Reproducibility of Deceased Donor Kidney Procurement Biopsies

S. Ali Husain1,2, Kristen L. King1,2, Ibrahim Batal3, Geoffrey K. Dube1, Isaac E. Hall2,4, Corey Brennan2,5, M. Barry Stokes1, R. John Crew1, Dustin Carpenter6, Hector Alvarado Verduzco1, Raphael Rosen7, Shana Coley3, Eric Campenot1, Dominick Santoriello7, Adler Perotte7, Karthik Natarajan7, Vivette D’Agati1, David J. Cohen1, Lloyd E. Ratner6, Glen Markowitz7, and Sumit Mohan1,2,8

Abstract

Background and objectives Unfavorable histology on procurement biopsies is the most common reason for deceased donor kidney discard. We sought to assess the reproducibility of procurement biopsy findings.

Design, setting, participants, & measurements We compiled a continuous cohort of deceased donor kidneys transplanted at our institution from 1/1/2006 to 12/31/2016 that had at least one procurement biopsy performed, and excluded cases with missing biopsy reports and those used in multiorgan transplants. Suboptimal histology was defined as the presence of advanced sclerosis in greater than or equal to one biopsy compartment (glomeruli, tubules/interstitium, vessels). We calculated k coefficients to assess agreement in optimal versus suboptimal classification between sequential biopsy reports for kidneys that underwent multiple procurement biopsies and used time-to-event analysis to evaluate the association between first versus second biopsies and patient and allograft survival.

Results Of the 1011 kidneys included in our cohort, 606 (60%) had multiple procurement biopsies; 98% had first biopsy performed at another organ procurement organization and their second biopsy performed locally. Categorical agreement was highest for vascular disease (κ=0.17) followed by interstitial fibrosis and tubular atrophy (κ=0.12) and glomerulosclerosis (κ=0.12). Overall histologic agreement (optimal versus suboptimal) was κ=0.15. First biopsy histology had no association with allograft survival in unadjusted or adjusted analyses. However, second biopsy optimal histology was associated with a higher probability of death-censored allograft survival, even after adjusting for donor and recipient factors (adjusted hazard ratio, 0.50; 95% confidence interval, 0.34 to 0.75; P=0.001).

Conclusions Deceased donor kidneys that underwent multiple procurement biopsies often displayed substantial differences in histologic categorization in sequential biopsies, and there was no association between first biopsy findings and post-transplant outcomes.


Introduction

Despite the shortage of kidneys available for transplantation in the United States, the proportion of deceased donor kidneys that are discarded is rising and has reached 20% (1–3). Prior analyses have demonstrated that organ discards are not a result of organ quality alone, and many discarded kidneys are likely viable (3–8). The most commonly cited reason for deceased donor kidney discard is unfavorable histology on procurement biopsies (i.e., frozen section biopsies performed around the time of organ recovery), accounting for 38% of all discards (3,8).

Significant emphasis continues to be placed on kidney biopsy findings during the allocation process, despite previous analyses that have demonstrated a lack of an association between procurement biopsy findings and post-transplant outcomes after adjusting for clinical donor factors that are considered during the allocation process (9–12). In contrast, favorable histologic findings on reperfusion biopsies (i.e., performed after implantation) correlate with longer allograft survival (12,13). The absence of robust associations of procurement biopsies with outcomes is likely multifactorial. Unlike formalin-fixed reperfusion needle core biopsies, procurement biopsies are typically frozen sections, oversample scarred subcapsular cortex due to wedge biopsy technique, use a single (hematoxylin and eosin) stain, and are interpreted by on-call pathologists with often limited expertise in kidney pathology (10,13–17).

Deceased donor kidneys may undergo multiple procurement biopsies prior to transplantation, typically when the organ is exported from one donation service area to another, reflecting clinical practice patterns at the importing organ procurement organization (OPO). Prior studies of sequential biopsies in...
experimental (16) or clinical (10,12) settings have demonstrated poor concordance between multiple biopsies of the same kidney, but these have been limited by small samples or lacked comparative outcomes analyses. We, therefore, identified a unique cohort of deceased donor kidneys that underwent multiple independent procurement biopsies to better assess the reproducibility and predictive value of histologic findings in sequential procurement biopsies. We hypothesized that there would be poor agreement between sequential biopsies of the same kidney and that neither the first nor second biopsy results would be associated with post-transplant outcomes.

Materials and Methods

Study Population

We compiled a continuous retrospective cohort of all deceased donor kidneys transplanted at Columbia University Medical Center from 1/1/2006 to 12/31/2016 that had at least one procurement biopsy (n=1049) (Figure 1). We

![Flow diagram of study cohort](image-url)

---

**Figure 1.** Flow diagram of study cohort. CUMC, Columbia University Medical Center; OPO, organ procurement organization.

---

Abbreviations: CUMC, Columbia University Medical Center; OPO, organ procurement organization.
excluded kidneys with missing biopsy reports ($n=6$) or follow-up data ($n=1$) and those transplanted as part of multiorgan transplants (i.e., dual kidney transplant and multiorgan transplants; $n=31$). This study was approved by the Columbia University Medical Center Institutional Review Board. All clinical and research activities associated with this study were consistent with the principles of the Declaration of Istanbul.

**Procurement Biopsy Results and Scoring**

All procurement biopsies were performed as part of routine clinical care, and the results of these biopsies were compiled retrospectively. Biopsy details were obtained from DonorNet. Information regarding glomerulosclerosis, interstitial fibrosis and tubular atrophy (IFTA), and vascular disease as reported by the interpreting pathologists was obtained for each biopsy directly from procurement biopsy reports. For each of these histologic parameters, we assigned a score of zero (most favorable) to three (least favorable) on the basis of thresholds outlined in Supplemental Table 1. These category thresholds were based upon biopsy report sheets at our local OPO. If a range of values was reported for any compartment, then the lower end of the range was used to assign the histologic score (e.g., IFTA “mild to moderate” was scored as one).

“Optimal” histology on a given biopsy was defined as having a score less than or equal to one for all three biopsy compartments, and bilateral concordance was defined as matching optimal versus suboptimal categorization for both kidneys (right and left) from the same donor. Compartment concordance for any individual biopsy was defined as homogeneity in the presence or absence of advanced sclerosis (i.e., histologic score greater than or equal to two) in each of the three components in that biopsy (13).

In cases where the kidney’s biopsy report was missing information for one of the biopsy compartments in either the first biopsy ($n=0$ for glomerulosclerosis, $n=44$ of 1011 for IFTA, and $n=85$ of 1011 for vascular disease [of which 36 were missing both IFTA and vascular disease]) or second biopsy ($n=0$ for glomerulosclerosis, $n=12$ of 606 for IFTA, and $n=15$ of 606 for vascular disease), histology was considered optimal if each of the remaining compartments had a score of less than or equal to one. Compartment concordance for these cases was defined as concordance of the compartments that were recorded.

**Donor and Recipient Variables**

We obtained demographic and clinical data for all donors and recipients from the medical record. Transplant-specific variables and outcomes were also obtained from the medical record. We calculated the kidney donor risk index (KDRI) and subsequently, the kidney donor profile index for each donor using a 2015 reference as recommended by the Organ Procurement and Transplantation Network (18). The primary end point of interest was death-censored allograft survival. Secondary end points were overall allograft survival (a composite of death and allograft failure) and recipient survival.

**Statistical Analyses**

Descriptive statistics were used to compare donor and recipient demographic variables between kidneys that underwent one versus multiple biopsies. We similarly compared histologic findings between groups. All subsequent analyses included only kidneys that underwent multiple procurement biopsies. A small number ($n=10$) of kidneys in our cohort underwent three procurement biopsies; in these cases, the first and last biopsies were used for comparisons of sequential biopsies.

Concordance between sequential biopsies for each histologic parameter was calculated using a $4 \times 4$ frequency table, and percentage concordance was calculated as the proportion of kidneys with the same score for each parameter. Agreement for each of these parameters was assessed by calculating a simple $\kappa$ coefficient, which (unlike percentage concordance) takes into account the likelihood of agreement by chance (note that $\kappa$ ranges from zero to one, where one indicates perfect agreement).

Unadjusted death-censored graft survival analysis and patient survival analysis were performed using the Kaplan–Meier method and log rank test. We performed univariable and multivariable time-to-event analyses for death-censored graft failure using Cox proportional hazards models. Patients who were lost to follow-up were censored on the last date that follow-up was available. Analyses were performed using Stata MP 15 (StataCorp, College Station, TX). Statistical significance was identified by a two-sided $\alpha<0.05$.

**Results**

A total of 1011 kidneys transplanted at our institution between 1/1/2006 and 12/31/2016 after undergoing $\geq 1$ procurement biopsy were included in our analysis, with a mean of 1.6 biopsies per kidney (Figure 1). Of these, 606 (60%) had multiple procurement biopsies performed before transplantation. Almost all (98%) kidneys with multiple procurement biopsies had their first biopsy performed at an outside OPO and their second biopsy performed at our local OPO.

**Characteristics of Kidneys That Underwent One Versus Multiple Biopsies**

Donors of kidneys that underwent a single biopsy at our local OPO were significantly younger and had a lower mean KDRI compared with donors whose kidneys underwent multiple procurement biopsies or a single biopsy elsewhere (Table 1). Kidneys in each of these groups had a similar number of glomeruli sampled. Biopsy technique could not be determined for 44% of first biopsies; when this information was included, wedge biopsy technique was most common among kidneys that underwent multiple biopsies.

Kidneys that underwent a single biopsy at our local OPO were also the least likely to have advanced sclerosis in each of the histologic compartments (glomeruli, interstitium, and vessels) (Table 1). Additionally, these kidneys were mostly likely to have biopsy reports of concordant histology between histologic compartments as well as histologic concordance with their contralateral mate kidney.

**Agreement between Sequential Biopsies**

Among the 606 kidneys that underwent multiple procurement biopsies, there was poor overall agreement
between sequential biopsy findings in each of the three histologic compartments (Figure 2). Categorical agreement was highest for vascular disease with $\kappa=0.17$ (59% concordance) followed by IFTA ($\kappa=0.12$; 71% concordance) and glomerulosclerosis ($\kappa=0.12$; 44% concordance) (Figure 2). Because glomerulosclerosis was recorded as a continuous variable in all biopsy reports, we also compared the correlation between raw glomerulosclerosis on first and second biopsies. The correlation between first and second biopsy glomerulosclerosis was weak ($r^2=0.12$) (Supplemental Figure 1).

Optimal histology was identified in 383 kidneys (63%) on first biopsy and 495 kidneys (82%) on second biopsy; 332 of these (55% of all kidneys) had optimal histology on both biopsies. Only 10% ($n=60$) of kidneys had suboptimal histology on both biopsies, a bias inherent in our design. Overall histologic concordance (optimal versus suboptimal) was 65% ($\kappa=0.15$), including 85% for kidneys with first biopsy that was categorized as optimal but only 27% for those with first biopsy that was categorized as suboptimal. Kidneys with discordant histology on sequential biopsies were more likely to have had a first biopsy with advanced glomerulosclerosis (66% versus 15% for kidneys with concordant histology), IFTA (5% versus 1%), and vascular disease (22% versus 6%, $P<0.01$) than those with concordant second biopsy histology (all $P<0.01$) (Supplemental Table 2). They were also less likely to have concordance on their first biopsy within histologic compartments (18% versus 53%) or with their contralateral partner (63% versus 85%) and tended to be from older donors (50±8 versus 45±12 years) and donors with lower final creatinine (1.59±1.02 versus 2.26±1.58 mg/dl) and higher KDRI (1.27±0.27 versus 1.19±0.29; all $P<0.01$) (Supplemental Table 2). The large proportion of these biopsies missing information about technique (54%) precluded meaningful comparison of core versus wedge biopsy.

### Association with Post-Transplant Outcomes

Optimal histologic classification on first biopsy was not associated with a higher likelihood of death-censored allograft survival compared with suboptimal first biopsy.
histology in univariate analysis \((P=0.72)\). In contrast, optimal histology on second biopsy was associated with a higher likelihood of death-censored allograft survival (hazard ratio for graft failure, 0.54; 95% confidence interval, 0.36 to 0.81) (Figure 3). Furthermore, among kidneys with optimal histology on the second biopsy, there was no difference in death-censored allograft survival among those with first biopsy that showed suboptimal histology compared with those with optimal histology on both biopsies \((P=0.96)\) (Supplemental Figure 2). Results were similar using overall (rather than death-censored) graft survival (Supplemental Figure 3). The association between second biopsy histology and death-censored allograft survival persisted after adjusting for KDRI, cold ischemia time, number of HLA mismatches, recipient age, and recipient dialysis time (adjusted hazard ratio for graft failure, 0.50; 95% confidence interval, 0.34 to 0.75) (Table 2).

Neither first nor second biopsy histology was associated with patient survival.

**Discussion**

There are nearly 100,000 candidates on the kidney transplant waitlist in the United States, of which over 8000 are removed from the waitlist annually due to death or deterioration of health (19). Despite this severe organ shortage, nearly 20% of kidneys procured from deceased donors are discarded, with unfavorable procurement biopsy findings being the most common reason for discard (1,3). Although there is belief that procurement biopsy findings provide reliable, objective information about kidney quality, prior studies have suggested that they are poorly reproducible and have limited or no association with post-transplant outcomes (10,12,14). Here, we demonstrate in a...
large cohort of deceased donor kidneys that sequential procurement biopsies on the same kidney often yield considerably discordant results and that only repeat biopsies using a consistent, standardized approach at our local OPO yielded histologic findings associated with allograft survival. The biopsies performed at our local OPO are needle core biopsies that are examined by a single center’s pathology group, calling into question the current lack of standardization in how biopsies are obtained, processed, and read across the country.

Kidneys that underwent a first biopsy at another OPO were declined for use at local transplant centers prior to export to our OPO, suggesting that these organs were viewed unfavorably and shipped out as a result. The second biopsy performed at our OPO appears to have allowed the recognition that these kidneys could in fact be appropriately used to the benefit of the patients who received them. Although the reason for performing a second biopsy on the kidneys in our study was not explicitly recorded, given that this decision was made in the clinical setting (requested by the accepting surgeon), it is likely that our cohort over-represents kidneys for which results of the first procurement biopsy were suspected to be inaccurate. Furthermore, it is also likely that a subset of kidneys with concordant first and second biopsies—those with suboptimal histology on both—was discarded and therefore, largely excluded from the analysis. Additional studies that identify these kidneys and their biopsy findings are needed.

However, although this selection bias may overestimate the degree of discordance between sequential biopsies, it nevertheless highlights the problems in organ quality assessment that can result from over-reliance on procurement biopsy results. Overall histologic categorization was discordant in 35% of repeat biopsies, and κ agreement statistics for each histologic compartment ranged from just 0.12 to 0.17. The fact that almost all of the kidneys with multiple biopsies in our study were imported from other OPOs suggests that—despite a lack of association with death-censored allograft survival even after adjusting for donor and recipient characteristics. Furthermore, the consideration of first and second biopsy results together did not provide any additional value to the information from second biopsy results alone as demonstrated by the fact that there was no difference in allograft survival among kidneys with optimal second biopsy histology when their first biopsy showed optimal versus suboptimal histology. Prior studies have provided mixed results about the association of procurement biopsy results with allograft outcomes (especially after adjusting for donor age, which is associated with nephrosclerosis independent of kidney function), including our recent study that also demonstrated poor concordance with reperfusion biopsy findings (10,12,14,20–27). Our findings suggest that these inconsistent findings are potentially due to differences in biopsy quality at different OPOs, and standardization of technique and interpretation to establish best practices may allow procurement biopsies to yield more meaningful results to aid clinicians during the organ allocation process.

Specific features of the biopsies performed at our local OPO may provide further insight into how this goal can be achieved. By policy, procurement biopsies performed by our OPO are 14-gauge needle core biopsies. Prior analysis found that needle biopsies were superior to wedge biopsies for evaluating vascular disease in kidneys from living donors and that wedge biopsies often oversample scarred subcapsular tissue (15,16). Despite a theoretical concern about increased bleeding risk with needle core biopsies (28), the consistent demonstration of the limitations of

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. with Optimal Histology</th>
<th>Total No. of Graft Failures</th>
<th>Unadjusted Model</th>
<th>Adjusted Model*</th>
</tr>
</thead>
<tbody>
<tr>
<td>First biopsy optimal</td>
<td>373</td>
<td>117</td>
<td>0.93 0.64 to 1.36 0.72</td>
<td>0.94 0.64 to 1.38 0.75</td>
</tr>
<tr>
<td>Second biopsy optimal</td>
<td>489</td>
<td>117</td>
<td>0.54 0.36 to 0.81 0.003</td>
<td>0.50 0.34 to 0.75 0.001</td>
</tr>
</tbody>
</table>

*HR, hazard ratio; 95% CI, 95% confidence interval.
*Adjusted for kidney donor risk index, recipient age at transplant, recipient dialysis time at transplant, cold ischemia time, and number of HLA mismatches.

from the same kidney, with over-representation of glomerulosclerosis in the subcapsular region. Additionally, Kasiske et al. (10) found marked discordance in sequential biopsies performed on 64 deceased donor kidneys as part of routine care, with an r² for glomerulosclerosis of only 0.25. We similarly found that sequential biopsies often yielded very different results, and we demonstrate that a second biopsy was more likely to display discordant results when the first biopsy demonstrated advanced sclerosis, a lack of concordance with findings from the mate kidney from the same donor, and inconsistent degree of sclerosis across biopsy compartments. The presence of these features should alert OPOs and clinicians that a second biopsy may be warranted before biopsy results are released to clinicians making organ disposition decisions.

Although first biopsy histology had no association with outcomes, second biopsy histology was associated with death-censored allograft survival even after adjusting for donor and recipient characteristics. Furthermore, the consideration of first and second biopsy results together did not provide any additional value to the information from second biopsy results alone as demonstrated by the fact that there was no difference in allograft survival among kidneys with optimal second biopsy histology when their first biopsy showed optimal versus suboptimal histology. Prior studies have provided mixed results about the association of procurement biopsy results with allograft outcomes (especially after adjusting for donor age, which is associated with nephrosclerosis independent of kidney function), including our recent study that also demonstrated poor concordance with reperfusion biopsy findings (10,12,14,20–27). Our findings suggest that these inconsistent findings are potentially due to differences in biopsy quality at different OPOs, and standardization of technique and interpretation to establish best practices may allow procurement biopsies to yield more meaningful results to aid clinicians during the organ allocation process.

Specific features of the biopsies performed at our local OPO may provide further insight into how this goal can be achieved. By policy, procurement biopsies performed by our OPO are 14-gauge needle core biopsies. Prior analysis found that needle biopsies were superior to wedge biopsies for evaluating vascular disease in kidneys from living donors and that wedge biopsies often oversample scarred subcapsular tissue (15,16). Despite a theoretical concern about increased bleeding risk with needle core biopsies (28), the consistent demonstration of the limitations of
wedge biopsies suggests that they should not be used for procurement biopsies. The large volume of missing data that we observed regarding core versus wedge technique also suggests that this aspect of the biopsy is not adequately considered at the time of allocation. Although we lack details of the training of each pathologist, interobserver variability may also have contributed to the differences that we observed, possibly reflecting different levels of experience among the examining pathologists (14). OPOs should be selective about who they allow to interpret procurement biopsies, especially in an era when slides can be digitized and read remotely. Using this information to improve the quality of the initial procurement biopsy can potentially reduce the need for repeat biopsies, thereby avoiding risks of additional biopsies, including extended cold ischemia, increased cost, and risk of bleeding from additional biopsy sites. Further studies are needed to determine whether such standardization can yield procurement biopsy results that resemble the histology seen on high-quality postimplantation biopsies. Until then, it seems that performing a second procurement biopsy to confirm unfavorable histologic findings is reasonable for kidneys with first biopsy that is performed at an OPO that uses less than ideal biopsy techniques. Finally, although we found that second biopsy findings were associated with post-transplant outcomes, it is critical to note that even kidneys with suboptimal histology demonstrate acceptable post-transplant outcomes and likely provide an overall benefit when transplanted in appropriate recipients (13).

Strengths of our study include the largest cohort of deceased donor kidneys with multiple procurement biopsies and extended follow-up. However, potential selection bias at three points should be considered. First, not every kidney undergoes a procurement biopsy, and criteria for these biopsies or repeat procurement biopsies are not standardized. It is likely that kidneys that underwent multiple biopsies were those in which there was suspicion that the first histologic interpretation was unreliable. However, even if this is true, the discordant findings on second biopsies provide evidence that, as currently performed and interpreted, many procurement biopsies provide findings that are not reproducible. Second, kidneys that underwent multiple biopsies were overwhelmingly first biopsied at other OPOs before being exported to ours. Because almost all of the second biopsies were performed and interpreted at our local OPO, it is possible that the results would differ at another center affiliated with a different OPO. Third, it is possible that kidneys with advanced sclerosis on either biopsy were discarded and that the absence of these kidneys and their potential worse outcomes biased our results. Our study is also limited by the large number of biopsies missing information about biopsy technique, which decreases our ability to draw stronger conclusions about the relationship between biopsy technique and reproducibility.

In conclusion, deceased donor kidneys that underwent multiple procurement biopsies prior to transplantation often displayed substantial differences in histologic categorization in sequential biopsies, and there was no association between first biopsy findings and post-transplant outcomes. This poor reproducibility of procurement biopsy histology provides further evidence against the over-reliance on procurement biopsy results in organ utilization decisions without improved standardization in biopsy technique and interpretation. Standardization in procurement biopsy technique and interpretation to reflect best practices may improve the utility of procurement biopsies in aiding organ utilization decisions.

Acknowledgments

Preliminary results of this analysis were presented in abstract form at the National Kidney Foundation Spring Clinical Meeting 2019.

Disclosures

Dr. Cohen reports a position on the scientific advisory board for the Global aHUS Registry and one-time advisory board positions at ITB Pharmaceuticals and Natera outside of the submitted work. Dr. Mohan reports a position on the scientific advisory board at Angion Pharmaceuticals and receiving personal fees from Kidney International Reports. All other authors have nothing to disclose.

Funding

Dr. Husain is supported by National Center for Advancing Translational Sciences grant KL2 TR001874 and a National Kidney Foundation Young Investigator Grant. Dr. Batal is supported by a grant from the American Society of Transplantation Research Network. Dr. Natarajan is supported by grants from the National Institutes of Health and the National Center for Advancing Translational Sciences. Dr. Mohan is supported by National Institute of Diabetes and Digestive and Kidney Disorders grants R01-DK114893-01 and U01-DK116066-01.

Supplemental Material

This article contains the following supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.09170819//DCSupplemental.

Supplemental Figure 1. Correlation between first and second biopsy glomerulosclerosis (dashed line is a reference line for concordance).

Supplemental Figure 2. Unadjusted Kaplan–Meier curves of death-censored allograft survival for kidneys with optimal second biopsy histology on the basis of first procurement biopsy histology.

Supplemental Figure 3. Unadjusted Kaplan–Meier curves of overall allograft survival on the basis of (A) first procurement biopsy histology and (B) second procurement biopsy histology (B).

Supplemental Table 1. Histologic scoring for each biopsy parameter.

Supplemental Table 2. Characteristics of kidneys with concordant versus discordant histology on sequential biopsies.

References


Received: August 6, 2019 Accepted: December 5, 2019

Published online ahead of print. Publication date available at www.cjasn.org.