

# Reproducibility of Chronic Changes on High-Quality Deceased Donor Kidney Allograft Biopsies



Hector Alvarado Verduzco<sup>1</sup>, Ibrahim Batal<sup>2</sup>, Sumit Mohan<sup>1,3</sup> and S. Ali Husain<sup>1</sup>

<sup>1</sup>Division of Nephrology, Department of Medicine, Columbia University Vagelos College of Physicians & Surgeons and New York Presbyterian Hospital, New York, New York, USA; <sup>2</sup>Department of Pathology & Cell Biology, Columbia University Vagelos College of Physicians & Surgeons and New York Presbyterian Hospital, New York, New York, USA; and <sup>3</sup>Department of Epidemiology, Columbia University Mailman School of Public Health, New York, New York, USA

**Correspondence:** S. Ali Husain, Division of Nephrology, Department of Medicine, Columbia University Vagelos College of Physicians & Surgeons and New York Presbyterian Hospital, 622 West 168th Street PH4-124, New York, New York 10032, USA. E-mail: [sah2134@cumc.columbia.edu](mailto:sah2134@cumc.columbia.edu)

Received 6 November 2021; revised 13 December 2021; accepted 14 December 2021; published online 27 December 2021

*Kidney Int Rep* (2022) 7, 889–891; <https://doi.org/10.1016/j.ekir.2021.12.022>

© 2021 International Society of Nephrology. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The assessment of chronic changes in deceased donor kidney procurement biopsies contributes to organ underutilization, but these biopsies have poor reproducibility and inconsistent associations with post-transplant outcomes despite the frequency with which they are used to determine whether donated kidneys are transplantable.<sup>1–3</sup> These limitations have been attributed to differences in the way these biopsies are performed and interpreted compared with other kidney biopsies used in clinical practice, including the frequent use of wedge rather than core sampling technique, frozen section rather than formalin-fixed paraffin-embedded processing, and interpretation by on-call nonrenal pathologists. However, it is unclear whether it is these factors or sampling variation inherent to any biopsy that is responsible for the poor reproducibility of procurement biopsies.<sup>4,5</sup>

The reproducibility of chronic changes on gold-standard biopsies is not well understood because sequential biopsies are not typically performed in scenarios other than kidney allocation. However, at our center, deceased donor kidney transplant recipients with pretransplant donor-specific antibodies undergo extensive scheduled post-transplant protocol biopsies. These biopsies are performed without the purported limitations of procurement biopsies: all are 18-gauge needle core biopsies, formalin-fixed, paraffin-embedded, and stained with hematoxylin and eosin, periodic acid–Schiff, trichrome, and Jones methenamine silver stains before being interpreted by fellowship-trained renal pathologists using Banff criteria. Of these protocol biopsies, those scheduled for

post-transplant days 7 and 14 fall closest together in time. Studying the concordance of these biopsies, between which the true development of large changes in chronic scarring is less likely, can help identify whether the histologic assessment of chronic changes remains poorly reproducible even under ideal circumstances.

We identified 69 consecutive kidneys transplanted at our center (2015–2019) that underwent day 7 and day 14 protocol biopsies and obtained the reported glomerulosclerosis, interstitial fibrosis/tubular atrophy (IFTA), and chronic vascular disease (arteriosclerosis/arteriolosclerosis). For IFTA values presented in numerical ranges (e.g., 5%–15%), the mean of the range was included. For vascular disease presented in categorical ranges (e.g., mild-to-moderate), the higher value was included. To compare day 7 and day 14 findings, correlation coefficients were calculated for glomerulosclerosis and IFTA, and Cohen's  $\kappa$  was calculated for arteriosclerosis and arteriolosclerosis. Given the potential impact of the presence of severe acute tubular injury or allograft rejection on biopsy reproducibility, we repeated these assessments after stratifying the cohort by the presence of each of these findings on either day 7 or day 14 biopsy. This study was approved by the Institutional Review Board of Columbia University Medical Center.

Among the 69 kidneys included in our study, day 7 biopsies sampled a median 17 glomeruli (interquartile range [IQR]: 14–25, range: 8–47) and day 14 biopsies sampled a mean 18 glomeruli (IQR: 14–23, range: 8–41). Median glomerulosclerosis was 5% (IQR: 0%–13%) at

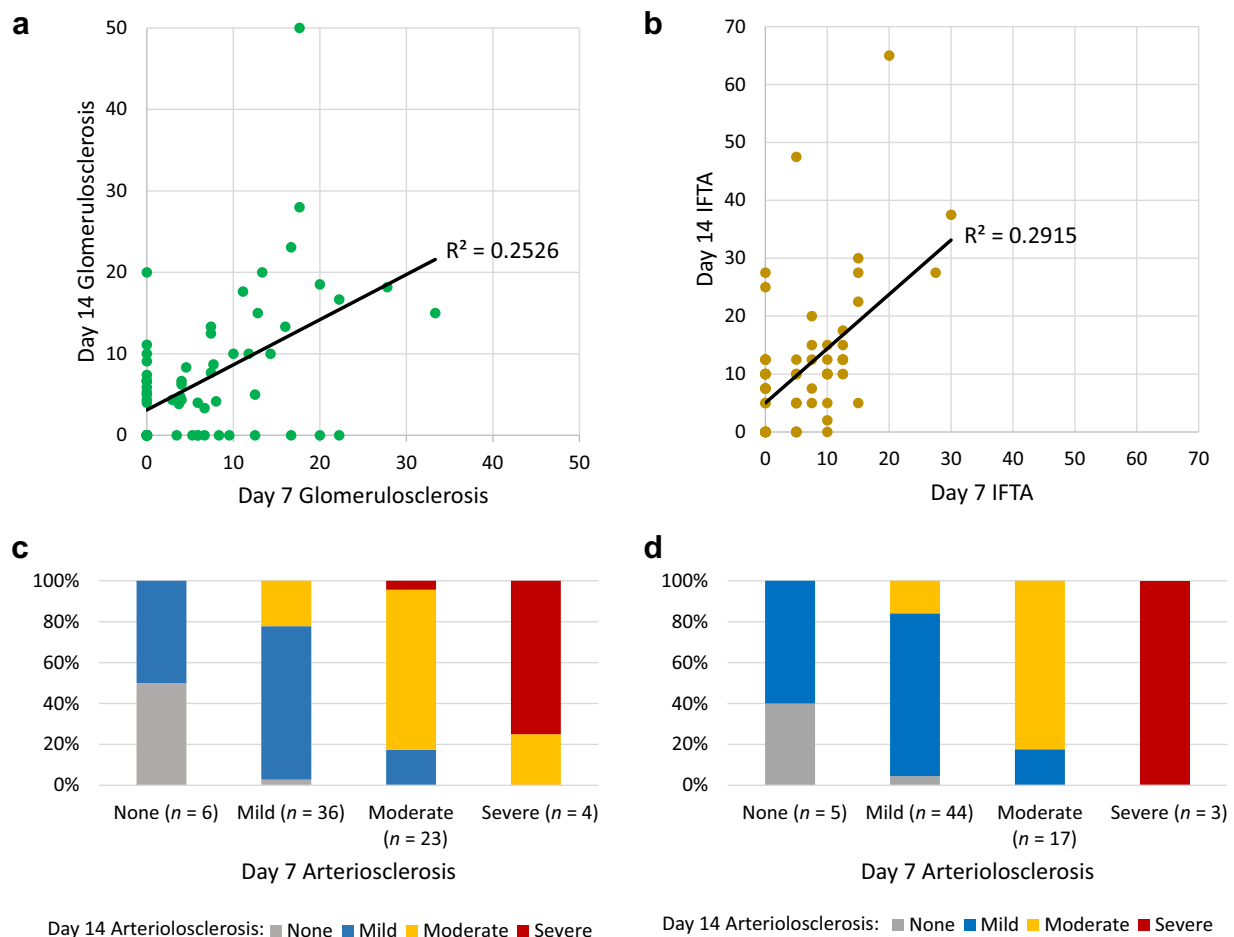
day 7 and 4% (IQR: 0%–11%) at day 14. Median absolute difference between day 7 and day 14 was only 4% (IQR: 0.1%–7%, range: 0%–32%), with moderate correlation between both time points ( $R^2 = 0.25$ ; Figure 1a). Median IFTA was 5% (IQR: 0%–10%) at day 7 and 10% (IQR: 2%–13%) at day 14, with median absolute difference 5% (IQR: 0%–10%, range: 0%–45%), again with moderate correlation ( $R^2 = 0.29$ ; Figure 1b). Agreement between day 7 and day 14 biopsies was best for arteriosclerosis (concordance 74%,  $\kappa = 0.57$ ; Figure 1c) and arteriolosclerosis (concordance 78%,  $\kappa = 0.60$ ; Figure 1d).

Moderate or severe acute tubular injury was identified on 23 (33%) day 7 and 15 (22%) day 14 biopsies. Correlation between biopsy results on day 7 and day 14 was similar when analyzing only the 41 allografts without moderate/severe acute tubular injury on either biopsy (glomerulosclerosis  $R^2 = 0.23$ , arteriosclerosis  $\kappa = 0.52$ , arteriolosclerosis  $\kappa = 0.54$ ). Analysis for IFTA ( $R^2 = 0.05$ ) in this subgroup was limited because no day 7 biopsy and only 2 day 14 biopsies had >15% IFTA. Results were also similar among the remaining 28 allografts with moderate/severe acute tubular injury on at least 1 biopsy

(glomerulosclerosis  $R^2 = 0.33$ , IFTA  $R^2 = 0.34$ , arteriosclerosis  $\kappa = 0.59$ , arteriolosclerosis  $\kappa = 0.63$ ).

Acute antibody-mediated, cellular, or borderline cellular rejection was seen in 15 (23%) day 7 biopsies (13 antibody-mediated, 1 Banff 1A cellular, 1 borderline cellular) and 35 (49%) day 14 biopsies (13 antibody-mediated, 9 Banff 1A–2A, 13 borderline). Among 31 allografts with no rejection on either biopsy, we observed glomerulosclerosis ( $R^2 = 0.32$ , IFTA  $R^2 = 0.46$ ), arteriosclerosis ( $\kappa = 0.50$ ), and arteriolosclerosis ( $\kappa = 0.60$ ) between day 7 and day 14. Among those with rejection on at least 1 biopsy, there was lower correlation between assessments of glomerulosclerosis ( $R^2 = 0.23$ ) and IFTA ( $R^2 = 0.16$ ), but the superior agreement for arteriosclerosis ( $\kappa = 0.62$ ) and arteriolosclerosis ( $\kappa = 0.60$ ) persisted.

Given that kidney biopsies are typically not performed without cause or in close temporal succession, protocol biopsies performed 1 week apart at our center are uniquely suited to assess whether reproducibility of chronic changes on procurement biopsies might be optimized by using improved sampling and processing techniques. These data suggest that sequential allograft



**Figure 1.** Association between day 7 and day 14 assessments of (a) glomerulosclerosis, (b) interstitial fibrosis/tubular atrophy, (c) arteriosclerosis, and (d) arteriolosclerosis in 69 deceased donor kidney transplants that underwent protocol biopsies.

biopsies performed, processed, and interpreted using ideal technique yield modestly differing assessment of chronic changes, with agreement that is particularly strong for vascular disease compared with published reports of sequential procurement biopsy reproducibility. In comparison, we previously reported inferior concordance/K values of 44%/0.12, 71%/0.12, and 59%/0.17 for categorized glomerulosclerosis, IFTA, and vascular disease for sequential deceased donor kidney procurement biopsies.<sup>2</sup> Together, these findings indicate that updating procurement biopsy techniques to best practices and relying on pathologists with kidney pathology expertise can potentially modestly increase the reproducibility—and therefore the utility—of their findings.

Understanding the mechanisms contributing to the superior reproducibility of these biopsy findings compared with procurement biopsy findings is an important step needed to improve the utility of procurement biopsies in the evaluation of deceased donor kidneys. Additional study of which of the differences between the way these biopsies are obtained, prepared, and interpreted contribute to differences in reproducibility may include the use of prospective trials of different biopsy sampling and processing techniques as well as blinded review by different types of pathologists, given mixed results from previous retrospective analyses about the roles of factors such as pathologist training and sampling technique.<sup>4,6,7</sup> In addition, unlike simpler potential changes such as standardizing tissue sampling technique, the implementation of tissue preparation techniques used for other kidney biopsies may not be practical to incorporate in a time-limited setting of organ allocation. In particular, the time required for traditional formalin fixation may limit the feasibility of replacing frozen section processing, although alternative techniques such as rapid formalin fixation or microwave fixation might be considered.<sup>8,9</sup> Similarly, the feasibility of restricting procurement biopsy interpretation to specific types of pathologists may be practically limited by availability and cost.

However, we observed large differences between day 7 and day 14 histology in a minority of cases even though pathologists evaluating day 14 biopsies were aware of the day 7 biopsy and the reported findings. This would be expected to result in a bias toward increased agreement between the reported results—a bias that is also present when sequential procurement biopsies are analyzed. Therefore, it is possible that the true reproducibility of these biopsies is actually lower than what we found. Given these considerations, it seems unlikely that even improving procurement biopsy technique to reflect best practices will yield histologic evaluations during the allocation process that clinicians can fully rely on to make organ utilization

decisions. This finding is also relevant for the care of post-transplant patients, whose providers should note that the degree of fibrosis seen in a clinical biopsy may be subject to a sampling bias and should be interpreted within the larger clinical context. While factors such as inflammation may contribute to these variations, clinicians utilizing biopsy results should continue to take the potential for sampling error into account even when using ideal biopsy technique.

## DISCLOSURE

SM also receives grant funding and personal fees from Angion Biomedica and personal fees from *KI Reports* outside of the submitted work. All the other authors declared no competing interests.

## ACKNOWLEDGMENTS

SAH is supported by the National Center for Advancing Translational Sciences (KL2 TR001874). SM is supported by National Institute of Diabetes and Digestive and Kidney Diseases grants R01-DK114893, R01-MD014161, R01-DK126739, and U01-DK116066.

## REFERENCES

1. Carpenter D, Husain SA, Brennan C, et al. Procurement biopsies in the evaluation of deceased donor kidneys. *Clin J Am Soc Nephrol*. 2018;13:1876–1885. <https://doi.org/10.2215/CJN.04150418>
2. Husain SA, King KL, Batal I, et al. Reproducibility of deceased donor kidney procurement biopsies. *Clin J Am Soc Nephrol*. 2020;15:257–264. <https://doi.org/10.2215/CJN.09170819>
3. Mohan S, Chiles MC, Patzer RE, et al. Factors leading to the discard of deceased donor kidneys in the United States. *Kidney Int*. 2018;94:187–198. <https://doi.org/10.1016/j.kint.2018.02.016>
4. Husain SA, Shah V, Alvarado Verduzco H, et al. Impact of deceased donor kidney procurement biopsy technique on histologic accuracy. *Kidney Int Rep*. 2020;5:1906–1913. <https://doi.org/10.1016/j.ekir.2020.08.004>
5. Reese PP, Aubert O, Naesens M, et al. Assessment of the utility of kidney histology as a basis for discarding organs in the United States: a comparison of international transplant practices and outcomes. *J Am Soc Nephrol*. 2021;32:397–409. <https://doi.org/10.1681/ASN.2020040464>
6. Azancot MA, Moreso F, Salcedo M, et al. The reproducibility and predictive value on outcome of renal biopsies from expanded criteria donors. *Kidney Int*. 2014;85:1161–1168. <https://doi.org/10.1038/ki.2013.461>
7. Liapis H, Gaut JP, Klein C, et al. Banff histopathological consensus criteria for preimplantation kidney biopsies. *Am J Transplant*. 2017;17:140–150. <https://doi.org/10.1111/ajt.13929>
8. Mac-Moune Lai F, Lai KN, Chew EC, Lee JC. Microwave fixation in diagnostic renal pathology. *Pathology*. 1987;19:17–21. <https://doi.org/10.3109/00313028709065129>
9. Chafin D, Theiss A, Roberts E, Borlee G, Otter M, Baird GS. Rapid two-temperature formalin fixation. *PLoS One*. 2013;8:e54138. <https://doi.org/10.1371/journal.pone.0054138>