with graft loss were similar in sensitivity analyses limited to 698 patients with at least 7 glomeruli on a biopsy, 137 patients with <7 glomeruli on a biopsy, 791 patients with *cg* = 0, and 753 patients with *g* = 0 (**Table S4, SDC,** http://links.lww.com/TXD/A661). To assess the possibility that death with function might be a competing event with graft loss, we performed Fine-Gray Cox Models and found that Banff *mm* score and 3 morphometry ME measures had a similar performance (**Table S5, SDC,** http://links.lww.com/TXD/A661).

Finally, after adjusting for baseline donor and recipient characteristics, recipient kidney function, and recurrent disease at 5 y, Banff *mm* score and all morphometry ME measures are still associated with graft loss (Table 4). The only exception was that after adding HbA1c at 5 y, the Banff *mm* score was no longer significant.

TABLE 2.

Spearman correlations between mesangial expansion scored by pathologist or morphometry with baseline clinical characteristics, those before and around the time of the 5-y biopsy

	Banff *mm* score		$\%ME_{mm}$		$\%ME_{any}$		%ME area	
	1-y biopsy	5-y biopsy	1-y biopsy	5-y biopsy	1-y biopsy	5-y biopsy	1-y biopsy	5-y biopsy
Baseline donor characteristic								
Donor's age	0.02	0.00	–0.01	0.01	0.00	0.02	0.00	0.02
	(0.88)	(0.95)	(0.87)	(0.85)	(0.94)	(0.64)	(0.97)	(0.64)
Donor's BMI	–0.01	0.02	-0.01	0.04	0.00	0.05	–0.01	0.05
	(0.82)	(0.62)	(0.76)	(0.28)	(0.90)	(0.14)	(0.80)	(0.12)
Living donor	**–0.13**	**–0.10**	-0.05	**–0.10 (0.003)**	**–0.12**	**–0.09**	**–0.11**	**–0.10**
	(0.0003)	**(0.005)**	(0.21)		**(0.0006)**	**(0.01)**	**(0.002)**	**(0.004)**
Baseline recipient characteristics								
Recipient's age	0.07	0.05	0.05	**0.09**	**0.13**	**0.10**	**0.12**	**0.11**
	(0.07)	(0.14)	(0.14)	**(0.008)**	**(0.0002)**	**(0.006)**	**(0.001)**	**(0.002)**
Recipient's BMI	**0.09**	**0.15**	**0.14**	**0.21 (<0.0001)**	**0.15**	**0.21**	**0.15**	**0.19**
	(0.01)	**(<0.0001)**	**0.0001**		**<0.0001**	**(<0.0001)**	**<0.0001**	**<0.0001**
White vs other races	–0.05	0.03	0.00	–0.03	0.01	–0.04	0.02	–0.03
	(0.18)	(0.36)	(0.97)	(0.42)	(0.69)	(0.31)	(0.67)	(0.31)
Hypertension	0.02	**0.07**	–0.01	0.00	0.03	0.02	0.02	0.01
	(0.66)	**(0.04)**	(0.78)	(0.94)	(0.34)	(0.66)	(0.56)	(0.69)
Pretransplant diabetes	0.06	**0.10**	**0.08**	**0.28**	**0.09**	**0.28**	**0.10**	**0.27**
	(0.07)	**(0.005)**	**(0.03)**	**(<0.0001)**	**(0.01)**	**(<0.0001)**	**(0.006)**	**(<0.0001)**
HLA mismatch	0.01	0.00	0.00	0.02	0.04	-0.01	0.04	0.00
	(0.76)	(0.92)	(0.98)	(0.58)	(0.29)	(0.83)	(0.31)	(0.92)
Delayed graft function	0.05	**0.09**	0.00	**0.09**	0.05	**0.09**	0.04	**0.10**
	(0.14)	**(0.008)**	(0.93)	**(0.008)**	(0.17)	**(0.01)**	(0.26)	**(0.003)**
Early posttransplant characteristics								
Any rejection within the first y	**0.09**	–0.01	0.00	0.06	0.00	**0.09**	0.00	**0.09**
	(0.01)	(0.70)	(0.97)	(0.11)	(0.96)	**(0.009)**	(0.99)	**(0.006)**
Borderline rejection	0.04	–0.05	0.01	0.01	–0.03	0.04	–0.03	0.04
	(0.27)	(0.16)	(0.88)	(0.69)	(0.34)	(0.26)	(0.38)	(0.20)
Acute T cell–mediated rejection	0.06	–0.01	0.00	0.04	0.01	0.06	0.01	0.07
	(0.12)	(0.72)	(0.91)	(0.23)	(0.82)	(0.06)	(0.80)	(0.05)
Prednisone maintenance therapy	0.01	0.01	–0.02	0.03	0.04	0.05	0.02	0.05
	(0.80)	(0.77)	(0.51)	(0.34)	(0.23)	(0.11)	(0.56)	(0.16)
Characteristics at 5-y posttransplant								
HbA1c at 5-y biopsy	0.01	**0.12**	0.08	**0.18**	**0.10**	**0.18**	**0.10**	**0.18**
	(0.75)	**(0.002)**	(0.06)	**(<0.0001)**	**(0.01)**	**(<0.0001)**	**(0.02)**	**(<0.0001)**
Recurrent disease at 5-y biopsy	**0.15**	**0.16**	**0.08**	**0.18**	**0.11**	**0.17**	**0.11**	**0.18**
	(<0.0001)	**(<0.0001)**	**(0.04)**	**(<0.0001)**	**(0.003)**	**(<0.0001)**	**(0.003)**	**(<0.0001)**

BMI, body mass index; HbA1c, hemoglobin A1c. Values in bold indicate correlations that were statistically significant ($P < 0.05$).

TABLE 3.

Banff scores and morphometry measures of ME at 5-y surveillance biopsies as predictors of allograft failure

Banff scores and morphometry measures at 5-y biopsy	Unadjusted		Multivariable adjusted							
			Banff scores only		Banff scores + $\%ME_{mm}$		Banff scores + $\%ME_{any}$		Banff scores + %ME area	
	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
ci score	**2.11 (1.71- 2.60)**	**<0.0001**	**1.52 (1.00-2.31)**	**0.05**	1.47 (0.95-2.28)	0.09	1.45 (0.94-2.24)	0.09	1.44 (0.94-2.22)	0.10
ct score[a]	**1.99 (1.62-2.45)**	**<0.0001**	–	–	–	–	–	–	–	–
cg score	**1.69 (1.52-1.88)**	**<0.0001**	**1.33 (1.07-1.66)**	**0.01**	**1.36 (1.10, 1.68)**	**0.004**	**1.37 (1.10-1.71)**	**0.005**	**1.29 (1.05-1.58)**	**0.01**
g score	**1.49 (1.31-1.69)**	**<0.0001**	0.92 (0.74-1.14)	0.44	0.99 (0.79-1.23)	0.91	1.00 (0.81-1.24)	0.97	1.03 (0.83-1.28)	0.77
cv score[b]	**1.72 (1.35-2.20)**	**<0.0001**	1.10 (0.80-1.50)	0.56	1.14 (0.83-1.57)	0.43	1.11 (0.81- 1.51)	0.52	1.12 (0.82-1.54)	0.48
ah score	**1.69 (1.33-2.14)**	**<0.0001**	**1.39 (1.07-1.79)**	**0.01**	**1.32 (1.02-1.71)**	**0.03**	**1.29 (1.00-1.67)**	**0.05**	**1.37 (1.07-1.75)**	**0.01**
i score	**1.22 (1.03-1.45)**	**0.02**	0.83 (0.65-1.06)	0.14	0.81 (0.63-1.04)	0.09	0.81 (0.64-1.04)	0.10	0.86 (0.68-1.10)	0.24
ptc score	**1.50 (1.32-1.70)**	**<0.0001**	1.10 (0.91-1.33)	0.32	1.14 (0.94-1.39)	0.18	1.12 (0.92-1.37)	0.24	1.16 (0.96-1.41)	0.12
v score	**1.15 (1.04-1.26)**	**0.005**	**1.23 (1.09-1.38)**	**0.0007**	**1.26 (1.11- 1.42)**	**0.0002**	**1.23 (1.08-1.39)**	**0.001**	**1.23 (1.09-1.38)**	**0.0006**
t score	1.08 (0.86-1.92)	0.50	–	–	–	–	–	–	–	–
ti score	**1.91 (1.59-2.31)**	**<0.0001**	1.25 (0.84-1.88)	0.27	1.28 (0.85-1.95)	0.24	1.28 (0.85-1.93)	0.24	1.22 (0.80-1.84)	0.36
mm score	**1.94 (1.65-2.28)**	**<0.0001**	**1.36 (1.04-1.78)**	**0.02**	1.01 (0.76-1.36)	0.92	0.95 (0.71-1.28)	0.74	1.02 (0.75-1.39)	0.89
$\%ME_{mm}$	**1.89 (1.60-2.23)**	**<0.0001**	–	–	**1.81 (1.43-2.30)**	**<0.0001**	–	–	–	–
$\%ME_{any}$	**2.30 (1.88-2.82)**	**<0.0001**	–	–	–	–	**2.13 (1.61-2.80)**	**<0.0001**	–	–
%ME area	**1.77 (1.55-2.03)**	**<0.0001**	–	–	–	–	–	–	**1.58 (1.27-1.95)**	**<0.0001**
c-statistic[c]	–		0.809/0.791		0.850/0.832		0.872/0.854		0.842/0.828	

[a]Because of high colinearity with the *ci* score, the *ct* score was not used in multivariable models.
[b]*cv* score could not be graded in 9 5-y biopsies.
[c]Second value is a mean c-statistic after 10-fold cross-validation.
ah, arteriolar hyalinosis; cg, transplant glomerulopathy; ci, interstitial fibrosis; CI, confidence interval; ct, tubular atrophy; cv, arterial fibrointimal thickening; HR, hazard ratio; g, glomerulitis; i, inflammation; %ME, percentage of mesangial expansion; ptc, peritubular capillaritis; t, tubulitis; ti, total inflammation; v, intimal arteritis.Values in bold indicate associations that were statistically significant ($P < 0.05$).

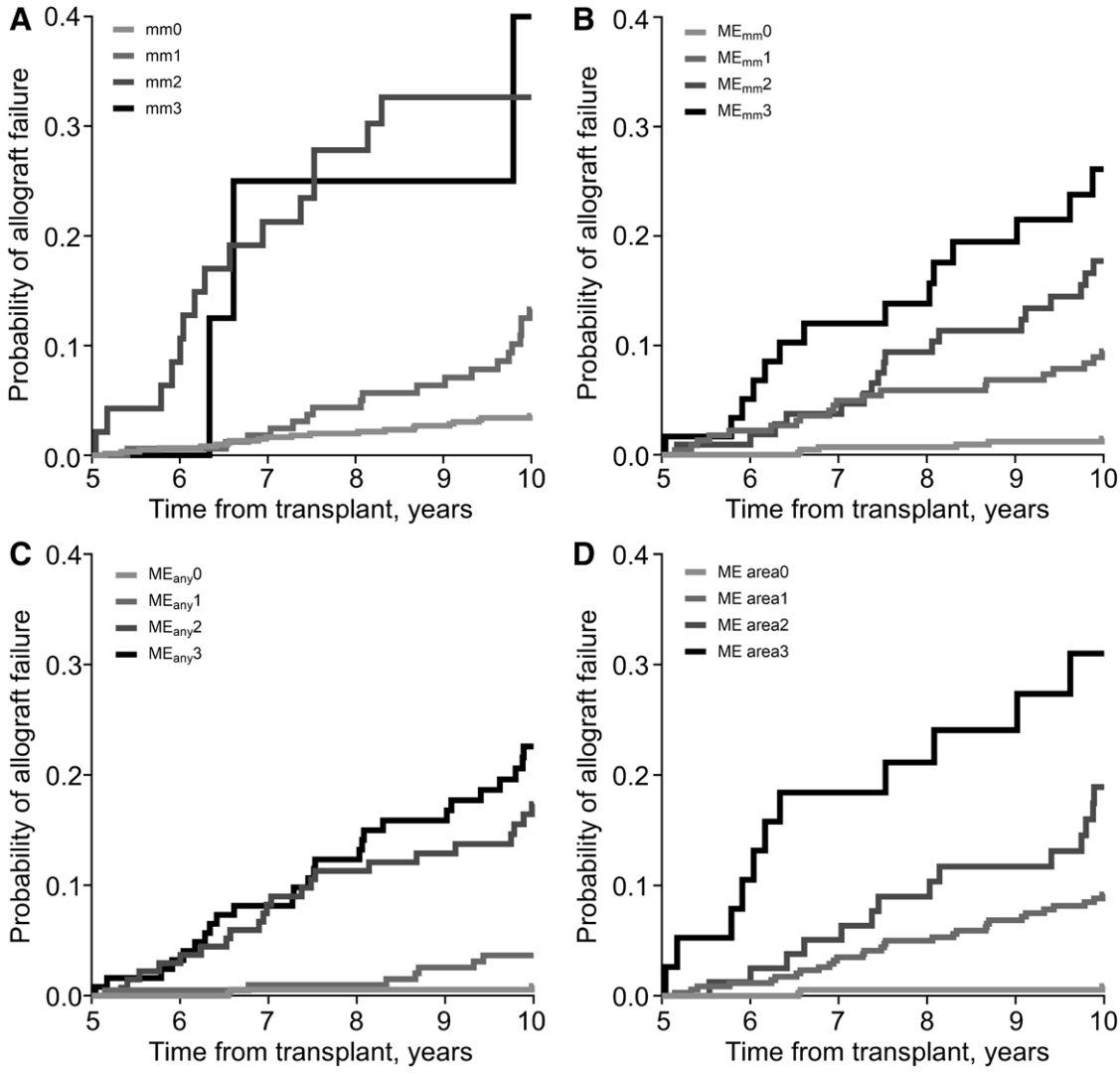

FIGURE 3. Cumulative incidence of allograft failure increased with (A) higher Banff *mm* score, (B) higher morphometric %ME$_{mm}$ score, (C) higher morphometric %ME$_{any}$ score, and (D) higher %ME area score on 5-y surveillance biopsy (log rank $P < 0.0001$ for all). %ME, percentage of mesangial expansion.

Risk Factors for Severe ME

The quartile (n = 209) of biopsies with the highest/worst %ME$_{mm}$ measures (>19.4% of affected glomeruli) was used to determine what putative risk factors were presented in the most severe cases. Of 209, 133 patients (63.6%) had diabetes and/or obesity (BMI >30). Of 76 patients who had neither diabetes nor obesity, 27 biopsies (35.5%) had some history or histologic findings associated with alloimmunity (any rejection in the first year, *cg* > 0 or *g* > 0); 19 biopsies (25%) had no alloimmunity evidence but had DGF, deceased donor or %ischemic glomeruli >25%; and 30 biopsies (39.5%) had no known risk factors. Findings were largely similar when we looked at the distribution of risk factors among 209 biopsies with the worst %ME$_{any}$.

Graft survival varied by the presence of risk factors. In 64 biopsies (30.6%) with the worst %ME$_{mm}$ (quartile 4) with some history or histologic findings of alloimmunity, with or without diabetes and/or obesity, the graft survival in the next 5 y was 71.8%. In 96 patients (45.9%) who had diabetes and/or obesity without evidence of alloimmunity, the graft survival in the next 5 y was 84.2%. Finally, in 49 (23.4%) biopsies in recipients with no diabetes/obesity without any evidence of alloimmunity, the graft survival in the subsequent 5 y was 87.7%.

DISCUSSION

The current study is the largest study to examine ME on both 1- and 5-y kidney transplant surveillance biopsies to determine its incidence, putative risk factors, and impact on graft survival within the first 5 y after the 5-y surveillance biopsy. We found that ME measured by pathologist or morphometry increased from 1 to 5 y, and by 5 y, 55.6% of biopsies had at least 1 focus of ME. Severe ME (>50% of affected glomeruli) was less common; however, 20.1% of biopsies had at least 25% of their glomeruli affected by ME (generally correlating with moderate-to-severe Banff *mm* scores). The correlation between 3 morphometric ME measures was very high; however, at the higher percentage of ME, the total area of ME significantly varied, suggesting that these 2 morphometric ME measures provide 2 dimensions of this lesion, severity and pervasiveness.

TABLE 4.

Measures of mesangial expansion as predictors of graft loss during the first 5 y after the 5-y biopsy. Several multivariable models were performed: (1) adjusted for baseline donor and recipient clinical characteristics, (2) adjusted for recipient kidney function at 5 y, and (3) adjusted for recipient kidney function and HbA1c at 5 y.

	Adjusted for baseline characteristics[a]		Adjusted for characteristics at 5 y[b]		Further adjusted for HbA1c at 5 y[c]	
	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
mm score	1.99 (1.69-2.34)	<0.0001	1.34 (1.09-1.65)	0.006	1.25 (0.97-1.61)	0.08
$\%ME_{mm}$	2.00 (1.66-2.40)	<0.0001	1.35 (1.08-1.69)	0.009	1.36 (1.06-1.74)	0.02
$\%ME_{any}$	2.49 (1.99-3.11)	<0.0001	1.64 (1.29-2.10)	<0.0001	1.67 (1.25-2.22)	0.0005
%ME area	1.90 (1.63-2.22)	<0.0001	1.25 (1.04-1.51)	0.02	1.37 (1.08-1.75)	0.01

[a]Living donor, recipient age, recipient BMI, pretransplant diabetes, delayed graft function, and any rejection in the first year.
[b]eGFR, 24-h protein and recurrent disease.
[c]Analysis limited to 638 patients with available HbA1c.
BMI, body mass index; CI, confidence interval; eGFR, estimated glomerular filtration rate; HbA1c, hemoglobin A1c; HR, hazard ratio; %ME, percentage of mesangial expansion.

A higher level of ME was associated with increased rates of allograft loss in both unadjusted and adjusted analyses, which accounted for other Banff scores. Of all ME measures, $\%ME_{any}$ most optimally predicted allograft loss. Most of these surveillance biopsies were in well-functioning kidneys and the overall survival in the cohort was 92.6% by 10 y posttransplant. However, allografts with the highest ME scores had graft failure rates of 20%–25% by 10 y.

Baseline clinical characteristics that associated with ME included older recipient age, pretransplant diabetes, obesity, DGF, deceased donor, and a history of rejection. It has been previously shown that ME increases with aging.[14-17] This may be because of multiple factors including gene expression alterations, the presence of comorbid conditions, such as hypertension and diabetes, and excessive caloric intake.[16] These changes may be exacerbated in the single kidney model in the kidney transplantation because years after the transplant, glomerular features are more reflective of recipient's than donor's characteristics.[12] Almost all biopsies most severely affected by ME had at least 1 of these putative risk factors. In kidney allografts, ruling out alloimmunity as a cause of injury is difficult, and we found that 30.6% of the biopsies with the worst quartile of $\%ME_{mm}$ had some history or histologic findings associated with alloimmunity. However, almost half of the most severely affected biopsies had only diabetes or obesity as the major risk factor, and these had lower graft survival compared with a similar cohort with mild or no ME. The association between DGF and ME is a novel finding. There is a link between DGF and cellular damage in tubular compartment because of ischemic/reperfusion injury[18]; however, the long-term sequelae of ME in glomeruli are currently unclear. Studies have shown that tubular epithelial cells and mesangial cells in glomeruli can express toll-like receptors without the complement system as a response to ischemic injury.[19,20] Another study demonstrated that a single nucleotide polymorphism in toll-like receptor 3 was more prevalent in patients with DGF.[21] Thus, it is possible that a combination of all donor-derived (ischemic injury, inflammatory signaling) and recipient-derived contributors (reperfusion injury, innate and/or adaptive immune response)[18] over a long period of time leads to an increase in ME. Taken together, these data suggest that most cases of moderate-to-severe ME are associated with 3 processes: diabetes/obesity, alloimmunity, and ischemic injury. In addition, these data agree with our prior studies suggesting that diabetes is a risk factor for graft loss and extend the concept that ME is one of the major histologic findings associated with graft loss in patients who, at the time of transplant, were diabetic and/or obese. Moreover, in several multivariable models, we showed that ME scored by either pathologist or morphometry associated with allograft loss independently of baseline clinical characteristics, as well as kidney function and HbA1c at 5 y. Thus, %ME at 5 y might be used as a surrogate endpoint for therapeutic interventions in patients with diabetes or obesity aimed at improving outcomes and deserves further study. Future studies are needed to explore pathophysiology in nearly 40% of biopsies with the worst ME measures that had no known risk factors.

It is well recognized that the interobserver reproducibility of Banff *mm* scores is poor to modest, with kappa values up to 0.37,[10] similar to the reproducibility found in this study. In clinical practice, the mm score may be overscored because of the presence of other histologic features or because it is influenced by clinical history/presentation. In contrast, it is difficult to identify early mesangial matrix expansion, especially in cases with ≥20 glomeruli, which may lead to underscoring. In a routine clinical setting, a pathologist needs to review 10 slides per case on average and does not have time to closely inspect every glomerulus for ME lesions. Conversely, the interobserver intraclass correlation of 3 morphometric ME measures was around 84%. Our proof-of-concept study supports the notion that a more quantitative assessment of mesangial matrix may be more sensitive in identifying early mesangial matrix expansion. Similarly, the outcomes analysis indicates that these early lesions are associated with adverse allograft survival.

This study had several potential limitations. First, our study population was predominantly White, so we could not meaningfully study race differences. Morphometry was performed only on a single Periodic acid Schiff–stained slide, compared with the traditional review of up to 10 diagnostic slides in usual pathology practice. Electron microscopy and immunofluorescence were not available in

these biopsies. Despite better interobserver reproducibility of morphometric ME measures, there is still a possibility that individual ME lesions were not always precisely annotated, as it is sometimes difficult to ascertain what is mesangium and what is not, and when does the ME stop and the endothelium or glomerular basement membrane begins. Nevertheless, this imprecision would affect the measure of %ME area but not the other 2 ME measures that quantify the percentage of affected glomeruli. We did not distinguish between cellular mesangial proliferation and mesangial matrix expansion, as this would significantly extend the time required to perform manual morphometry. Morphometry takes about 10 min to complete 1 case; thus, as such morphometry is impractical for routine clinical practice. However, it is necessary to take an interim step to better understand the importance of ME lesions, and it is a permanent record that can be used for the training of future deep learning models[22-25] which is designed to reduce the analysis time. Despite the extensive follow-up of transplant recipients, we could not assess rejection episodes after 1 y, and whether recurrent or de novo immune complex diseases, posttransplant immunologic (donor-specific antibody) or nonimmunologic (drug toxicity) events were associated ME. Finally, findings cannot be generalized to all transplant biopsies, including indication biopsies. The most notable strengths of our study are: (1) this is the largest study to date that investigated ME in such detail on 5-y surveillance biopsies and its effects on graft survival within the first 5 y; and (2) detailed follow-up on many patients allowing careful graft outcome analyses.

In summary, ME increased from 1- to 5 y, and is relatively common at 5 y after kidney transplantation. Scoring using a more quantitative morphometric method is more accurate than visual estimation by a pathologist and better predicts graft failure after adjusting for other Banff scores and kidney function. Although ME appears related to alloimmune-mediated damage in some instances, it could also be an important biomarker to study nonalloimmune causes of graft loss, such as long-term diabetes. These data suggest that improving long-term renal allograft survival may benefit from the protocols aimed to prevent or ameliorate ME.

ACKNOWLEDGMENTS

The authors thank Miloš Denić for assistance with computer algorithms for batch processing and extraction of biopsy annotations data. We also thank Mollie J. Luhman for her assistance in scanning the biopsy slides.

REFERENCES

1. Matas AJ, Leduc R, Rush D, et al. Histopathologic clusters differentiate subgroups within the nonspecific diagnoses of CAN or CR: preliminary data from the DeKAF study. *Am J Transplant*. 2010;10:315–323.
2. Matas AJ, Fieberg A, Mannon RB, et al. Long-term follow-up of the DeKAF cross-sectional cohort study. *Am J Transplant*. 2019;19:1432–1443.
3. Kukla A, Alexander MP, Turkevi-Nagy S, et al. Mesangial expansion at 5 years predicts death and death-censored graft loss after renal transplantation. *Clin Transplant*. 2021;35:e14147.
4. Itabashi Y, Oguchi H, Mikami T, et al. Prognostic value of the 7-year protocol biopsy of adult kidney allografts: impact of mesangiosclerosis and proteinuria. *Ren Fail*. 2023;45:2197499.
5. Dalla Vestra M, Saller A, Mauer M, et al. Role of mesangial expansion in the pathogenesis of diabetic nephropathy. *J Nephrol*. 2001;14(Suppl 4):S51–S57.
6. Abrass CK. Diabetic nephropathy. Mechanisms of mesangial matrix expansion. *West J Med*. 1995;162:318–321.
7. Steffes MW, Osterby R, Chavers B, et al. Mesangial expansion as a central mechanism for loss of kidney function in diabetic patients. *Diabetes*. 1989;38:1077–1081.
8. Racusen LC, Solez K, Colvin RB, et al. The Banff 97 working classification of renal allograft pathology. *Kidney Int*. 1999;55:713–723.
9. Denic A, Morales MC, Park WD, et al. Using computer-assisted morphometrics of 5-year biopsies to identify biomarkers of late renal allograft loss. *Am J Transplant*. 2019;19:2846–2854.
10. Ozluk Y, Blanco PL, Mengel M, et al. Superiority of virtual microscopy versus light microscopy in transplantation pathology. *Clin Transplant*. 2012;26:336–344.
11. Gaston RS, Fieberg A, Hunsicker L, et al. Late graft failure after kidney transplantation as the consequence of late versus early events. *Am J Transplant*. 2018;18:1158–1167.
12. Denic A, Bogojevic M, Subramani R, et al. Changes in glomerular volume, sclerosis, and ischemia at 5 years after kidney transplantation: incidence and correlation with late graft failure. *J Am Soc Nephrol*. 2023;34:346–358.
13. Conover WJ. *Practical Nonparametric Statistics*. 3rd ed. Wiley; 1999.
14. Bolton WK, Sturgill BC. Spontaneous glomerular sclerosis in aging Sprague-Dawley rats. II. Ultrastructural studies. *Am J Pathol*. 1980;98:339–356.
15. Lim JH, Kim EN, Kim MY, et al. Age-associated molecular changes in the kidney in aged mice. *Oxid Med Cell Longev*. 2012;2012:171383.
16. Wiggins JE. Aging in the glomerulus. *J Gerontol A Biol Sci Med Sci*. 2012;67:1358–1364.
17. Noordmans GA, Caputo CR, Huang Y, et al. Genetic analysis of mesangial matrix expansion in aging mice and identification of Far2 as a candidate gene. *J Am Soc Nephrol*. 2013;24:1995–2001.
18. Siedlecki A, Irish W, Brennan DC. Delayed graft function in the kidney transplant. *Am J Transplant*. 2011;11:2279–2296.
19. Anders HJ, Banas B, Schlondorff D. Signaling danger: toll-like receptors and their potential roles in kidney disease. *J Am Soc Nephrol*. 2004;15:854–867.
20. Wu H, Chen G, Wyburn KR, et al. TLR4 activation mediates kidney ischemia/reperfusion injury. *J Clin Invest*. 2007;117:2847–2859.
21. Kruger B, Banas MC, Walberer A, et al. A comprehensive genotype-phenotype interaction of different toll-like receptor variations in a renal transplant cohort. *Clin Sci (Lond)*. 2010;119:535–544.
22. Ginley B, Lutnick B, Jen KY, et al. Computational segmentation and classification of diabetic glomerulosclerosis. *J Am Soc Nephrol*. 2019;30:1953–1967.
23. Hermsen M, Ciompi F, Adefidipe A, et al. Convolutional neural networks for the evaluation of chronic and inflammatory lesions in kidney transplant biopsies. *Am J Pathol*. 2022;192:1418–1432.
24. Lutnick B, Manthey D, Becker JU, et al; Kidney Precision Medicine Project. A user-friendly tool for cloud-based whole slide image segmentation with examples from renal histopathology. *Commun Med (Lond)*. 2022;2:105.
25. Lutnick B, Murali LK, Ginley B, et al. Histo-fetch—on-the-fly processing of gigapixel whole slide images simplifies and speeds neural network training. *J Pathol Inform*. 2022;13:7.