Published in final edited form as:
*Am J Transplant*. 2012 October ; 12(10): 2573–2574. doi:10.1111/j.1600-6143.2012.04250.x.

# Is the Time Ripe for Genomic Diagnosis and Prediction of Rejection?

**J. S. Bromberg**[a,*] and **D. Iklé**[b]

[a]Department of Surgery, University of Maryland, Baltimore, MD

[b]Rho Federal Systems Division, Chapel Hill, NC

In this issue of the *American Journal of Transplantation*, Li et al. describe a peripheral blood signature for acute rejection in a primarily pediatric population of renal transplant recipients (1). Since the first such finding published in AJT in 2004 (2), we still have not yet taken such a biomarker from the lab into clinical practice. Therefore, we discuss here some general challenges that we still face in advancing the clinical practice of transplantation, and specific challenges of the current study, through the use of noninvasive biomarkers to diagnose acute rejection in order to improve long-term graft survival.

First, we are faced with a choice of an imperfect "gold standard" with inherent measurement issues, whether it is clinical rejection, expressed as a rise in serum creatinine levels beyond some cutoff, or biopsy-proven acute rejection, with its problems of inter-rater reliability and frequent "borderline" findings. In addition, if we are to rely on biopsy-proven acute rejection as the standard, do we accept potentially less reliability with pathologists at clinical multiple centers, or require that biopsies be read centrally by multiple pathologists and require adjudication of differences? The latter may provide greater inter-rater reliability, but the former may better reflect best clinical practice. As imperfect as biopsies may be, they may still be the best we can do at this time to clearly define an endpoint that can be successfully modeled to produce a biomarker with acceptable measurement accuracy. In Li et al., there are the additional clinical limitations that the majority of experimental material is from children, young adults, and Caucasians. Thus, the generalizability of this work to other patient groups is not known.

While many recent studies consider the biomarker to be a noninvasive substitute for a biopsy, the more clinically relevant objective may be the prediction of future rejection that would permit the customization of maintenance therapy to the individual patient before rejection occurs. However, the study designs and statistical methods suitable to develop such predictive models are substantially more complex than those suitable for the development of biomarkers that diagnose concurrent rejection. As a result, we should probably focus on solving the concurrent diagnosis problem first, rather than attempting to take on the more difficult future prediction problem, which is not necessarily solved by simply observing how the predictions from a diagnostic model behave over time in relation to observed rejection events. In this regard, Li et al. is simply a correlation between rejection diagnosed by the pathology gold standard and rejection diagnosed by a proposed new standard of peripheral blood biomarkers. Thus, if successful, the new test may be useful for diagnosis of rejection, but the data do not permit conclusions about the prediction of rejection. Both diagnosis and



*Corresponding author: Jonathan S. Bromberg, jbromberg@smail.umaryland.edu.

**Disclosure**

The authors of this manuscript have no conflicts of interest to disclose as described by the *American Journal of Transplantation.*

prediction will require not only a far more robust clinical and statistical design, but also the specific biological analysis of the quantitative and statistical relationship between an allograft biopsy finding and a peripheral blood measurement, and how that correlation changes before versus during rejection.

Once the study objectives are well-defined, the choice of an appropriate experimental design with adequate sample size that is implemented prospectively is critical to the successful generalization of the results to larger populations of patients. Unfortunately, it has been far more common to develop and test statistical models on samples of convenience, rather than on those that have been well-defined in advance to meet specified eligibility criteria, that have assessments made at well-defined intervals, and that have been sized according to careful power and sample size considerations. It is equally important to include some method of validation of the proposed statistical model and its associated discrimination and calibration metrics, whether it is through external validation in an independent sample or internal validation using established methods for cross-validation in the same sample in which the model is developed (e.g. bootstrap resampling). These design limitations are acknowledged in Li et al., which relied more on the strictures of patient and sample availability than on prospective design and implementation. While the authors are to be lauded for their immense efforts and success in bringing together the samples and data for this study, validation of their hypotheses will now require adherence to more defined prospective criteria.

Of equal importance is the development of an appropriate statistical analysis plan that is tailored to the study objectives and the experimental design and includes a complete assessment of the statistical properties of the proposed classifier. The extensive statistical literature on these subjects and the software available to implement them are readily available, although they do require substantial statistical expertise to implement them appropriately. Far too often, either the analysis plan is simply inadequate or the statistical methods are so poorly described that it is impossible for the sophisticated reader to understand what was done. We suspect that failure to meet this challenge, in particular, may explain why many published biomarkers for rejection have not made it into practice. In Li et al., it is sometimes difficult to determine exactly how the 5 genes specifically included in the signature were identified out of all those available for selection.

In conclusion, the present study by Li et al. is a landmark in furthering our ability to diagnose acute rejection in renal allograft recipients using peripheral blood biomarkers. Many challenges remain to validate the signature prospectively in many other patient populations and clinical settings, and to show that it produces clinically meaningful data that really influence patient care. Whether the signature will supplant biopsy or be added to biopsy data remains to be seen, and whether the signature will be useful for predicting rejection will require a far more robust prospective trial design.

## References

1. Li L, Khatri P, Sigdel TK, et al. A five-gene peripheral blood diagnostic test for acute rejection in renal transplantation. Am J Transplant. 2012; 12:2710–2718. [PubMed: 23009139]
2. Flechner SM, Kurian SM, Head SR, et al. Acute kidney transplant rejection and tissue injury revealed in gene profiling signatures of biopsies and peripheral blood lymphocytes. Am J Transplant. 2004; 4:1475–1489. [PubMed: 15307835]