Received: 15 May 2022 | Revised: 31 May 2022 | Accepted: 4 June 2022

DOI: 10.1111/ctr.14747

Clinical TRANSPLANTATION
The Journal of Clinical and Translational Research
WILEY

OPINION

The MMDx® diagnostic system: A critical re-appraisal of its knowledge gaps and a call for rigorous validation studies

Parmjeet Randhawa

Division of Transplantation Pathology, Department of Pathology, The Thomas E Starzl Transplantation Institute, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Correspondence
Parmjeet Randhawa, Division of Transplantation Pathology, Department of Pathology, The Thomas E Starzl Transplantation Institute, University of Pittsburgh, E737 UPMC-Montefiore Hospital, 3459 Fifth Ave, Pittsburgh, PA 15213, USA.
Email: randhawapa@upmc.edu

Social Media: 200 character tweet: The molecular microscope has great potential to enhance the care of transplant patients. However, at this time it is not sufficiently validated to be included in standard of care recommendations.

Abstract

Transcriptomics generates pathogenetic insights not obtainable by histology, but translation of these insights into diagnostic tests is not a trivial task. This opinion-piece critically appraises declarative MMDx statements, such as the infallibility of machine learning algorithms, measurements of gene expression with >99% precision, and "unambiguous reclassifications" of contentious biopsies such as those with borderline change, polyomavirus nephropathy, chronic active T-cell or mixed rejection, isolated intimal arteritis, and renal medullary pathology.

It is shown that molecular diagnoses that do not agree with histology cannot be attributed primarily to pathology reading errors. Neither can all molecular calls derived from arbitrary binary thresholds be automatically accepted as the ground truth. Important other sources of discrepancies between clinico-pathologic and molecular calls include: (a) organ being studied, (b) disease definition, (c) clinical histologic, and gene expression heterogeneity within the same diagnostic label, (d) size and composition of comparator groups, (e) molecular noise, (f) variability in output of different machine learning algorithms, and (g) the nonavailability of a molecular classifier for chronic active TCMR. Carefully designed clinical trials are needed to determine which of the proposed indications of MMDx provide incremental value over existing standard of care protocols.

KEYWORDS
gene expression, microscope, molecular, pathology, transcriptomics, transplantation

1 | INTRODUCTION

Histopathology is the primary tissue-based modality for investigation of graft dysfunction in all solid organs. It is unable to recognize the earliest stages of rejection and nonimmune injury at the molecular level. There is also a currently unmet need for tests that could provide quantitative measures of inflammation and cell damage. The Molecular Microscope (MMDx) offers potential solutions to these problems. It is also a powerful tool for clarifying disease pathogenesis and potentially recognizing new therapeutic targets. As is true for other currently available molecular tests, practicing physicians seeking to understand MMDX capabilities need to be able to make a distinction between commercial hype and rigorous science. An opinion piece published in this journal last year highlighted the limitations of this promising technology.[1] McCloskey et al. have now responded to this article expressing disagreement with several of the points made in that article.[2] I welcome this opportunity to respond to the issues raised, and re-emphasize the substantial validation studies that are needed before MMDx can be regarded as standard of care. I also comment on some other areas relevant to molecular diagnosis, and point out

TABLE 1 Considerations that support continued confidence in utilization of pathology readings in the era of molecular medicine

1.	Molecular heterogeneity within the same diagnosis is now recognized as a key factor in misclassification of samples.
2.	Pathology has been in use for > 4 decades during which time it has guided substantial improvements in graft outcome.
3.	It remains the basis of international standard of care TCMR recommendations such as KDIGO.
4.	Turnaround time for pathology report and in-person or Zoom video conference as little as 6 h, allowing treatment on the same day as clinic visit
5.	Pathology of protocol biopsies has substantial associated outcome data on subclinical TCMR/ABMR.
6.	Diagnostic value of pathology lesions is supported by extensive body of human research and animal models.
7.	Histology provided the basis for pathogenesis-based transcripts (PBTs) used in MMDx
8.	Pathology interpretations go beyond the basic MMDx categorization of biopsies into just three categories, namely, Not Rejection, ABMR, and TCMR[a]
9.	Histology can provide scores for i-IFTA, t-IFTA and recognize chronic active TCMR which is typically interpreted as "No TCMR" by MMDX.
10.	Pathology breaks kdown ABMR into C4d -ve and C4d +ve variants which may not necessarily be equally responsive to anti-complement therapy
11.	Light microscopy identifies non-renal tissue, necrosis, and dense scars which can potentially result in false -ve MMDX calls.
12.	Morphology provides a global structural context to pathologic changes and localizes disease to specific tissues, cell types and organelles
13.	Pathology will be essential for follow up studies needed to understand the meaning of MMDX-histology discrepancies.[b]

[a]A biopsy can diagnose infectious pathology, calcineurin inhibitor toxicity, diabetic nephropathy, focal segmental glomerulosclerosis, glomerulonephritis, non-immune interstitial nephritis, vascular disease, malignancy, metabolic disorders, paraprotein deposits, and amyloid.
[b]This includes ~50% of biopsies with histologic TCMR that are not confirmed by MMDX, and biopsies with high molecular ABMR scores in the absence of histologic rejection or clinical graft dysfunction.

the continued importance of pathology readings in the era of Omics technologies and machine learning (Table 1).

2 | SAMPLING PROBLEMS IN MMDX

The MMDx assay was initially standardized on a separate core of renal allograft tissue. Due to logistical issues in the setting of routine clinical work, analyses are now performed on a small 3 mm tissue fragment taken from one end of the needle biopsy taken for diagnostic purpose. The effect of this reduction in tissue analyzed on potentially missing important lesions is never discussed. In a weekly kidney biopsy conference conducted by me, I see noticeable regional lesion variations in 10–20% of all kidney biopsies. Biopsy sampling problems are an order of magnitude higher in heart than in kidney biopsies. It has been estimated that the true negative rates of histologic cardiac rejection corresponding to one, two and three biopsy fragments are 42%, 63% and 79%, respectively.[3] Thus, it is expected that MMDx analysis of insufficient tissue would contribute to discrepancies between histologic and MMDX assessment. The frequency of these discrepancies may also differ according to the size of the core (16 gauge vs. 18 gauge) and the case mix.

It is hard to understand McCloskey et al.'s assertion that gross and microscopic pathology lesions are subject to sampling error, but the molecular representation of those lesions is not. Certainly, molecules would diffuse beyond the epicenter of the pathology, but with substantial fall in concentration. This would affect MMDX performance because molecular classifiers use actual gene counts in developing predictive equations. The potential for erroneous results using small biopsy samples is illustrated in Figures 1–3. Parenthetically, the phenomenon of sampling error is very well appreciated by infectious disease physicians, who diagnose presumptive nephropathy with virus negative biopsies two to three times more commonly than biopsy proven BKV nephropathy.[4] Renal pathologists are also acutely aware that making submicroscopic and subcellular observations without ensuring that the material examined is representative runs the risk of missing the forest for the trees. As an example, electron micrographic observations made at 100 000 magnifications become meaningful only if correlated with changes at the light microscopic 10x–400x magnification level.

3 | PRECISION AND REPRODUCIBILITY OF MMDX

MMDx literature reports >99% precision and/or >99% reproducibility for its ability to express the 3-dimensional relationship of each new biopsy to a phenotyped reference set, and to measure transcriptomics changes in biopsy samples.[5–8] The methodology used to assess the correctness of assignments in 3-dimensional space is not clear. Moreover, the terminology used does not follow conventions that are broadly accepted in laboratory medicine. At times reproducibility seems to refer to the precision of measurement, while on other occasions it seems to allude to the accuracy of the diagnostic class. When used in the sense of precise measurement, there is no description of the coefficient of variation of the assay signals that go into the test interpretation.

I find it surprising that statistical measures of variability of principal component scores, archetype scores, molecular classifier scores, and individual lesion scores that are a prominent part of the MMDx

FIGURE 1 Illustration of sampling issues in kidney pathology at the macroscopic level: Gross photograph of a deceased donor kidney being evaluated for transplantation. The lower pole (right lower corner) has a circumscribed infarct. It stands to reason that a sample taken from the center of the infarct may have no viable RNA to analyse; a second sample next to the infarct would show a gene expression profile corresponding to severe ischemic injury, which would not be as apparent in a third sample taken several cm away from the infarct

report are missing in several landmark publications. This is not a trivial point. For example, extensive experience with quantitative RT-PCR for transplant-associated viruses has shown that 2–5 fold variability is common in the measurement of molecular signals. If there is greater precision in MMDX scores in replicate samples or serial biopsies from patients being treated for TCMR or ABMR, then that data should be shared with the transplant community on the Thermo Fisher website. Indeed, close examination of a publication commonly cited for claiming high MMDX precision actually shows many kidney biopsies with 2–5 decile variations in technical as well as biologic replicates of ABMR biopsies.[9–11]

The MMDX report contains an impressive list of lesion-classifiers that continues to grow. There is no peer reviewed publication attesting to the performance of many of these classifiers, including probability scores for interstitial inflammation, tubulitis, tubular atrophy, peritubular capillaritis, and arteriolar hyalinosis. These caveats represent a significant impediment to the very desirable goal of incorporating these scores into clinical decision making and therapeutic trials. For example, one cannot determine whether TCMR scores of .3 and .5 applied to two different biopsies represent a true difference in disease state, or are within the range of measurement error.

Considering that ~4500 biopsies have now been analyzed by MMDx, the amount of quality assurance data is indeed very sparse. One recent report presents correlations as high as .99 to 1.0 between principal component, archetype and classifier scores based on replicate analyses of just 11 samples.[7] In these replicate analyses, a small fragment was divided into even smaller fragments, and one cannot expect such iterative mincing to capture the variation in pathology seen across the whole biopsy core (Figure 2). The diagnoses of the 11 samples is not known, but in an earlier publications that studied 26 biopsies with ABMR or TCMR, the best reproducibility was seen in samples with relatively little pathology (molecular scores close to the diagnostic MMDx threshold).[9] This is analogous to pathologists being easily to correctly identify near-normal biopsies. On the other hand, biopsies with worse inflammation or cortical versus medullary locations showed several decile variations and did not fit the description of "99% precision." The proportion of such cases with appreciable technical variation in any given dataset would depend on the case-mix, and would be substantial if biopsies shown to have focal pathology were specifically examined at opposite ends.

In summary, the reproducibility of MMDx needs to be further defined by a stress test that includes larger numbers of challenging biopsies that are difficult to interpret with molecular as well as histologic techniques. Importantly, correlation coefficients should not be conflated with the concept of precision. High overall correlations can be obtained even when samples differ several fold in replicate measurements with totally unacceptable coefficients of variation exceeding 80%, as long as an overall trend is maintained (Figure 4). Data on correlation coefficients analysis has value in population based research studies, but does not help the management of individual patients.

4 | INTER-OBSERVER AGREEMENT BETWEEN MMDX EXPERTS

This is another area in MMDx literature that calls for critical comment. Agreement between three MMDx experts closely working together is reportedly as high as 92–94% for prototypical biopsies.[11] It should be noted that this is an overall agreement in a biopsy set wherein more than half the biopsies had no rejection. The agreement rates were much lower in problematic areas. Thus, in the category of probable TCMR (molecular equivalent of borderline change) experts 1 and 2 had a discrepancy rate of 7/18 or 38%. For probable ABMR (suspicious for antibody mediated injury) the discrepancy rate was 53% (27/51 biopsies). These latter agreement rates are comparable to the 50% interpathologist TCMR disagreement noted in a study that is frequently quoted in MMDx presentations as evidence of variability in histology readings.[12] To extend this comparison, a departmental

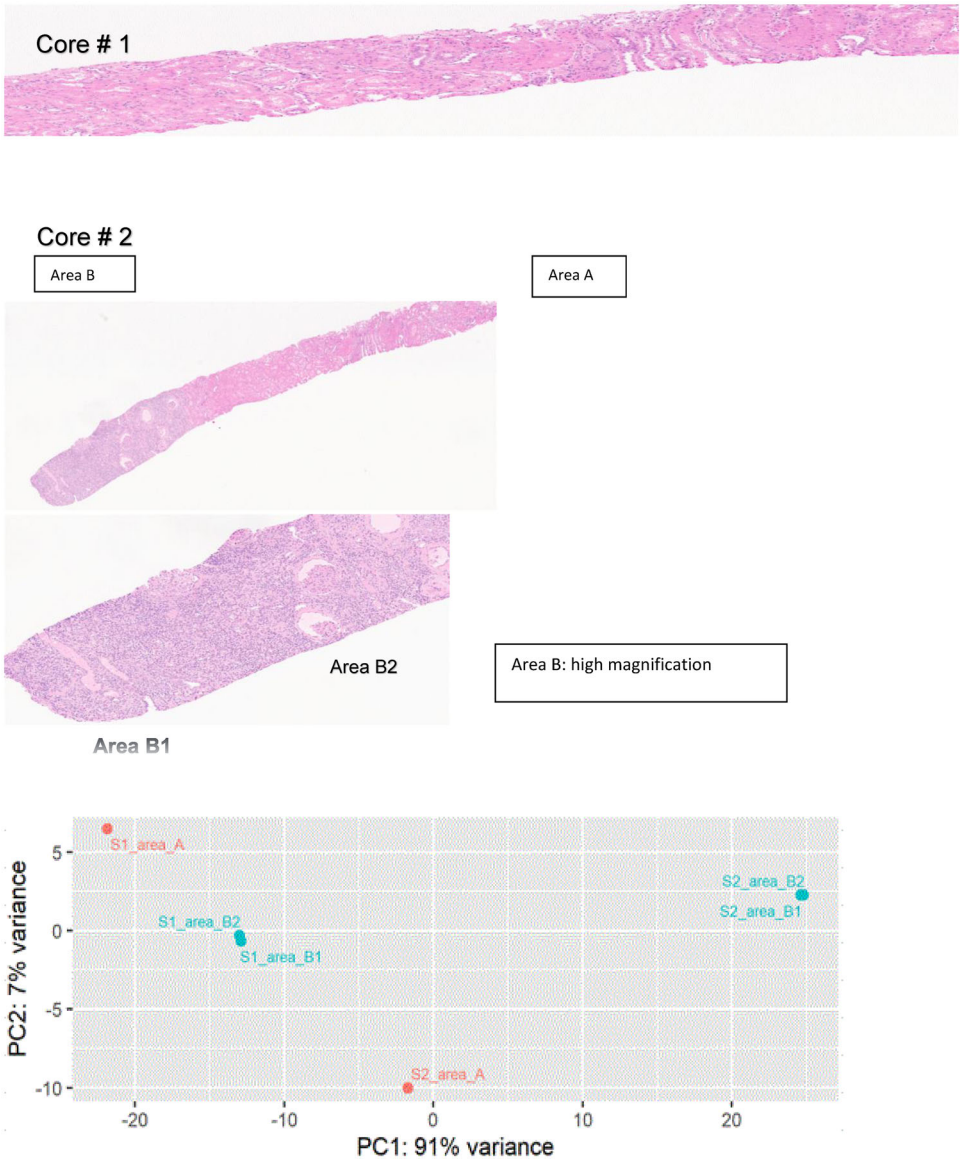

FIGURE 2 Illustration of sampling issues in renal allograft biopsies examined at the light microscopic level: This biopsy sample consisted of two cores, respectively, measuring 1.5 cm (Core #1) and 1.2 cm (Core #2). Core #1 is virtually free of inflammation. Core #2 has severe inflammation in approximately 30% of the core (Area B), while the remaining 70% (Area A) is free of inflammation of and tubulitis. Gene expression analysis by RNAseq using a published technique (PMID 30296518) showed 1198 differentially expressed genes between Area A and Area B of the smaller core. Replicates B1 and B2 within area B were quite concordant, with a correlation coefficient of .99, which created a false sense of reproducibility in the molecular analysis. On clustering analysis (lowest panel), using the CLC Genomics Workbench, areas B1 and B2 within two different samples S1 and S2 clustered together with each other but not with the area A of the same core in the corresponding sample. This is a clear illustration of how 3 mm cores taken for evaluation by the molecular microscope have the potential to label one portion of the biopsy as TCMR and another portion as normal. A correlation coefficient as high as 99% may be obtained between replicate analyses in a small core, and automated molecular sign-outs by different molecular experts may achieve 100% concordance; yet the diagnostic categorization of this biopsy may not be correct for the purposes of clinical management. It is also apparent that molecular scores for inflammation and associated pathology lesions derived from a small fragment can be quite nonrepresentative of the whole biopsy

histology quality assurance program overseen by me achieves ~ 80% interpathologist for Banff grade 1A and 90% for grade 1B TCMR.

It is important to note that even the 92–94% agreement between three MMDx experts working closely together applies only to the final step of analyzing these signals in an optimized bioinformatics pipeline. It does NOT account for the variability inherent in the generation of these molecular signals. Pathologists given an excel table of lesion-scores would achieve a similarly high rate of agreement for canonical cases, if they were to be tested only on the translation of these scores to a Banff grade of rejection. The numerous "upstream" factors that can contribute to signals generated by MMDx are summarized in Figure 5. If all the variables were to be considered, and

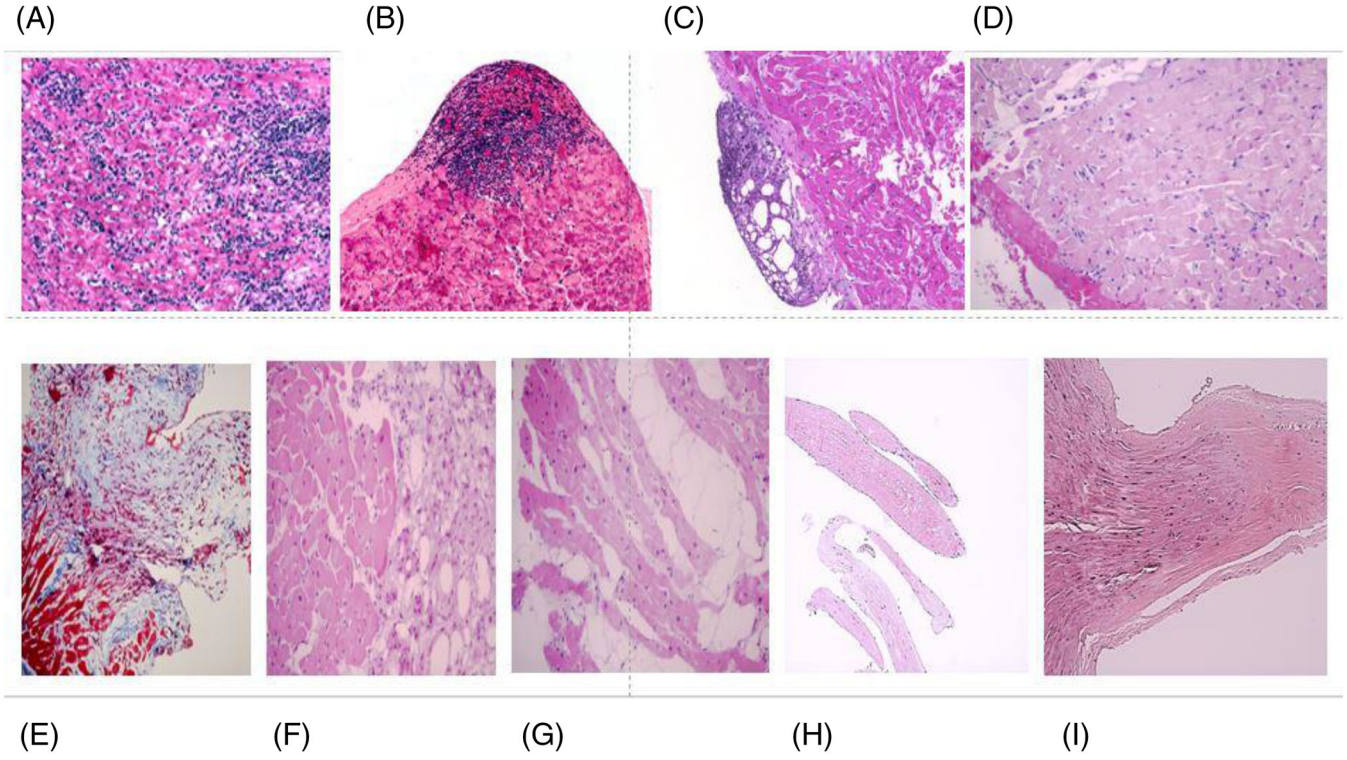

FIGURE 3 Illustration of sampling issues in heart allograft biopsies examined at the light microscopic level: Heart allograft biopsies are typically 1–2 mm in diameter and even more subject to sampling variations than a kidney biopsy. Some landmark molecular studies in peer reviewed literature have been based on examination of a single fragment. Different myocardial fragments in the same biopsy can vary in the severity and localization of inflammation (A-D). Furthermore, endomyocardial biopsies contain nonmyocardial tissue fragments about 20% of the time (E-I). Most often, this is a prior biopsy site (E); less often portions of pericardium (F), intramyocardial fat (G), chordae tendineae (H), and papillary muscle may be included (I). The degree to which biopsy related trauma (E,C) and clinically insignificant inflammation such as Quilty lesions (B) confound molecular MMDX adjudication of rejection and tissue injury is unknown. This is because the tissue taken for MMDX analysis is not first examined for the underlying histopathologic lesions

allowance was made for even a low % of errors at each step, the overall MMDX test performance would be well below 99%. The oncology literature has shown that microarray based assays can result in misclassification rates of 31–49%.[13] Bioinformatics analyses of public transcriptomics data derived from the human allograft kidney generates sample misclassification rates of 27.9–46.9%.[14] Notably, these estimates are in the same range as reported discrepancies between MMDX and histology.

In the light of these considerations, it should be apparent that disagreements between Histology and MMDx cannot be described as being largely due to the known noise of histology assessments. Indeed, when different molecular classifiers are used, the interalgorithmic variability becomes comparable to that seen among pathologists (see section below on use of Ensembles). Important other sources of discrepancies between clinico-pathologic and molecular calls include: (a) disease definition, (b)clinical histologic, and gene expression heterogeneity within the same diagnostic label, (c) size and composition of comparator groups, (d) molecular noise, (e) variability in output of different machine learning algorithms, and (e) the nonavailability of reliable molecular classifiers for some disease such as polyomavirus nephropathy and chronic active TCMR (Sections 1.11, 1.15).

5 | SPECIFICITY OF MMDX DIAGNOSES

It is well known that while some pathology lesions are relatively specific, while others can be produced in more than one disease. Thus, microvascular injury can occur in antibody-mediated rejection, thrombotic microangiopathy and membranoproliferative glomerulonephritis. If we accept the premise that gene expression profiles reflect disease pathogenesis, the same lack of absolute specificity should be assumed for transcriptomics signatures, till otherwise proven. It is currently unknown if the MMDx signature for ABMR will result in false positive diagnoses in biopsies with non-ABMR causes of microvascular inflammation. MMDX treats C4d positive and C4d negative ABMR in the same diagnostic category, but these may have complement and antibody-independent mechanisms such as those mediated by T-cells, macrophages, NK cell recognition of "missing self," and innate immune mechanisms involving paired immunoglobulin-like receptors.[15,16,17,18] Likewise, DSA positive and DSA negative ABMR both get the same molecular diagnostic label, but a formal study is needed to ascertain if both groups of patients show a similar response to plasmapheresis and Rituximab therapy. MMDx reports acknowledge that this system is not suitable for the diagnosis of glomerulonephritis. Whether it

(A)

	Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Standard Deviation	Coefficient of Variation
Data Value #1	0.1	0.5	2.4	2.2	1.1	1.02	0.81
Data Value #2	0.3	0.9	5	6.2	2.5	2.56	0.86
Data Value #3	0.7	1.7	10.2	14.2	5.3	5.72	0.89
Data Value #4	1.1	2.5	15.4	22.2	8.1	8.90	0.90
Data Value #5	1.5	3.3	20.6	30.2	10.9	12.08	0.91
Data Value #6	1.8	3.9	24.5	36.2	13	14.47	0.91
Data Value #7	2.5	5.3	33.6	50.2	17.9	20.04	0.91
Data Value #8	4.1	8.5	54.4	82.2	29.1	32.76	0.92
Data Value #9	4.8	9.9	63.5	96.2	34	38.33	0.92
Data Value #10	5.3	10.9	70	106.2	37.5	42.31	0.92

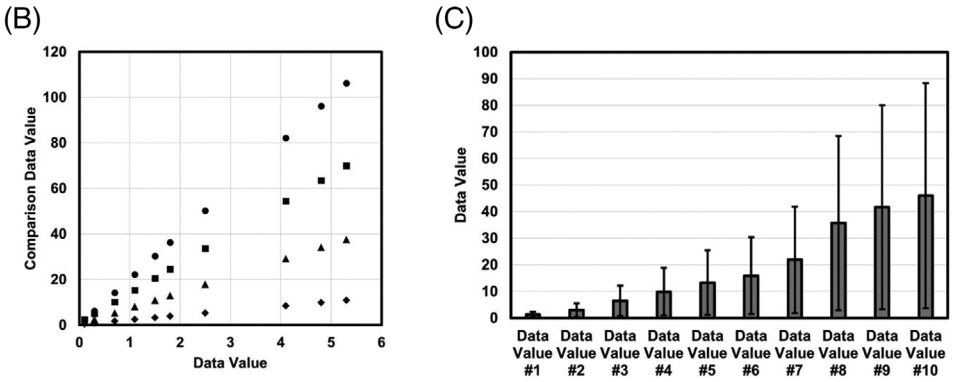

FIGURE 4 Correlation is not equivalent to accuracy and reproducibility. Hypothetical datasets have been created to illustrate this key point. (A) A table containing five perfectly correlated datasets (Pearson's/Spearman's correlation coefficient equal to 1). Each of the five datasets is composed of 10 data values per set. A wide range of standard deviations (1.02 to 42.31) and coefficients of variation (81% to 92%) are listed for each data value. (B) Scatter plot showing Dataset#2 (rhombus), Dataset #3 (square), Dataset #4 (circle), and Dataset #5(triangle) (y-axis) plotted against Dataset #1 (x-axis). C) Bar graph of mean and standard deviation of data values across the five datasets shown in A

can distinguish TCMR from autoimmune interstitial nephritis or drug hypersensitivity reactions that are seen in many of these patients is not known. Thus, further definition of the specificity of MMDx diagnoses is needed. This would require substantial additional microarray profiling of a broad array of glomerular, tubulo-interstitial and vascular kidney diseases that can recur in the renal allograft.

5.1 | Objectivity of MMDx diagnoses

It is commonly asserted that histologic diagnoses are subjective in nature while diagnoses based on gene expression profiling are more objective. This is a valid argument since an eye-balling approach can never equal a precise measurement in the biochemical or molecular biology laboratory. However, this should not blind us to the fact that there are inherent limitations in machine learning algorithms, all the more so when the learning is dependent on diagnoses captured from pathology reports. For example, if one pathologist identifies mixed rejection in 10% of a set of ABMR biopsies, while another pathologist designates 90% of the same biopsies as "mixed rejection," molecular classifiers built from these two sets of pathology reports will have very different performance. Even more importantly, molecular score thresholds used by these algorithms to distinguish between diagnostic categories are not as "objective" as one might think. Indeed, these thresholds are also the result of eye-balling gene expression data plots and assignment of arbitrary cut-offs.[11,19] Currently used thresholds that are used to separate TCMR and ABMR from nonrejection vary widely from .1 to .6. In fact, the thresholds used in MMDX signout have changed over time (like Banff rules!), since the most appropriate threshold can vary with the molecular classifier used, the case mix studied, and the sample size.

It goes without saying that the actual threshold used will markedly affect molecular diagnostic calls. This concept can be illustrated by reference to a recent MMDX analysis of liver biopsies, wherein as few as 26 and as many as 58 biopsies could be categorized as TCMR depending on the cut-off chosen.[20] It is obvious that molecular calls have no inherent meaning until they are correlated with the clinical picture, biopsy findings, response to treatment and ultimate graft outcome.

A second problem to note is that some of the published molecular classifiers suffer from circular reasoning. Thus, a classifier built for TCMR excluded borderline cases during the test development phase,[21] but was subsequently used to proclaim ~70% of borderline (BL) biopsies as 'no molecular rejection.' It is not a surprise that most histologic BL calls were identified as no rejection by this classifier. The more important question is why 30% of BL cases still got that label. The latter number should have been zero, if the classifier had performed as originally trained. Circular reasoning is also apparent in labeling biopsies

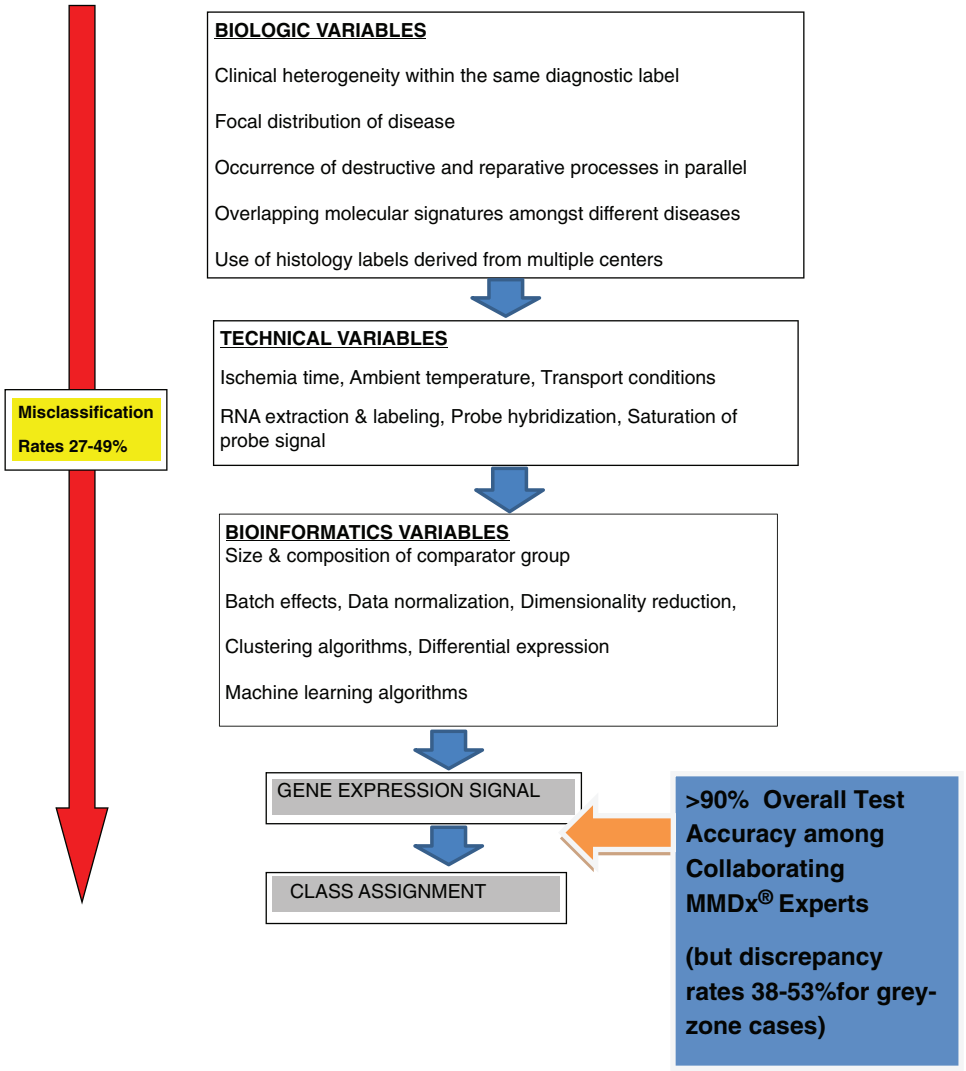

FIGURE 5 Sources of data variability in MMDX Assays. Biologic, technical, and bioinformatics factors all contribute to data variability in DNA microarray technology. Closely collaborating MMDX experts focusing only on the last analytic step (box at lower right) can attain overall concordance> of > 90%, but discrepancy rates as high as 38–53% can be seen for gray zone cases such as probable TCMR (molecular equivalent of borderline change) of biopsies suspicious but not diagnostic of ABMR. By comparison, oncology studies that that take into account the interplay of all the steps in the diagnostic pipeline (vertical red arrow) report interinstitutional sample misclassification rates in the 31–49% range.[13] For renal allograft transcriptomics data available in public databases overall classification error rates vary between 27.9 to 46.9%.[14.] Biopsy samples are subject to biologic and technical but not bioinformatics related factors. Biopsies are also subject to sampling error but with an average kidney biopsy length of 1.5 cm, this would be substantially less than that seen in MMDx analysis, which typically examines ~3 mm of kidney tissue

with isolated v-lesions as 'no-rejection, interpreting i-IFTA as having little TCMR, and in the derivation of the BKV probability classifier, as is noted in subsequent sections. Unsupervised archetype analysis does not eliminate this problem, as it is still dependent on differential gene expression data that is derived from comparator groups that retain similar labeling problems.

A third point to make is that the output of MMDX machine learning classifiers is colored by personal opinions lurking behind a façade of 'objective' molecular diagnoses. Thus, MMDx investigators believe that tubulitis in ABMR is always a secondary phenomenon. Accordingly, it has built classifiers that have attributed ~60% of renal graft loss to ABMR.[22] Other investigators reasonably argue that these are mostly mixed T-cell and antibody rejection, and argue that a proportion of these cases reflect TCMR as the primary injury and antibodies being a secondary (but still important) player. In support of this argument, RNA-seq based machine learning applied to gene expression data in the public domain indicate that 62% of biopsies labeled ABMR satisfy molecular criteria for TCMR.[23]

6 | USE OF ENSEMBLES IN THE MMDX SYSTEM

Machine learning algorithms are an excellent tool for performing repetitive tasks. However, literally, scores of such algorithms are available,

and these can differ significantly in output. Recently, Dr. R.N. Smith at the Harvard Medical School downloaded the Banff Human Organ Transplant (BHOT) Panel genes from archival public microarray data. This panel readily identified renal rejection and nonrejection using in silico statistical analyses. However, different modeling algorithms showed a highly variable pattern of misclassifications per sample in the range of 27.9–46.9%. The error rate could be reduced by increasing the number of clusters analyzed from 6 to 9. However, the additional statistical clusters did not have a readily definable clinical phenotype.[14] These observations clearly show that gene expression heterogeneity within a given diagnosis can make a substantial contribution to sample misclassification.

MMDx has taken a reasonable approach to correct for machine learning model variability by analyzing each biopsy with an ensemble of 12 different algorithms, and using the median for molecular sign outs. However, using an ensemble only creates a new error matrix while retaining per sample variations at a level that is still problematic for clinical use. Thus, the MMDx ensemble produces ABMR probability correlations varying over a wide range from about .4 to .9. It is concerning that the IQR variability and confidence intervals of the ensemble estimate median values generated is typically not included in scientific publications. The choice of the 12 algorithms used is not explained, and if were to use different set of algorithms even lower correlations might ensue. Thus, diagnostic classes defined by ensembles cannot per se be regarded as a new gold standard without further clinical validation in terms of superior therapeutic responses. With respect to the use of median or mean values as a substitute for hard clinical outcomes, I am reminded of the old adage that a 6-foot tall man can drown in a river that is on the average only 5 feet deep!

6.1 | The "Unambiguous Nature" of MMDX diagnoses

The unambiguous nature of MMDx interpretations is stated to be one of its strengths. In reality, this air of authority is created by defining fixed thresholds in continuously variable data, and use of ensemble medians and random forest-based majority vote calls. Moreover, principal component plots show overlapping clusters for different diagnostic classes, and disagreement rates between MMDx experts can be as high as 38–53% for samples that lie in boundary zones.[7] A constantly expanding reference biopsy set is a potential strength of MMDx, but adding more biopsies into the mix with the same diagnostic discrepancies as previous ones cannot be expected to improve test accuracy. Current molecular sign outs include archetype analysis, which is described as being independent of histology lesions. However, the MMDx system retains an important connection to pathology. Principal components, classifier scores and pathogenesis-based transcripts that are used in generating molecular calls are derived from differential gene expression that used histology labels derived from human or experimental investigations. For example, one ABMR classifier chose .2 as a diagnostic threshold in an attempt to achieve 90% specificity for diagnosing ABMR as defined by histology and DSA.[24]

7 | POTENTIAL UTILLTY OF MMDX IN ABMR

It has become clear that MMDx calls ABMR more frequently than pathology, while histologic TCMR is often not confirmed on molecular analysis. This reflects, at least in part, the rules that were used to develop ABMR classifiers. For example, $g > 0$, ptc > 1, and cg > 0 were key elements in classifier development, and we know that these lesions are not absolutely specific for ABMR, which is therefore at risk of being overestimated. In some MMDX publications, it appears that biopsies with even $g = 1$, C4d $= 0$, and asymptomatic DSA were called C4d negative ABMR, and used to define the top 20 probe sets for diagnosis of ABMR.[24] It is also important to keep in mind that there are substantial reproducibility issues between distinguishing histology grade 0 from grade 1 microvascular injury.[25] This uncertainty and inaccuracy in data input will affect the ability of molecular classifiers to predict clinical events, even if these algorithms are based on rigorously applied mathematical principles.

To highlight the more frequent ABMR calls by MMDX, I will first quote one recent publication, which detected elevated ABMR transcripts in DSA positive patients with biopsies not showing histologic rejection.[26] There was an increased risk of 3-year graft failure in these patients. This study is a good illustration of the potential of MMDx to improve patient care in the ABMR arena. On the other hand, in a series of biopsies with isolated intimal arteritis, high ABMR scores were associated with a benign clinical course, as judged by a graft survival that was no worse than a comparator set of biopsies with no vascular inflammation.[27] More studies are needed to better define the circumstances in which a purely molecular diagnosis of ABMR should trigger therapeutic intervention. It will be necessary to determine if some cases resolve spontaneously, and can be spared exposure to expensive and potentially toxic mediations such as plasmapheresis, intravenous immunoglobulins and rituximab.[26]

8 | RELEVANCE OF MMDX FOR TCMR

Whereas MMDx frequently diagnoses ABMR, it fails to confirm histologic TCMR in about one half of renal transplant biopsies. MMDx presentations frequently mention how participating INTERCOMEX physicians agree somewhat more with molecular over histology calls in discrepant cases. The difference was slight (87% vs. 80%), based on subjective opinion, possibly within the realm of statistical variation, and expected from investigators who had invested so much intellectual energy into this study. I am certainly aware of large medical centers outside of this team, who dismiss these discrepancies, and continue treating biopsy proven histologic TCMR according to national guidelines formulated by KDIGO and other professional organizations. These professionals believe that MMDx has not produced sufficient evidence to implement a departure from TCMR standard of care, which would require that we stop treating 50% of our histologic TCMRs according to a steroid-based treatment protocol supported by 4–5 decades of clinical experience.

The problem of unconfirmed histologic TCMR is worse in heart transplantation wherein 25/28 of ISHLT 1990 grade 1A, 1B and grade 2 (ISHLT 2005 grade 1R) biopsies analyzed by us were not recognized by MMDx.[28] Parenthetically, not all grade 1R rejection in the heart is inconsequential: persistent 1B and ISHLT1990 grade 2 biopsies should get consideration of treatment if there is persistent graft dysfunction without another explanation. This latter view point acknowledges the substantial sampling issues that are seen in endomyocardial biopsies, Figure 3. One of four biopsies with grade 3A TCMR was also labeled as "no rejection" by MMDx, while a second biopsy was said to have ABMR but not TCMR.

The reasons why some definitive, histologic calls of TCMR are MMDX-negative is not clear. First, there is room for histologic misinterpretation in the setting of infectious interstitial nephritis or post-transplant lymphoproliferative disease. However, the frequency of these diseases does not explain the magnitude of the discrepancy. A second very relevant issue is the threshold chosen by MMDx to discriminate rejection from no rejection. It that threshold were to be lowered, the agreement rate between histology and MMDX would obviously increase. Thirdly, the focal nature of TCMR is a consideration as discussed earlier (Section 1.2, Figures 2 and 3). Consistent with the importance of irregular lesion distribution, histology-positive and MMDX-negative are biopsies are much less common for ABMR, which is often a more diffuse pathology than TCMR. A final potential explanation for fewer MMDX calls of TCMR may be that it is a relative decrease secondary to more biopsies getting labeled as ABMR (Section 1.9). The clinical implications of the apparent false negative MMDx calls of TCMR remain to be determined. Short-term follow up of small series of patients from centers that regularly use MMDX has not resolved the issue.

9 | CHRONIC ACTIVE TCMR

A chronic active phase of TCMR with progressive interstitial fibrosis and chronic allograft arteriopathy has been suspected for a long time, but was formally recognized only in the Banff 2017 Schema of renal allograft pathology.[29,30] The diagnosis of chronic active TCMR requires a Total-i score > 1, i-IFTA score > 1 and t score > 1. MMDx cannot currently diagnose this entity, as the corresponding molecular classifiers have not yet been developed. It typically assigns these biopsies to the No Rejection or ABMR category.[31] This mis-assignment has significant implications for the performance of both ABMR and TCMR classifiers, and increases the discrepancy rates between MMDX and histology calls.

10 | CAN MMDX ASSESS MIXED ABMR-TCMR?

The idea that archetype analysis or molecular classifier scores could provide separate estimates of the TCMR and ABMR components of alloimmune injury in a biopsy is quite attractive. However, some MMDx estimates of mixed rejection are as low as 5%,[24] while histology based estimates are as high as 63%, or even 96% if borderline cases are excluded.[32] In Pittsburgh, ~90% of cases with ABMR have T-cell infiltrates, tubulitis, tubular apoptosis and tubular cytoplasmic injury that we believe should be targeted by a TCMR treatment regimen. One way for resolving these two disparate viewpoints would be for INTERCOMEX study investigators (ClinicalTrials.gov NCT01299268) to accrue data showing that only the 5% of cases that are diagnosed by MMDx need anti-T-cell therapy. I consider that an unlikely scenario and postulate that the truth probably lies somewhere between 5% and 90%.

11 | APPLICATIONS OF MMDX TO THE BORDERLINE CATEGORY

In keeping with the higher threshold set by MMDx for TCMR, MMDx finds no rejection in 70% of biopsies with BL change on histology.[21] There is speculation that MMDx can distinguish nonspecific injury ("wound healing") from immunologic injury, although data on serial MMDx scores providing proof of concept are not available. From time-to-time patients with prolonged ischemia/perfusion injury do progress to acute rejection. Whether MMDx profiles can predict such events in some patients or produce false positive signals of TCMR in others remains to be determined. There is a lot of recent literature that shows that a proportion of BL progress clinically to Banff 1A TCMR, i-IFTA, chronic active TCMR, lower eGFR and graft loss.[33–35] The key MMDx validation study to perform in this setting would be to show that these deleterious events do not occur in biopsies that fail to satisfy MMDx criteria for TCMR.

A related question that needs to be tackled is the management of patients identified as molecular TCMR by MMDx and not so recognized by histology. The clinical value of treating subclinical histologic TCMR is still debated after many years of investigation. Hence, the desirability of treating or not treating subclinical and subhistologic molecular TCMR may also be difficult to determine without well-powered randomized clinical trials and long patient follow up.

12 | CONSIDERATION OF V-LESIONS

Intimal arteritis is considered a serious lesion in renal allograft pathology since it is associated with worse graft outcomes. Traditional histology considered it to be a manifestation TCMR, but later it was found to be associated with DSA and/or positive C4d peritubular capillary staining in the majority of cases. One MMDx study found that 5/21 v-lesion biopsies with significant inflammation and 13/28 biopsies without significant inflammation had high molecular ABMR scores.[27] Evaluation of these studies should keep in mind that V-lesions may not be actually present in the tissue fragment taken for MMDX analysis. In general, MMDx is not sensitive to the presence of rare lesions; probe

signals with the least variation are actually removed during the normalization process. This creates a bias for the detection of only major transcriptional disturbances.

13 | DISTINCTION BETWEEN TCMR AND BKV NEPHROPATHY

The distinction between TCMR and BKVN is a vexing problem in clinical transplantation. As such, it was hoped that molecular methods would help make this distinction. Unfortunately, many investigators using a variety of technical platforms have found the gene expression patterns of TCMR and BKVN to be very similar.[36–39] A limited number of molecules (different in each study) were found to be expressed more often in BKVN, but the difference was relative, and not of a magnitude that would permit robust differentiation in the clinical setting.

MMDx TCMR scores are claimed to be able to detect and quantify alloimmune injury even in the presence of BKVN. However, the MMDx TCMR classifier did not include any BKVN biopsies in the training set. Therefore, one cannot claim that it can distinguish between cognate T-cell responses to viral versus allogeneic antigens. The ability of this classifier to predict TCMR after resolution of BKVN has also not been formally interrogated. Most recently, MMDx has incorporated a BKVN probability classifier into its diagnostic system.[40] Its expected clinical value is limited by the fact that it was derived from training sets containing TCMR and BKVN biopsies with unknown proportions of both diseases.

An MMDx RT-PCR protocol is now in place to measure viral VP2 mRNA in the renal allograft biopsy as a surrogate for virus activity. While such a measurement can be expected to reliably identify viral infection in kidney biopsy tissue, a threshold of detection that corresponds to a clinically relevant and actionable therapeutic plan remains to be validated. It is also noteworthy that currently there is no MMDx tool that can assist in the diagnosis of presumptive nephropathy. This is a clinical syndrome seen much more commonly than prototypical BKVN that presents with persistent BK viremia, but with sampling errors precluding a definite tissue based diagnosis.[4]

14 | USE OF MMDX FOR DIAGNOSIS OF MEDULLARY SPECIMENS

A commonly made assertion in presentations of MMDx is that it is superior to histopathology in being able to diagnose rejection in the medulla. In fact, pathologists can diagnose rejection in the medulla: they are just cautious about doing so, to avoid a false positive diagnosis in other diseases such as BKVN, bacterial interstitial nephritis and drug-induced hypersensitivity reactions. MMDx has never profiled the latter diseases and its ability to diagnose these conditions is doubtful. I am aware of MMDx false positive diagnosis of TCMR in granulomatous disease as well as acute pyelonephritis. Parenthetically, the Banff Schema does not interdict interpretation of medullary inflammation as TCMR in the right clinical setting. For the sake of consistency across medical centers, it only requires that inflammation and tubulitis be scored in the cortex, which differs from the medulla in terms of both its baseline mononuclear infiltrate as well as the amount of stromal matrix. MMDX staff should also note that histologic lesion scoring is not always necessary or sufficient for making pathologic diagnoses. Indeed, blind scoring without reference to the clinical context risks misdiagnoses such as confusing membranoproliferative glomerulonephritis or thrombotic microangiopathy for antibody-mediated rejection.

15 | ISSUES WITH HEART BIOPSIES

MMDX has utilized gene expression profiles derived from the kidney to develop a diagnostic system for the heart. Broadly speaking this is a valid approach, but there are caveats that would result in a less than perfect system.[10] Specifically, correlations for rejection-associated genes in these two organs are variable, and range from minus .4 to plus .8 for ABMR related genes. The possibility of organ specific differences in the pathogenesis of TCMR and ABMR in the heart also deserves consideration. Discrepancies between histologic and MMDX calls of rejection are substantial, as is the case for the kidney.[28] Additionally, the extent to which clinically insignificant infiltrates such as Quilty lesions and prior biopsies sites can confound molecular classifiers of TCMR is unknown. From my pathology practice, I estimate that these considerations could potentially affect the results of MMDx analyses in approximately 20% of biopsies. Some MMDX publications assert histologic mis-calls of rejection in biopsies labeled by MMDx as "injury[41]." Given the fragment-to-fragment variation in pathology lesions, this could also be the result of molecular-level tissue damage extending beyond the areas of active T-cell infiltration. Conversely, the injury related S4 score in endomyocardial biopsies is, in part, a reflection of macrophage-associated transcripts. Macrophages are pleuripotent cells, and I have seen biopsies with macrophage rich, C4d positive, DSA positive ABMR which yielded false negative results on MMDx analysis.

16 | ISSUES WITH LUNG ALLOGRAFTS

MMDx has been applied to transbronchial biopsies (TBx) for diagnosis of acute rejection in the lung.[42] There are numerous discrepancies with histology as has been observed in kidney and heart transplants. There is a need to better define the confounding effects of infection, sepsis, aspiration and smoking associated injury. Neutrophil and eosinophil rich pathology is particularly not well suited for MMDx analysis since RNAase enzymes cause rapid degradation of mRNAs in TBx.

Gene expression profiling of TBx is now being explored for the diagnosis of chronic lung allograft dysfunction (CLAD).[43] CLAD consists of two distinct syndromes, namely, bronchiolitis obliterans (BOS) and restrictive allograft syndrome (RAS). One MMDx classifier for CLAD had an AUC of .7, which improved to .86 after correction for the time dependence of gene expression. High AUC values during binary classifier development frequently do not result in good diagnostic tests

in complex clinical settings wherein multiple and often overlapping diseases are in the differential diagnosis. Moreover, the sensitivity of prediction was .56 (barely better than chance). The specificity was reportedly .9, but the correlation of molecular prediction with actual biopsy pathology is not reported. It is not even clear if the classifier performs equally well for BOS and RAS, which would raise the question of classifier being a nonspecific marker of advanced lung injury. The CLAD classifier conferred a hazard ratio of 1.33 for graft failure within 1 year of biopsy. This was much lower than a hazard ratio of 4.63 associated with a clinical diagnosis of CLAD.

Another recent study explored the prospects of diagnosing CLAD by taking mucosal biopsies from the third bifurcation of the bronchi (3BMBs).[44] These biopsies are larger than TBx, less subject to sampling error, and associated with fewer complications such as bleeding and pneumothorax, particularly in patients who are unsuitable for TBx, for example due to advanced respiratory compromise. This approach also offers greater opportunities for sampling the bronchial mucosal epithelium, albeit at a level that is more proximal to BOS and ROS associated pathology. DNA microarray analysis of 3BMBs has confirmed many "CLAD-selective" genes that were also increased in TBB, but the overall correlations were poor (r = .08–.24). Moreover, correlations were observed primarily with TBx bronchial pathology (so-called B-lesions), and not with airway lesions (A-lesions), which form the current basis for the diagnosis and treatment of acute rejection in lung allografts. A 3BMB based classifier for CLAD was not reported in this study. It remains to be seen if transcriptional profiling of 3BMB mucosa can act as a reliable surrogate for diagnosis of diseases that predominantly affect the terminal bronchioles (BOS) and alveolar interstitium (RAS). Thus, neither TBx nor 3BMB can be currently considered to be a diagnostic system that can inform clinical decision making.[44]

17 | THE UNIVERSITY OF PITTSBURGH EXPERIENCE WITH MMDX

The liver, lung, kidney, and adult transplant programs at Pittsburgh do not regard MMDx as a clinically validated system that qualifies as standard of care. The experience of the pediatric heart transplant program, which transplanted 10 patients in 2021, has been published in abstract form.[28] In essence, histologic 1R and 2R rejections were not recognized by MMDx in 25/28 and 1/4 biopsies, respectively. Molecular ABMR was diagnosed in 8/36 biopsies designated as pAMR0, and 11/18 biopsies with pABMR 1h+. Two biopsies with pAMR histology, diffuse C4d and circulating DSA were not recognized by MMDx.

This year the pediatric transplant program sent five biopsies for MMDX analysis. The "within 48 hours" turnaround time quoted by MMDx has been up to 1 week for biopsies shipped to them midweek or later. Banff grade 1A/1B acuteTCMR and chronic active TCMR were not recognized by MMDx. Examples were seen of biopsies with histologic grade t3 tubulitis, grade ct3 tubular atrophy and grade 3 arteriolar hyalinosis in which the corresponding molecular scores derived from a small tissue fragment were in the normal range. One biopsy with a high molecular DSA probability score (.71) was associated with negative DSA tests on four different occasions. Data asserting "excellent correlations" between histology rejection scores and MMDx scores has been recently published, but again expressed as overall (across-the-board) correlation coefficients, which range from -.53–.62 and cannot be equated with high precision.[45] Examination of this data with respect to biopsies that had high histologic inflammation and fibrosis scores reveals wide IQR and minimum-maximum range (as high as ~9 deciles) in the corresponding molecular scores. This reflects the patchy nature of inflammation and atrophy-fibrosis in biopsy tissue, and highlights the need for great caution in using MMDx molecular scores to assess serial changes in graft histology or graft function.

All cases sent for MMDx analysis were treated based on histology within 24 h of the biopsy, as is the current practice in the vast majority of transplant programs in the USA. Results of immunohistochemistry (C4d, BK virus) were available within this time frame. Electron microscopy was not indicated in any of these patients. Subsequent receipt of the MMDx report did not change the management in any case to justify the additional cost of $3159 per biopsy. With respect to costs, a recent Thermo Fisher sponsored study conducted by Boston Health Care Associates suggests that MMDx related costs may be mitigated as early as the 2nd year after transplantation.[46] The basic premise behind the estimates was that the ability of MMDx to predict ABMR better than histology can be assumed to translate into 25% reduction in graft loss over a period of 5 years. Factors not considered included the implications of providing MMDx-based ABMR therapy to asymptomatic patients with no histologic findings, and the consequences of potentially withholding TCMR treatment for the ~50% of histologic TCMR episodes that are not confirmed by MMDx. A scenario for multiple MMDx tests was also modeled, but excluded the cost of the for-cause biopsy, and allowed for only a single set of antirejection treatments over 5 years.

18 | CONCLUSIONS

In conclusion, I agree with McCloskey et al.'s call for MMDx to be further studied to determine which of its proposed indications can be implemented as a part of routine clinical care. In fact, the preceding sections outline several clinical settings for potential investigation. However, these studies be done with full appreciation of the caveats associated with MMDx bioinformatics algorithms. It will be important to define the incremental value of molecular methods over established standard of care protocols. Appropriate study endpoints can be modeled after other recent biomarker studies in transplantation.[47,48] A recent publication where a very promising molecular test could not be validated in an independent multicenter study should serve as reminder of why such studies are critical.[49]

ACKNOWLEDGMENTS

Dr. Ellis Mitchell, PhD, Bioinformatics Scientist, contributed Figure 4 to this manuscript. This work is supported by The Division of Transplant

Pathology, Department of Pathology, The Thomas E Starzl Transplantation Institute, University of Pittsburgh.

CONFLICT OF INTEREST

None.

DISCLOSURES

The author of this manuscript is a member of the Board of Advisors for The Renal Pathology Society

DATA AVAILABILITY STATEMENT

Not Applicable.

ORCID

Parmjeet Randhawa https://orcid.org/0000-0002-7523-3437

REFERENCES

1. Randhawa P. The Molecular Microscope (MMDX(R)) interpretation of thoracic and abdominal allograft biopsies: putting things in perspective for the clinician. *Clin Transplant*. 2021;35(4):e14223.
2. McCloskey C, Zubrycki M, Lawrence C. The Molecular Microscope Diagnostic (MMDx (R)) System: interpretation of solid organ biopsies: restoring the perspective. *Clinical Transplant*. 2022; June 6; e14711.
3. Spiegelhalter DJ, Stovin PG. An analysis of repeated biopsies following cardiac transplantation. *Statist Med*. 1983;2(1):33-40.
4. Schaub S, Hirsch HH, Dickenmann M, et al. Reducing immunosuppression preserves allograft function in presumptive and definitive polyomavirus-associated nephropathy. *Am J Transplant*. 2010;10(12):2615-2623.
5. Halloran PF. Integrating molecular and histologic interpretation of transplant biopsies. *Clin Transplant*. 2021;35(4):e14244.
6. Halloran PF, Madill-Thomsen KS. group Is. The Molecular Microscope((R)) Diagnostic System meets eminence-based medicine: a clinician's perspective. *Am J Transplant*. 2020;20:2964-2965.
7. Madill-Thomsen K, Perkowska-Ptasinska A, Bohmig GA, et al. Discrepancy analysis comparing molecular and histology diagnoses in kidney transplant biopsies. *Am J Transplant*. 2020;20(5):1341-1350.
8. Halloran PF, Reeve J, Madill-Thomsen KS, et al. The Trifecta Study: comparing plasma levels of donor-derived cell-free DNA with the molecular phenotype of kidney transplant biopsies. *J Am Soc Nephrol*. 2022;33(2):387-400.
9. Madill-Thomsen KS, Wiggins RC, Eskandary F, Bohmig GA, Halloran PF. The effect of cortex/medulla proportions on molecular diagnoses in kidney transplant biopsies: rejection and injury can be assessed in medulla. *Am J Transplant*. 2017;17(8):2117-2128.
10. Halloran PF, Potena L, Van Huyen JD, et al. Building a tissue-based molecular diagnostic system in heart transplant rejection: the heart Molecular Microscope Diagnostic (MMDx) System. *J Heart Lung Transplant*. 2017;36(11):1192-1200.
11. Reeve J, Bohmig GA, Eskandary F, et al. Generating automated kidney transplant biopsy reports combining molecular measurements with ensembles of machine learning classifiers. *Am J Transplant*. 2019;19(10):2719-2731.
12. Reeve J, Sellares J, Mengel M, et al. Molecular diagnosis of T cell-mediated rejection in human kidney transplant biopsies. *Am J Transplant*. 2013;13(3):645-655.
13. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet*. 2005;365(9458):488-492.
14. Smith RN. In-silico performance, validation, and modeling of the Nanostring Banff Human Organ transplant gene panel using archival data from human kidney transplants. *BMC Med Genom*. 2021;14(1):86.
15. Koenig A, Mezaache S, Callemeyn J, et al. Missing self-induced activation of NK cells combines with non-complement-fixing donor-specific antibodies to accelerate kidney transplant loss in chronic antibody-mediated rejection. *J Am Soc Nephrol*. 2021;32(2):479-494.
16. Callemeyn J, Senev A, Coemans M, et al. Missing self-induced microvascular rejection of kidney allografts: a population-based study. *J Am Soc Nephrol*. 2021;32(8):2070-2082.
17. Zhao D, Abou-Daya KI, Dai H, Oberbarnscheidt MH, Li XC, Lakkis FG. Innate allorecognition and memory in transplantation. *Front Immunol*. 2020;11:918.
18. Dai H, Lan P, Zhao D, et al. PIRs mediate innate myeloid cell memory to nonself MHC molecules. *Science*. 2020;368(6495):1122-1127.
19. Parkes MD, Aliabadi AZ, Cadeiras M, et al. An integrated molecular diagnostic report for heart transplant biopsies using an ensemble of diagnostic algorithms. *J Heart Lung Transplant*. 2019;38(6):636-646.
20. Madill-Thomsen K, Abouljoud M, Bhati C, et al. The molecular diagnosis of rejection in liver transplant biopsies: first results of the INTERLIVER study. *Am J Transplant*. 2020;20(8):2156-2172.
21. de Freitas DG, Sellares J, Mengel M, et al. The nature of biopsies with "borderline rejection" and prospects for eliminating this category. *Am J Transplant*. 2012;12(1):191-201.
22. Sellares J, de Freitas DG, Mengel M, et al. Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence. *Am J Transplant*. 2012;12(2):388-399.
23. Liu P, Tseng G, Wang Z, Yuchen Huang BS, Randhawa P. Diagnosis of T-cell mediated kidney rejection in formalin fixed paraffin embedded tissues using RNA-Seq based machine learning algorithms. *Hum Pathol*. 2018.
24. Sellares J, Reeve J, Loupy A, et al. Molecular diagnosis of antibody-mediated rejection in human kidney transplants. *Am J Transplant*. 2013;13(4):971-983.
25. Smith B, Cornell LD, Smith M, et al. A method to reduce variability in scoring antibody-mediated rejection in renal allografts: implications for clinical trials - a retrospective study. *Transpl Int*. 2019;32(2):173-183.
26. Madill-Thomsen KS, Bohmig GA, Bromberg J, et al. Donor-specific antibody is associated with increased expression of rejection transcripts in renal transplant biopsies classified as no rejection. *J Am Soc Nephrol*. 2021;32(11):2743-2758.
27. Salazar ID, Merino Lopez M, Chang J, Halloran PF. Reassessing the significance of intimal arteritis in kidney transplant biopsy specimens. *J Am Soc Nephrol*. 2015;26(12):3190-3198.
28. Randhawa P, Seitz A, Huang Y, Feingold B. The pathology of heart allograft biopsies: discrepancies in interpretation between conventional histology and the Molecular Microscope Diagnostic (MMDx (R)) system. *Am J Transplant*. 2021;21:361-362.
29. Haas M, Loupy A, Lefaucheur C, et al. The Banff 2017 Kidney Meeting Report: revised diagnostic criteria for chronic active T cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials. *Am J Transplant*. 2018;18(2):293-307.
30. Lefaucheur C, Gosset C, Rabant M, et al. T cell-mediated rejection is a major determinant of inflammation in scarred areas in kidney allografts. *Am J Transplant*. 2018;18(2):377-390.
31. Halloran PF, Chang J, Famulski K, et al. Disappearance of T cell-mediated rejection despite continued antibody-mediated rejection in late kidney transplant recipients. *J Am Soc Nephrol*. 2015;26(7):1711-1720.

32. Randhawa PS. T-cell mediated rejection of the kidney in the era of donor-specific antibodies: diagnostic challenges and clinical significance. *Curr Opin Organ Transplant*. 2015;20(20):325-332.
33. Mehta RB, Tandukar S, Jorgensen D, et al. Early subclinical tubulitis and interstitial inflammation in kidney transplantation have adverse clinical implications. *Kidney Int*. 2020;98(2):436-447.
34. Sood P, Randhawa PS, Mehta R, Hariharan S, Tevar AD. Donor kidney microthrombi and outcomes of kidney transplant: a single-center experience. *Clin Transplant*. 2015;29(5):434-438.
35. Nankivell BJ, Agrawal N, Sharma A, et al. The clinical and pathological significance of borderline T cell-mediated rejection. *Am J Transplant*. 2019;19(5):1452-1463.
36. Pan L, Lyu Z, Adam B, et al. Polyomavirus BK nephropathy-associated transcriptomic signatures: a critical reevaluation. *Transplant Direct*. 2018;4(2):e339.
37. Mannon RB, Hoffmann SC, Kampen RL, et al. Molecular evaluation of BK polyomavirus nephropathy. *Am J Transplant*. 2005;5:2883-2893.
38. Adam BA, Kikic Z, Wagner S, et al. Intragraft gene expression in native kidney BK virus nephropathy versus T cell-mediated rejection: prospects for molecular diagnosis and risk prediction. *Am J Transplant*. 2020.20:3486-3501
39. Sigdel TK, Bestard O, Salomonis N, et al. Intragraft antiviral-specific gene expression as a distinctive transcriptional signature for studies in polyomavirus-associated nephropathy. *Transplantation*. 2016;100(10):2062-2070.
40. Halloran PF, Madill-Thomsen K, Bohmig GA, et al. A two-fold approach to polyoma virus (BK) nephropathy in kidney transplants: distinguishing direct virus effects from cognate T-cell-mediated inflammation. *Transplantation*. 2021;105:2374-2384.
41. Halloran PF, Reeve J, Aliabadi AZ, et al. Exploring the cardiac response to injury in heart transplant biopsies. *JCI insight*. 2018;3(20):e123674.
42. Halloran KM, Parkes MD, Chang J, et al. Molecular assessment of rejection and injury in lung transplant biopsies. *J Heart Lung Transplant*. 2019;38(5):504-513.
43. Parkes MD, Halloran K, Hirji A, et al. Transcripts associated with chronic lung allograft dysfunction in transbronchial biopsies of lung transplants. *Am J Transplant*. 2022;22(4):1054-1072.
44. Halloran K, Parkes MD, Timofte IL, et al. Molecular phenotyping of rejection-related changes in mucosal biopsies from lung transplants. *Am J Transplant*. 2020;20(4):954-966.
45. Gupta G, Moinuddin I, Kamal L, et al. Correlation of donor-derived cell-free DNA with histology and molecular diagnoses of kidney transplant biopsies. *Transplantation*. 2022;106(5):1061-1070.
46. Fusfeld L, Menon S, Gupta G, Lawrence C, Masud SF, Goss TF. US payer budget impact of a microarray assay with machine learning to evaluate kidney transplant rejection in for-cause biopsies. *J Med Econ*. 2022;25(1):515-523.
47. Friedewald JJ, Kurian SM, Heilman RL, et al. Development and clinical validity of a novel blood-based molecular biomarker for subclinical acute rejection following kidney transplant. *Am J Transplant*. 2019;19(1):98-109.
48. Van Loon E, Gazut S, Yazdani S, et al. Development and validation of a peripheral blood mRNA assay for the assessment of antibody-mediated kidney allograft rejection: a multicentre, prospective study. *EBioMedicine*. 2019;46:463-472.
49. Van Loon E, Giral M, Anglicheau D, et al. Diagnostic performance of kSORT, a blood-based mRNA assay for noninvasive detection of rejection after kidney transplantation: a retrospective multicenter cohort study. *Am J Transplant*. 2021;21(2):740-750.

How to cite this article: Randhawa P. The MMDx® diagnostic system: A critical re-appraisal of its knowledge gaps and a call for rigorous validation studies. *Clin Transplant*. 2022;e14747. https://doi.org/10.1111/ctr.14747