Kidney Transplant Rejection Clusters and Graft Outcomes: Revisiting Banff in the Era of "Big Data"

George Vasquez-Rios[1] and Madhav C. Menon[1,2]
[1]Division of Nephrology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, New York
[2]Division of Nephrology, Yale University School of Medicine, New Haven, Connecticut

JASN 32: ●●●–●●●, 2021.
doi: https://doi.org/10.1681/ASN.2021030348

The Banff Classification of Allograft Pathology systematically categorizes histologic injury on the basis of acute and chronic renal compartment lesions. The Banff schema was developed by an iterative process involving expert consensus, mainly incorporating data from studies that mapped intercorrelated individual lesion scores to known diagnoses, such as T cell–mediated rejection or antibody-mediated rejection (ABMR).[1] Although the Banff classification has served as a major advance in the management of allograft recipients, it exhibits some intrinsic limitations. Current Banff diagnoses are composites of categoric lesion scores associated with a diagnosis from a histologic standpoint alone. Underlying pathogenetic mechanisms are woven into existing Banff diagnoses only in a few instances, such as the inclusion of anti-HLA donor-specific antibody (HLA-DSA) assays and/or C4d staining for ABMR, or the incorporation of SV40 staining for polyomavirus nephropathy. Moreover, demonstrable prognostic heterogeneity exists within the same Banff diagnostic group, which is frequently identified in newer data.[2] Sequential Banff classification updates[3] have therefore aimed to address these limitations by incorporating unbiased or multidimensional data.[4] Nonetheless, this remains an ongoing challenge.

In this issue of *JASN*, Vaulet *et al.*[5] used a semi-supervised clustering approach to retrospectively evaluate a large dataset (3622 biopsies from 949 patients). The authors used individual Banff acute lesion scores coupled with death-censored graft survival data to identify diagnostic clusters with prognostic relevance among these biopsies with histologic diagnoses of rejection. Acute lesion scores were incorporated in the modeling in a weighted manner on the basis of their individual associations with graft survival. Information regarding the presence or absence of DSA at the time of biopsy was also utilized. In addition, the authors expertly identified and minimized biases arising from the inclusion of both protocol and indication biopsies and multiple biopsies within the same patient (with varying lesion scores). Cluster stability was assessed and confirmed in a majority of biopsies.

In this manner, they identified six novel clusters with an acceptable level of diagnostic accuracy compared with the original Banff diagnoses (adjusted Rand index, 0.48). Clusters 1–3 were not associated with DSA, whereas clusters 4–6 were DSA positive. Cluster 1 (the "no rejection" cluster) had limited inflammation and was associated with a 10-year graft survival of 54.6%. Cluster 2 represented moderate to severe glomerulitis, with limited tubulointerstitial inflammation, whereas cluster 3 was characterized by moderate to severe degrees of tubulointerstitial inflammation (resembling T cell–mediated rejection). Compared with cluster 1, clusters 2 and 3 associated with poor graft outcomes (with 10-year graft survival of 33.3% and 39.8%, respectively). Among DSA-positive clusters, cluster 4 exhibited C4d activity and minimal inflammation, but associated with a markedly low graft survival rates (28.6%) when compared with cluster 1, despite sharing similar Banff scores. Although biopsies in cluster 5 had high glomerulitis g scores, reflective of the predominant microvascular inflammation in ABMR, they also had lower interstitial inflammation *i* and tubulitis *t* scores (and were thus categorized as "mixed borderline rejection"). Biopsies in cluster 6 had higher *t* and *i* scores, representing actual mixed rejection. Interestingly, both cluster 5 and cluster 6 had similar 10-year graft survival rates (6.1% and 6.2%, respectively), despite variable *i* and *t* scores. The association between these novel clusters and graft loss was validated in an external cohort comprising 5191 biopsies, which exhibited an adjusted Rand index of 0.35 (versus 0.48 in the training set). The lower agreement with Banff scores in the validation set as compared to the training set may have resulted from intrinsic differences between the cohorts, including a significantly higher proportion of patients allocated to cluster 4 as a result of a higher C4d prevalence in the validation set (26% versus 8.7%, $P<0.001$). Nevertheless, this study's clinical applicability could be synthesized as shown in Figure 1 (adapted from Supplemental Figure 4 in Vaulet *et al.*[5]), demonstrating the potential for these results to be incorporated into routine patient care.

Although Vaulet *et al.* provide a novel and comprehensive analysis, there are important considerations that could affect the interpretation of the study results. First, only those who completed a 5-year follow-up were included, adding a potential selection bias by missing patients with the highest all-cause

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

Correspondence: Dr. Madhav C Menon, Yale University School of Medicine, 300 Cedar St., New Haven CT 06519. Email: madhav.menon@yale.edu

Copyright © 2021 by the American Society of Nephrology

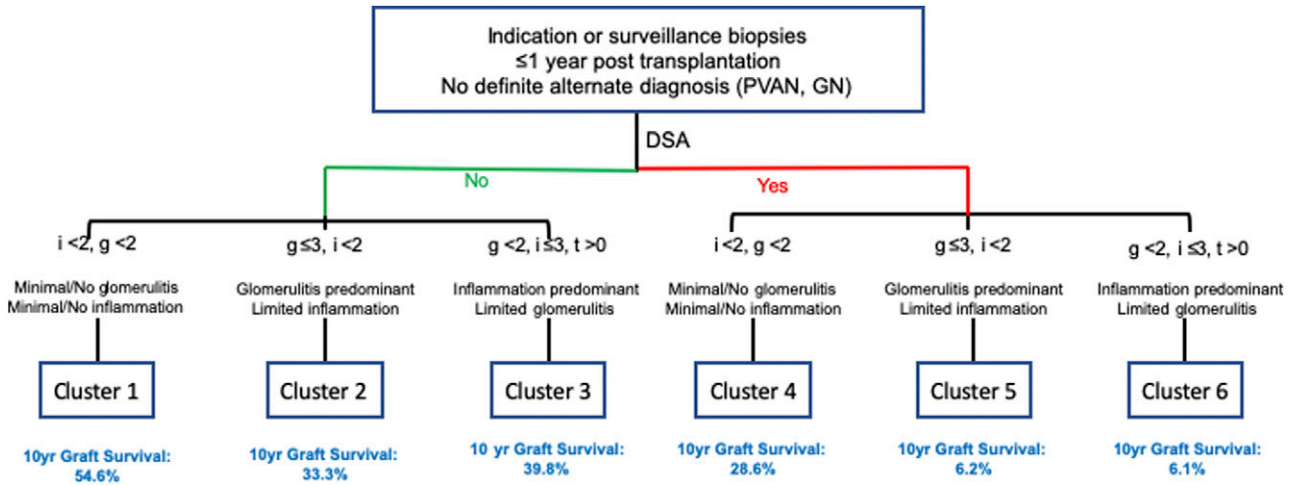

Figure 1. A simplified schema depicting the potential clinical utility of clustering approach (adapted from Supplemental Figure 4 in reference 5). Analyses of year-1 biopsies without any definite nonrejection diagnosis would start with DSA assessment. DSA-negative biopsies could be subdivided as shown into "no rejection," predominant glomerulitis, or predominant tubulo-interstitial inflammation clusters (clusters 1, 2, and 3, respectively). Similarly, DSA-positive biopsies could be subdivided as shown into quiescent, glomerulitis predominant "mixed-borderlines," or inflammation predominant (clusters 4, 5, and 6, respectively). The prognostic relevance of these clusters from Vaulet *et al.* are shown as 10-year graft survival rates. i, inflammation; g, glomerulitis; t, tubulitis; PVAN, Polyoma virus-associated nephropathy.

graft loss rates after transplant. Second, most biopsies included in the dataset were obtained within the first year (83.3%), and it is unknown whether cluster designation may be affected by late-rejection biopsies with late ABMR or predominantly interstitial fibrosis and tubular atrophy. Although the authors focused on acute lesions only, future clustering approaches must include data from chronic lesion scores, because these are reported[6] to associate with graft loss. In this regard, despite differing in inflammation, clusters 5 and 6 had similar graft survival rates, a finding that may point to a role for unmapped chronic scores. In the absence of additional pathogenetic differences between clusters 5 and 6, treatment strategies for both these clusters would be similar in most centers, limiting the utility of these two clusters in particular.

In addition, diagnostic and prognostic heterogeneity needs to be considered within the glomerulitis clusters as there was no evaluation of non–HLA-DSA. Similar heterogeneity may exist in the "no-rejection" cluster 1 because the graft loss rate was higher than nationally reported estimates in the United States.[7] This represents a downside of clustering approaches, which may oversimplify the heterogeneity of many characteristics into a limited number of profiles.[8] It also remains necessary to evaluate in greater depth underlying clinical risk factors within each cluster. For example, although cluster 1 and cluster 4 were histologically similar and differed only in the presence of HLA-DSA, 10-year graft survival rates were markedly worse in the latter (54.6% versus 28.6%). We conjecture that cluster 4 may be capturing an unmeasured parameter within patients, such as nonadherence, which has been associated with both DSA and graft loss.[9] Finally, as the authors caution, this approach should not replace comprehensive evaluation of individual biopsies by transplant physicians and pathologists.

Despite such limitations, this novel dataset from an expert group of investigators highlights the advantage of using semisupervised clustering that incorporates biopsy scores as high-dimensional continuous variables, guided by nonbiopsy-related information (in this case graft survival). The findings show this approach has the ability to relay meaningful clinical data and reduce noninformative factors from preexisting diagnostic clusters. Such innovative tools, enabled by machine learning, could also offer guidance for patients with overlapping histologic patterns on allograft biopsy, helping to reclassify them. In future studies of patients with serial biopsies, it would be particularly interesting to assess the prognostic relevance of dynamic clustering trends, namely, cluster reclassification within the same patient, to further assertain the utility of this approach. *Post hoc* analysis of previously completed trials using similar "big data" approaches could reveal novel means to stratify participants and facilitate cluster-based targeted therapeutics when planning interventional trials in transplantation.

DISCLOSURES

G. Vasquez-Rios has nothing to disclose. M. Menon reports having an ownership interest in Renalytix AI; reports being a scientific advisor or member of *JASN* Editorial board as Editorial fellow, *Journal of Clinical Medicine* Editorial board, and *Clinical Transplantation* Associate Editor.

FUNDING

M.C. Menon received funding from the National Institutes of Health (R01DK122164).

ACKNOWLEDGMENTS

The content of this article reflects the personal experience and views of the author(s) and should not be considered medical advice or recommendations. The content does not reflect the views or opinions of the American Society of Nephrology (ASN) or *JASN*. Responsibility for the information and views expressed herein lies entirely with the author(s).

REFERENCES

1. Solez K, Axelsen RA, Benediktsson H, Burdick JF, Cohen AH, Colvin RB, et al.: International standardization of criteria for the histologic diagnosis of renal allograft rejection: The Banff working classification of kidney transplant pathology. *Kidney Int* 44: 411–422, 1993
2. Loupy A, Aubert O, Orandi BJ, Naesens M, Bouatou Y, Raynaud M, et al.: Prediction system for risk of allograft loss in patients receiving kidney transplants: International derivation and validation study. *BMJ* 366: l4923, 2019
3. Loupy A, Haas M, Roufosse C, Naesens M, Adam B, Afrouzian M, et al.: The Banff 2019 Kidney Meeting Report (I): Updates on and clarification of criteria for T cell- and antibody-mediated rejection. *Am J Transplant* 20: 2318–2331, 2020
4. Mengel M, Loupy A, Haas M, Roufosse C, Naesens M, Akalin E, et al.: Banff 2019 Meeting Report: Molecular diagnostics in solid organ transplantation-Consensus for the Banff Human Organ Transplant (B-HOT) gene panel and open source multicenter validation. *Am J Transplant* 20: 2305–2317, 2020
5. Vaulet T, Divard G, Thaunat O, Lerut E, Senev A, Aubert O, et al.: Data-driven derivation and validation of novel phenotypes for acute kidney transplant rejection using semi-supervised clustering [published online ahead of print March 9, 2021]. *J Am Soc Nephrol* 32: XXX–XXX, 2021
6. O'Connell PJ, Zhang W, Menon MC, Yi Z, Schröppel B, Gallon L, et al.: Biopsy transcriptome expression profiling to identify kidney transplants at risk of chronic injury: A multicentre, prospective study. *Lancet* 388: 983–993, 2016
7. Hart A, Smith JM, Skeans MA, Gustafson SK, Wilk AR, Castro S, et al.: OPTN/SRTR 2018 Annual data report: Kidney. *Am J Transplant* 20: 20–130, 2020
8. Bair E: Semi-supervised clustering methods. *Wiley Interdiscip Rev Comput Stat* 5: 349–361, 2013
9. Wiebe C, Gareau AJ, Pochinco D, Gibson IW, Ho J, Birk PE, et al.: Evaluation of C1q status and titer of de novo donor-specific antibodies as predictors of allograft survival. *Am J Transplant* 17: 703–711, 2017

See related article, "Data-driven Derivation and Validation of Novel Phenotypes for Acute Kidney Transplant Rejection using Semi-supervised Clustering" on pages xxx–xxx.