**Improving diagnostic performance of kidney allograft rejection with a model combining relative fraction and absolute copies of donor-derived cell-free DNA - results from five independent cohorts**

Louise Benning[1¥], Aylin Akifova[2¥], Michael Oellerich[3], Bilgin Osmanodja[2], Christian Morath[1,4], Julia Beck[5], Ana Schuetz[6], Kirsten Bornemann-Kolatzki[5], Eva V. Schrezenmeier[2], Thuong Hien Tran[7], Vedat Schwenger[8], Maria Shipkova[9], Eberhard Wieland[9], Ekkehard Schuetz[5,6], Klemens Budde[2*]

¥ shared authorship

1 Department of Nephrology, Heidelberg University Hospital, Heidelberg, Germany
2 Department of Nephrology and Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany
3 Department of Clinical Pharmacology, George-August-University Medical Center Göttingen, Göttingen, Germany
4 Department of Nephrology and Hypertension, Klinikum Nuremberg, Paracelsus Medical University, Nuremberg, Germany
5 Chronix Biomedical GmbH, Göttingen, Germany
6 Insight Molecular Diagnostics Inc., Nashville TN, USA
7 Institute of Immunology, Heidelberg University Hospital, Heidelberg, Germany
8 Department of Nephrology, Transplant Center, Klinikum Stuttgart, Stuttgart, Germany
9 Bioscientia Ingelheim, Germany

*Corresponding author:

Professor Klemens Budde
Department of Nephrology and Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany
Charitéplatz 1, 10117 Berlin
klemens.budde@charite.de

**Running head: dd-cfDNA in kidney transplantation**

**Abstract**: 238/300 words

**Lay Summary (Clinical):** 147/150 words

**Word count of the Main Text**: 3666/4000

**Figures**: 4

**Tables**: 4

**References**: 40

**Abstract**

**Introduction**

Donor-derived cell-free DNA (dd-cfDNA) is a standard-of-care biomarker in kidney transplantation, reported as percentage of total cfDNA or copies/mL. We hypothesized that combining both metrics into a continuous composite score would reduce false classifications thus improving the accuracy of rejection diagnosis.

**Methods**

We analyzed 443 dd-cfDNA measurements in 383 individual patients from five independent kidney transplant cohorts with both percentage and copies/mL. A random discovery set of 31 biopsy-proven rejection and 29 non-rejection cases was used to derive a continuous composite score (CM-Score) integrating dd-cfDNA (%) and copies/mL. Fixed thresholds from the discovery group were validated in 75 rejections and 279 non-rejections for diagnostic performance and compared to twelve published cohorts.

**Results**

The CM-Score outperformed dd-cfDNA percentage and cp/mL alone. At the predefined threshold (at 25% prevalence), the CM-Score retained a high NPV of 91% (89–93%), while significantly improving PPV to 80% (78–83%; P<0.002), compared to the published values (weighted average NPV: 88%; 89–90%; PPV: 51% 50–53%, N=6,861). In addition, a decision curve analysis yielded a significantly higher net benefit for the CM-Score (*P*<0.003; prevalence 10-25% at threshold 15-50%). This superior diagnostic performance of the CM-Score can broaden the applicability of dd-cfDNA to real-world transplantation populations.

**Conclusions**

The CM-Score is a robust and clinically meaningful tool that improves diagnostic accuracy for both ruling-out and ruling-in rejection using dd-cfDNA and establishes itself as a superior stand-alone metric with clear potential to improve decision-making in kidney transplantation.

**Keywords**: donor-derived cell-free DNA, dd-cfDNA, kidney transplantation, rejection, biomarker

**Lay Summary (Clinical)**

Kidney transplant failure remains a major concern, and early detection of injury is crucial to prevent long-term damage. A less invasive alternative to a biopsy is a simple blood test that measures tiny fragments of DNA from the transplanted kidney, called donor-derived cell-free DNA (dd-cfDNA). These results can be reported either as a percentage of total DNA or as the number of DNA copies in the blood. We examined whether combining both measurements could improve the detection of rejection. Using data from five transplant cohorts, we developed a combined measure (CM-score) and validated it in 279 samples without rejection and 75 with rejection. At an assumed rejection rate of 25%, it correctly identified rejection with a positive predictive value of 80% and ruled it out with a negative predictive value of 91%. Overall, the CM-score may help doctors diagnose rejection more accurately and make timely treatment decisions.

## 1 Introduction

Donor-derived cell-free DNA (dd-cfDNA) has become an integral biomarker in kidney transplantation aftercare and standard-of-care (SOC) in clinical practice in both for-cause and routine surveillance. Most commercial tests report dd-cfDNA as percentage of total cell-free DNA (cfDNA), which is referred to as relative quantification of dd-cfDNA (dd-cfDNA%)[1]. Although dd-cfDNA% has been used to assess the risk of rejection for more than a decade, it has limitations. dd-cfDNA% is calculated by dividing dd-cfDNA by total cfDNA, the latter mainly derived from the recipient, thereby creating an inverse relation between cfDNA quantity and dd-cfDNA%. Fluctuations in recipient cfDNA will impact the dd-cfDNA%. Total cfDNA levels can fluctuate for various reasons, including physical activity, infections, malignancies, or intake of calcineurin inhibitors[2]. This can lead to falsely high or low dd-cfDNA% results that are independent of graft health and status[3]. Using a value that is unrelated to the graft as the denominator will add noise and overshadow the true amount of dd-cfDNA that is shed by the graft.

To address the limitations of dd-cfDNA%, several studies have evaluated absolute quantification of dd-cfDNA in genomic copies per milliliter of plasma (cp/mL), consistently demonstrating superior performance compared to dd-cfDNA% alone[4–6].

However, reporting two distinct values may lead to disagreement, producing divergent results compared to a single fixed threshold. Consequently, parallel testing approaches have been proposed[4,6,7] in which an elevated result from either assay would be interpreted as indicative of an increased injury risk. The latter has recently also applied in heart transplant recipients[8]. Such approaches may increase sensitivity by capturing more true cases of rejection, but at the expense of lowering specificity, since isolated false-positive results from either percentage or cp/mL would lead to a false positive classification.

We hypothesized that integrating dd-cfDNA% and copies/mL into a continuous combination score would overcome the limitations of single thresholds, enhance robustness and maintain high specificity for rejection diagnosis. By combining relative and absolute measures, this approach may reduce the influence of outliers and better distinguish true rejection from background variation. We used a robust discovery validation approach to prove our hypothesis.

## 2 Methods

### 2.1 Study Design

We combined 443 dd-cfDNA results from 383 patients of five independent cohorts where dd-cfDNA was reported as percentage and absolute dd-cfDNA[2,5,9–12]. Details of the patient populations are reported in the original publications. 265 samples were from patients undergoing indication biopsies. BK-virus associated nephropathy (BKVAN samples) included 32 samples with high BK-viraemia defined as >2.5-fold above the American Society of Transplantation (AST) threshold for presumptive BKVAN (>4 $\log_{10}$/mL), and 13 cases with biopsy-proven BKVAN. 25 Control samples with unremarkable histology were included. Histological findings (Table 1) were classified per Banff criteria at time of publication. The 146 stable samples (without biopsy) were from 86 clinically stable kidney recipients, defined by stable and normal serum creatinine over the last 6 months within a routine surveillance setting.

Biopsy indication was determined by the treating physicians. This combined cohort, therefore, reflects a broad spectrum of real-world complications commonly observed in both early and late post-transplant periods. This includes not only early delayed graft function (DGF), or borderline/T cell-mediated rejection (TCMR), but also rather late complications such as antibody-mediated rejection (ABMR) and drug toxicity (e.g., calcineurin-inhibitor (CNI) toxicity). The wide temporal range of biopsies in our cohort highlights its real-world relevance and distinguishes it from most studies, which often focus primarily on early events in the first year posttransplant (See Supplementary Figure S1).

Blood samples for dd-cfDNA assessment were obtained before or at least 12 hours after biopsy. dd-cfDNA was quantified using two versions[5,13,14] of a single-nucleotide polymorphism (SNP)-based assay at least ≥14 days post-engraftment, leveraging droplet-

digital PCR (ddPCR) technology described by the Supplementary Table S1. Both methods have been validated to show equivalence (Supplementary Figure S2).

**2.2 Combination of percent and copies/mL**

To combine percent and cp/mL into a continuous value, we first assessed correlation between cp/mL and dd-cfDNA percentage in the interval from LLoQ to 2× upper limit of normal (0.5% and 50 cp/mL[5]). No significant correlation was found (P=0.12, Supplementary Figure S3).

Linear, logarithmic, and quadratic models were evaluated in a discovery cohort of biopsy-proven rejection (N=29) and non-rejection cases (N=31), using ROC-AUC as the performance endpoint. The cohort size met recommended standards for two-variable model development, exceeding the minimum of 20 cases per group based on the guideline of including at least 10 times the number of variables[15–17].

The optimal model (Eq.1) achieved an AUC of 0.942 (95% CI 0.890–0.977).

$$\boldsymbol{Eq.1}:\ Model\ value = \left(\frac{Percent\ Value}{0.71}\right)^{0.5} + \left(\frac{cp/mL\ Value}{32}\right)^{0.6}$$

**2.3 Thresholds**

Optimal cut-offs that maximized the Youden index were 2.295 for the model, 0.47% for dd-cfDNA%, and 44 cp/mL. We converted the model output into the "CM-Score" by dividing the value by 2.295 and subtracting 1, so that negative scores indicate "not suspicious" and positive scores indicate "suspicious for rejection". We then validated this approach in 279 non-rejection samples and 75 rejection samples (Supplementary Table S2).

**2.4 Statistics**

Analyses used Python, R, or Excel. Statistical comparisons used Mann-Whitney *U*-test. Confidence intervals (CI) were calculated via inverse beta distribution[18]. Likelihood ratios at discrete CM-Score values were defined as the product of the positive ($LR^+$) and negative ($LR^-$

) likelihood ratios at the exact value. For CM-Score the values between -1 and +1 $LR^+$ and $LR^-$ were derived from the ROC curve. When the yielded post-test probability values were averaged for values above the cut-off of zero and for values below the cut-off, the correlation of post-test rejection probabilities (prevalences between 5% and 50%) with the 1-negative predictive value (NPV)-below zero and positive predictive value (PPV) - above zero was calculated with $R^2$=0.999 and $f(x)=1.01x - 0.01$, which is within the rounding error of the calculations, *quod erat demonstrandum*.

For the meta-analysis of published data, pooled estimates of sensitivity, specificity, PPV, and NPV were calculated using inverse variance weighting[19]. Pooled estimates were calculated as weighted means with inverse variance weighting ($wi = 1/SEi^2$), and standard error $SE = \sqrt{(1/\Sigma wi)}$. Between-group comparisons used two-sample Z-tests for independent proportions[20], with $Z = (\theta_1 - \theta_2) / \sqrt{(SE_1^2 + SE_2^2)}$, where $\theta_1$ and $\theta_2$ represent group estimates and $SE_1$ and $SE_2$ their standard errors. Two-tailed p-values were derived from the standard normal distribution. Resulting sensitivity/specificity estimates with CIs were used for decision curve analysis[21,22].

## 3 Results

### 3.1 Diagnostic Performance of the CM-Score Compared with Individual dd-cfDNA Metrics

We first present the univariate data evaluation by variable and clinical groups. Figure 1 shows the boxplots of dd-cfDNA% (A), cp/mL (B), and the CM-Score (C) segregated by biopsy group. The statistical significances are given in Table 2, which provides support that dd-cfDNA is specific for rejection and not influenced by other common histological findings, as published elsewhere[9,10,23]. The numeric values in the various biopsy groups were compared. Only the CM-Score could distinguish all rejection types from the Control group as well as distinguish all non-rejection pathologies from TCMR, which is notoriously difficult to detect using relative or absolute dd-cfDNA alone (Table 2). We demonstrate a clear and significant increase of dd-cfDNA with increasing severity of TCMR (Supplementary Figure S4).

Figure 2 visualizes the data of the ROC curves of the validation group. Although the overall AUC was not significantly different between the CM-Score and the individual dd-cfDNA metrics, visual inspection of the ROC curve shape suggests that the combination model provides improved diagnostic sensitivity and specificity at the thresholds defined in the Discovery Cohort.

As expected, dd-cfDNA demonstrated markedly superior performance compared with estimated glomerular filtration rate (eGFR), which showed a sensitivity of 39% (31–48%), a specificity of 82% (77–86%), an AUC of 0.62 (0.49–0.73), an NPV of 80% (77–83%), and a PPV of 42% (31–53%). As supplementary metrics, the areas under the precision–recall curves (PRAUC) were 0.76 (0.73–0.79) for the CM-Score, 0.73 (0.71–0.77) for dd-cfDNA percentage, and 0.71 (0.69–0.74) for copies/mL.

### 3.2 Performance Characteristics in the Validation Cohort

We then evaluated the most clinically relevant estimators, NPV and PPV, calculated at a presumed rejection prevalence of 25% for the validation cohort. Figure 3 shows the NPV and PPV for CM-Score, percentage, and cp/mL dd-cfDNA. Again, the single-point cut-offs established in the discovery group were applied to the validation group. While the NPV was maintained without significant decline, the PPV improved significantly when using the CM-Score ($P < 0.001$). dd-cfDNA expressed as copies per milliliter showed a better PPV than dd-cfDNA percentage, but the CM-Score outperformed either single metric alone. The full set of performance metrics, including AUC, sensitivity, specificity, PPV, and NPV, is provided in Table 3.

**3.3 Contextualization with Evidence from Other Clinical Cohorts**

To fully understand the value of the CM-Score, an additional analysis was performed to compare its established performance characteristics to 12 clinical cohorts from 11 published studies where percent and/or absolute values or values with added clinical parameters were reported in both for-cause and surveillance settings. Larger clinical studies with published sensitivity and specificity were used to calculate the predictive values at a rejection prevalence of 25%. A total of 6,861 samples with 1,809 rejection cases are incorporated (Supplementary Table S3 and Supplementary Figure S5). Figure 4 shows the results comparing the performance of the selected clinical cohorts to the CM-score[5,6,23–31].

The inclusion of both percent and cp/mL based on the CM-Score described herein significantly improved the PPV to 80% in the validation cohort while preserving the typically high NPV, confirming the hypothesis that a continuous combination of percentage and cp/mL would outperform other means of interpreting dd-cfDNA. The full analysis is given in Supplementary Table S4, which amongst other information also shows that there is no difference between the validation group and the full dataset. Therefore, subsequent analyses were based on the full dataset. A decision curve analysis further compared the clinical utility

of the validation cohort results with published data by quantifying net benefit across a range of threshold probabilities (Supplementary Figure S6 and Supplementary Table S5).

### 3.4 Clinical Application

Predictive values are key performance metrics for clinical practice. Calculated via Bayes' theorem, they depend heavily on prevalence as the pre-test probability. In the for-cause biopsy cohort, the prevalence of rejection was 36%, whereas surveillance cohorts typically report rates of 6–11%[32]. In general, the ABMR incidence ranges from 3% to 12%[33]. In DSA-positive patients, the prevalence of rejection might be up to 40% per year[34,35], associated with a high PPV[35].

Naesens and Wong recently outlined how dd-cfDNA results can be integrated into clinical decision-making[36]. Using the likelihood ratios (LR+ and LR–) associated with a test's defined threshold, the prior probability of rejection can be updated to a post-test probability[37]. The general assumption is that an $LR^+ > 10$ and an $LR^- < 0.33$ are considered necessary to confidently rule in or rule out disease. However, these criteria are not met in the published dd-cfDNA cohorts (see Supplementary Table S3).

To support context-specific interpretation, we can use the LRs to exemplify the gain in post-test probability. For the CM-Score at the threshold of 0, the $LR^+$ is 13.4 (7.7–19.0), and the $LR^-$ is 0.28 (0.2–0.36). Supplementary Figure S7 shows the Fagan nomogram with the respective post-test probability, which is a function of the likelihood ratios (LR). Using the single-point cut-off of 0 for the CM-Score and a pre-test probability of 5%, the post-test probabilities are 43% and 1.5%. At 25% pre-test probability, the post-test probabilities are 82% and 8.5% respectively.

Supplementary Table S6a shows the positive likelihood ratios across the range of expected CM-Score values and pre-test probabilities from 5% to 50%. The corresponding post-test probabilities of rejection for CM-Score values exceeding each threshold are also provided.

Conversely, Supplementary Table S6b shows the post-test probability when the CM-Score falls below the specified threshold.

Taking this further, LRs can be calculated for the spectrum of expected values of the CM-Score. This gives more reliable information, based on the actual value rather than a fixed single point delimitator (e.g., 0) as above. The continuous matrix for distinct CM-Score values is presented in Table 4, providing a more realistic and useful representation of the probabilities that a patient will have a rejection or not. For example, if the CM-Score result is ≥0.6 and the pre-test probability is 50% (a reasonable estimate in cases classified as Borderline, also referred to as suspicious for TCMR, where we can assume equal probabilities), the result would indicate a post-test probability of a rejection of 94%, resulting in a high probability of rejection. On the other hand, with a very low value of e.g. -0.8, the results would be 3%, suggesting that no rejection is present.

The same interpretation applies in a surveillance setting, where the probability of rejection is approximately 6-11%[32]. In this context, a CM-Score value of ≥0.6 raises the post-test probability to ≥62% whereas a negative score of ≤-0.8 reduces it to below 0.5%. Thus, comparison with pre-test probability should guide clinical interpretation: a pronounced increase in the positive post-test probability supports rejection, while a marked decrease argues against it.

Table 4 illustrates these dependencies of CM-Scores at specific threshold values. Individual transplant centers can apply their own experience to estimate the local prevalence of rejection in various clinical scenarios and then use Table 4 to determine the corresponding post-test probabilities, supporting more informed clinical decision-making.

## 4 Discussion

Several publications have demonstrated that adding dd-cfDNA absolute quantity to the percentage value improves diagnostic performance in transplant recipients, including but not limited to those with kidney transplants[4,6,8]. Our results support this finding: at a rejection prevalence of 25%, the CM-Score achieved a PPV of 82%, compared to a weighted average PPV of only 51.4% (50.1%–52.7%) reported in published studies (Figure 4 and Supplementary Table S3). Such low PPVs, however, complicate the interpretation of positive test results. Consequently, dd-cfDNA's primary clinical utility has been as a rule-out test, supported by consistently high NPVs across studies (weighted average NPV of 89%; 88.3%–89.6%) with commercially available assays. However, the low PPV limits utility, as elevated results provide limited guidance for clinical decision-making. The higher PPV achieved by combining percentage and absolute dd-cfDNA into the continuous CM-Score indicates that dd-cfDNA may also have value as a rule-in biomarker. With a PPV of 82%, the probability of correctly identifying ongoing rejection is substantially increased, improving confidence in positive results. This improvement reflects the superior performance of the continuous score over simple thresholds or parallel combinations and is supported by a corresponding 168% increase in LR+, indicating markedly stronger rule-in capability. This may be particularly relevant in surveillance settings, where rejection prevalence is lower (6-11%) than in for-cause settings[32].

With dd-cfDNA% alone or using simple combinations, even including additional clinical predictors, positive results often lead to overdiagnosis. Based on published sensitivities and specificities (Supplementary Table S3), the average PPV at 10% prevalence is only 28% (Figure 4). In current practice, dd-cfDNA replaces surveillance biopsies mainly due to its high NPV, allowing to forgo biopsies when results are below thresholds. However, values above the threshold lead to unnecessary biopsies (72%) due to low PPV. Adding longitudinal data

provided limited benefit; Bromberg et al. reported a PPV of 31% at assumed 10% rejection prevalence, within the range of studies without longitudinal data[29]. Figure 5 shows that, in published clinical cohorts, current dd-cfDNA-based surveillance approaches are associated with a high overdiagnosis ratio, with 2.6 unnecessary biopsies performed for every necessary one (odds of 2.6). Using the dichotomized CM-Score would increase the positive biopsies yield to 60%, reducing overdiagnosis odds to 0.68 and to ≤0.54 at a high CM-Score of ≥0.8.

Further, we show that the CM-Score best discriminates rejection (both TCMR and ABMR) from other pathologies compared to either metric alone. Negative values reliably exclude non-rejection pathologies (e.g., BKVAN, CNI toxicity) with 91% NPV, while positive values predict rejection of any type. Although dd-cfDNA is primarily a damage biomarker, our findings support its relative specificity for rejection. Prior studies likewise demonstrate dd-cfDNA's specificity for rejection[23], particularly in ABMR, with elevated values also observed in DSA-negative microvascular inflammation, likely reflecting endothelial injury[9,23]. Incorporating both absolute and relative values into the CM-Score captures dynamic changes more precisely, eventually improving utility for longitudinal and routine monitoring.

Aubert et al. recently demonstrated an increased AUC in a large cohort when dd-cfDNA was added to a standard of care clinical model[23]. However, dd-cfDNA percentage alone (AlloSure™/AlloSeq™) showed limited utility, with a PPV of 54% at the standard 1.0% threshold and an NPV of 86% (validation group results; calculated at a 25% rejection prevalence). When combined with clinical data, the NPV increased to 91%, whereas the PPV decreased to 52%. Clinical data alone had an NPV of 88% and a PPV of 50%. This suggests that percentage alone adds marginally to diagnostic value. For surveillance settings, assuming a 10% rejection prevalence, the PPV was 22% at the 0.5% threshold and 28% for the 1% threshold. This was confirmed in the Kidney Outcomes with AlloSure Registry (KOAR) study[29], which reported that dd-cfDNA >1% or >0.5% with ≥61% relative change (RCV)

yielded PPVs of 57% at 25% prevalence (and 60% using the 1.0% cut-off alone). For surveillance testing (and a 10% prevalence), the PPV drops to 31% and 34%, corresponding to overdiagnosis odds of 2.25 and 1.96, respectively.

As shown elsewhere, it should be understood that focusing solely on the ROC AUC (which is a summation value without considering the curve shape) provides limited information regarding the added clinical benefit of a diagnostic test[38,39]. The results shown by Aubert et al.[23] exemplify this by demonstrating a significant increase in AUC by adding dd-cfDNA without any effect on the PPV. In contrast, we show herein that even without a significant improvement of the AUC, the truly clinically important parameters of NPV and PPV can substantially and significantly improve (Figure 2).

Bromberg et al. reported results from the ProActive registry study[31], which included mainly surveillance biopsies. At a prevalence of 25%, dd-cfDNA% yielded an NPV of 93% and a PPV of 64%, which corresponds to a PPV of 37% in a surveillance setting. Halloran et al. previously evaluated the same assay platform in a 218-patient for-cause biopsy cohort, assessing both percentage and copies/mL measurements[6]. Percentage and copies/mL were combined using a parallel testing rule, in which exceeding either threshold (≥1% or ≥78 cp/mL) constituted a positive result for rejection. When compared with Banff-based histology, this approach yielded an NPV of 90% and a PPV of 56% at a 25% prevalence, consistent with other published findings. However, in a surveillance setting, the PPV would decrease to 30%.

Using a two-threshold approach in our study with optimized values (≥0.72% and ≥69cp/mL) yielded an NPV of 92% (90%–94%) and a PPV of 72% (68% –77%) in our cohort of 106 rejection and 307 non-rejection cases at a 25% rejection prevalence. This performance compared to the same approach reported for NGS-based assays[6] may be due to the use of digital PCR for cfDNA quantification, together with per-sample extraction efficiency measurement. The reliability of digital PCR for DNA quantification is well established and is

cited as an example of a primary reference measurement procedure in ISO17511:2020[40]. cfDNA extraction from plasma can vary substantially, and controlling for this adds reliability to quantification.

Furthermore, adding clinical values (e.g., eGFR, HLA-mismatch, age, time since transplantation) did not improve diagnostic performance over the CM-Score (data not shown), likely due to the already high sensitivity and specificity and because most conventional clinical parameters are prognostic rather than diagnostic. Thus, the CM-Score provides diagnostic information without requiring additional clinical parameters or composite calculators. With its high PPV and strong specificity for rejection, the CM-Score supports a rejection-directed diagnosis, helping distinguish rejection from other causes such as CNI toxicity in cases of rising creatinine or decreasing eGFR.

Table 5 shows the importance of considering the pre-test probability of rejection when interpreting results. The value of using a continuous score rather than dichotomized thresholds, such as the multiple percentage cut-offs proposed by Bromberg et al.[29], or the evaluation of percentage and cp/mL described by Halloran et al.[6], is evident. A continuous metric enables interpretation of dd-cfDNA values within the clinical context and is particularly useful in unselected surveillance settings, where the prevalence of rejection is substantially lower than in for-cause testing. In such scenarios, only clearly elevated results should prompt intervention to minimize the risk of overdiagnosis and overtreatment. For example, if a 50% post-test probability is required to justify biopsy and the baseline rejection rate is 10%, Table 5 indicates that a CM-Score >0.3 would represent the appropriate threshold.

The introduction of the clinical application guide, considering the pre-test probability of the event to evaluate the likelihood of the results, allows for a more personalized approach to patient care. It acknowledges that the same result may carry different implications depending

on the clinical context and patient's individual risk profile (e.g. high immunological risk), thereby moving away from a uniform, "one-size-fits-all" interpretation to apply.

In summary, the CM-Score is a robust tool that improves diagnostic accuracy for both ruling out and ruling in rejection, outperforming dd-cfDNA% alone or with incorporating traditional clinical parameters, establishing itself as a superior stand-alone metric in kidney transplantation.

**Limitations**

These data are derived from real-world practice rather than controlled studies, which may be seen as a limitation but more accurately reflects the use of dd-cfDNA in routine care. As most biopsies were for cause, extrapolation to surveillance settings may be limited, although pathologist interpretation by Banff criteria should not differ substantially. Thus, applying Bayes' theorem to estimate post-test probabilities at different pre-test probabilities appears valid. Notably, one publication[31] showed slightly improved LRs for surveillance compared to for-cause. Aubert et al. hypothesized that surveillance might result in lower dd-cfDNA values in ABMR based on differences observed in their cohorts[23]. Inspection of Figure 2 and Extended Figure 4 reveals a roughly twice as high dd-cfDNA for all Banff groups in the validation cohort, rather suggesting a methodological bias between the two assays. For our study, central pathology was not available, and molecular pathology (MMDx) was limited. Incorporating MMDx could have improved NPV and PPV, given its ~10% higher condordance[6].

**Disclosure Statement**

Kits and reagents for dd-cfDNA assessment were provided by Insight Molecular Diagnostics Inc. (iMDx; GraftAssure), which also supplied the testing and infrastructure used to generate the study results.

A.A. received travel support from Insight Molecular Diagnostics Inc.

M.O. and K.B. act as consultants to Insight Molecular Diagnostics Inc. M.O. received travel support from Insight Molecular Diagnostics Inc. BO received travel support from Insight Molecular Diagnostics Inc. and honoraria from AstraZeneca. J.B. and K.B.-K. are employees of Chronix Biomedical GmbH (an Insight Molecular Diagnostics Inc. company), E.S. and A.S. are employees of an Insight Molecular Diagnostics Inc., which holds intellectual property rights (EP 3004388B, EP3201361B1, US11155872B2, US10570443B2).

The Department of Nephrology and the Transplantation Immunology, University Hospital Heidelberg (L.B., C.M. and H.T.) has received kits and supplies for dd-cfDNA testing by CareDx and by Insight Molecular Diagnostics Inc.

EV.S. reports relationships with: AstraZeneca Pharmaceuticals LP (consulting or advisory and speaking and lecture fees), GlaxoSmithKline LLC (speaking and lecture fees), Otsuka Pharmaceutical Co Ltd (consulting or advisory), Chiesi (travel reimbursement), Novartis (speaking and lecture fees).

K.B. has received research funds and/or honoraria from: Aicuris, Alexion, Astellas, AstraZeneca, Biogen, Biohope, Carealytics, CareDx, Chiesi, CSL Behring, DTB GmbH, Eledon, HiBio, iMDx, MSD, Natera, Neovii, Oncocyte, Oska, Otsuka, Paladin, Pfizer, Pirche, Sanofi, smart care solutions, Stada, Takeda, Veloxis, Vifor and Xenothera.

**Declaration of Generative AI and AI-assisted technologies in the writing process**

During the preparation of this work the authors used ChatGPT (GPT-5, developed by OpenAI) for assistance with language refinement and rephrasing to improve readability. After using this

tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

**Data Sharing Statement**

The data that support the findings of this study are not publicly available, as public data sharing was not included in the ethics approval or patient consent. However, deidentified data underlying the results of this article can be made available upon reasonable request to the corresponding author, provided that such sharing is in accordance with institutional and ethical regulations. Summary statistics and analytic code may also be shared on reasonable request to enable reproducibility of the findings.

**Supplementary Material**

Supplementary material is available online.

**Funding**

Louise Benning is funded by the Olympia Morata Program of Heidelberg University.

**Author's contributions**

LB, AA, ES, and KB designed the study. LB, AA, JB, AS, KBK, ES, and KB analyzed and interpreted the data and drafted the manuscript. Patients were recruited and results generated by LB, AA, BO, CM, EVS, VS, and KB. JB, ES, THT, MS, and EW established and performed the quantification of dd-cfDNA in their respective laboratories. MO, CM, VS, and KB supervised the project and revised the manuscript. All the authors critically reviewed the manuscript.

**References**


1. Oellerich M, Sherwood K, Keown P, et al. Liquid biopsies: donor-derived cell-free DNA for the detection of kidney allograft injury. *Nat Rev Nephrol*. 2021;17(9):591-603. doi:10.1038/s41581-021-00428-0

2. Schütz E, Asendorf T, Beck J, et al. Time-Dependent Apparent Increase in dd-cfDNA Percentage in Clinically Stable Patients Between One and Five Years Following Kidney Transplantation. *Clin Chem*. 2020;66(10):1290-1299. doi:10.1093/clinchem/hvaa175

3. Osmanodja B, Akifova A, Budde K, et al. Absolute or Relative Quantification of Donor-derived Cell-free DNA in Kidney Transplant Recipients: Case Series. *Transplant Direct*. 2021;7(11):e778. doi:10.1097/txd.0000000000001237

4. Bunnapradist S, Homkrailas P, Ahmed E, Fehringer G, Billings PR, Tabriziani H. Using both the Fraction and Quantity of Donor-Derived Cell-Free DNA to Detect Kidney Allograft Rejection. *JASN*. 2021;32(10):2439-2441. doi:10.1681/asn.2021050645

5. Oellerich M, Shipkova M, Asendorf T, et al. Absolute quantification of donor-derived cell-free DNA as a marker of rejection and graft injury in kidney transplantation: Results from a prospective observational study. *Am J Transplant*. 2019;19(11):3087-3099. doi:10.1111/ajt.15416

6. Halloran PF, Reeve J, Madill-Thomsen KS, et al. Combining Donor-derived Cell-free DNA Fraction and Quantity to Detect Kidney Transplant Rejection Using Molecular Diagnoses and Histology as Confirmation. *Transplantation*. 2022;106(12):2435-2442. doi:10.1097/tp.0000000000004212

7. Whitlam JB, Ling L, Skene A, et al. Diagnostic application of kidney allograft-derived absolute cell-free DNA levels during transplant dysfunction. *Am J Transplant*. 2019;19(4):1037-1049. doi:10.1111/ajt.15142

8. Kim PJ, Olympios M, Sideris K, et al. A two-threshold algorithm using donor-derived cell-free DNA fraction and quantity to detect acute rejection after heart transplantation. *Am J Transplant*. 2025;25(9):1895-1905. doi:10.1016/j.ajt.2025.04.021

9. Akifova A, Budde K, Choi M, et al. Association of Blood Donor-derived Cell-free DNA Levels With Banff Scores and Histopathological Lesions in Kidney Allograft Biopsies: Results From an Observational Study. *Transplant Direct*. 2025;11(5):e1794. doi:10.1097/txd.0000000000001794

10. Benning L, Morath C, Fink A, et al. Donor-Derived Cell-Free DNA (dd-cfDNA) in Kidney Transplant Recipients With Indication Biopsy—Results of a Prospective Single-Center Trial. *Transpl Int*. 2023;36:11899. doi:10.3389/ti.2023.11899

11. Kühn T, Speer C, Morath C, et al. Immune Response to COVID-19 mRNA Vaccination in Previous Nonresponder Kidney Transplant Recipients After Short-term Withdrawal of

Mycophenolic Acid 1 and 3 Months After an Additional Vaccine Dose. *Transplantation*. 2023;107(5):1139-1150. doi:10.1097/tp.0000000000004516

12. Schrezenmeier E, Rincon-Arevalo H, Jens A, et al. Temporary antimetabolite treatment hold boosts SARS-CoV-2 vaccination-specific humoral and cellular immunity in kidney transplant recipients. *Jci Insight*. 2022;7(9):e157836. doi:10.1172/jci.insight.157836

13. Loi L, Reineke M, Zeier M, et al. #1269 Quantifying donor-derived cell-free DNA for monitoring kidney transplant recipients: a head-to-head comparison of NGS and ddPCR assays. *Nephrol Dial Transplant*. 2025;40(Supplement_3):gfaf116.0847. doi:10.1093/ndt/gfaf116.0847

14. Beck J, Bierau S, Balzer S, et al. Digital Droplet PCR for Rapid Quantification of Donor DNA in the Circulation of Transplant Recipients as a Potential Universal Biomarker of Graft Injury. *Clin Chem*. 2013;59(12):1732-1741. doi:10.1373/clinchem.2013.210328

15. Vittinghoff E, McCulloch CE. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. *Am J Epidemiology*. 2007;165(6):710-718. doi:10.1093/aje/kwk052

16. Smeden M van, Groot JAH de, Moons KGM, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Méd Res Methodol*. 2016;16(1):163. doi:10.1186/s12874-016-0267-3

17. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441. doi:10.1136/bmj.m441

18. Hilgers RA. Distribution-free confidence bounds for ROC curves. *Methods Inf Med*. 1991;30(2):96-101.

19. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177-188. doi:10.1016/0197-2456(86)90046-2

20. Altman DG, Bland JM. Standard deviations and standard errors. *BMJ*. 2005;331(7521):903. doi:10.1136/bmj.331.7521.903

21. Vickers AJ, Calster BV, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*. 2016;352:i6. doi:10.1136/bmj.i6

22. Calster BV, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol*. 2018;74(6):796-804. doi:10.1016/j.eururo.2018.08.038

23. Aubert O, Ursule-Dufait C, Brousse R, et al. Cell-free DNA for the detection of kidney allograft rejection. *Nat Med*. 2024;30(8):2320-2327. doi:10.1038/s41591-024-03087-3

24. Bloom RD, Bromberg JS, Poggio ED, et al. Cell-Free DNA and Active Rejection in Kidney Allografts. *J Am Soc Nephrol*. 2017;28(7):2221-2232. doi:10.1681/asn.2016091034

25. Bu L, Gupta G, Pai A, et al. Clinical outcomes from the Assessing Donor-derived cell-free DNA Monitoring Insights of kidney Allografts with Longitudinal surveillance (ADMIRAL) study. *Kidney Int*. 2022;101(4):793-803. doi:10.1016/j.kint.2021.11.034

26. Huang E, Sethi S, Peng A, et al. Early clinical experience using donor-derived cell-free DNA to detect rejection in kidney transplant recipients. *Am J Transplant*. 2019;19(6):1663-1670. doi:10.1111/ajt.15289

27. Park S, Guo K, Heilman RL, et al. Combining Blood Gene Expression and Cellfree DNA to Diagnose Subclinical Rejection in Kidney Transplant Recipients. *Clin J Am Soc Nephrol*. 2021;16(10):1539-1551. doi:10.2215/cjn.05530421

28. Sigdel TK, Archila FA, Constantin T, et al. Optimizing Detection of Kidney Transplant Injury by Assessment of Donor-Derived Cell-Free DNA via Massively Multiplex PCR. *J Clin Med*. 2018;8(1):19. doi:10.3390/jcm8010019

29. Bromberg JS, Brennan DC, Taber DJ, et al. dd-cfDNA Significantly Improves Rejection Yield in Kidney Transplant Biopsies. *Am J Transplant*. 2025;25(12):2529-2542. doi:10.1016/j.ajt.2025.07.2484

30. Hidalgo LG, Madill-Thomsen KS, Reeve J, et al. Improving the histologic detection of DSA-negative antibody-mediated rejection in kidney transplants. *Am J Transplant*. Published online 2025. doi:10.1016/j.ajt.2025.08.029

31. Bromberg JS, Bunnapradist S, Samaniego-Picota M, et al. Elevation of Donor-derived Cell-free DNA Before Biopsy-proven Rejection in Kidney Transplant. *Transplantation*. 2024;108(9):1994-2004. doi:10.1097/tp.0000000000005007

32. Tamargo CL, Kant S. Pathophysiology of Rejection in Kidney Transplantation. *J Clin Med*. 2023;12(12):4130. doi:10.3390/jcm12124130

33. Mayer KA, Budde K, Jilma B, Doberer K, Böhmig GA. Emerging drugs for antibody-mediated rejection after kidney transplantation: a focus on phase II & III trials. *Expert Opin Emerg Drugs*. 2022;27(2):151-167. doi:10.1080/14728214.2022.2091131

34. Mayer KA, Doberer K, Tillgren A, et al. Diagnostic value of donor-derived cell-free DNA to predict antibody-mediated rejection in donor-specific antibody-positive renal allograft recipients. *Transpl Int*. 2021;34(9):1689-1702. doi:10.1111/tri.13970

35. Akifova A, Budde K, Amann K, et al. Donor-derived cell-free DNA monitoring for early diagnosis of antibody-mediated rejection after kidney transplantation: a randomized trial. *Nephrol Dial Transplant*. 2024;40(7):1384-1395. doi:10.1093/ndt/gfae282

36. Naesens M, Wong G. Simply measuring what matters—toward clinical utility of donor-derived cell-free DNA testing? *Kidney Int*. 2025;107(5):799-805. doi:10.1016/j.kint.2025.01.039

37. Fagan TJ. Nomogram for Bayes's Theorem. *N Engl J Med*. 1975;293(5):257-257. doi:10.1056/nejm197507312930513

38. Wald N, Bestwick J. Is the area under an ROC curve a valid measure of the performance of a screening or diagnostic test? *J Méd Screen*. 2013;21(1):51-56. doi:10.1177/0969141313517497

39. Cook NR. Statistical Evaluation of Prognostic versus Diagnostic Models: Beyond the ROC Curve. *Clin Chem*. 2008;54(1):17-23. doi:10.1373/clinchem.2007.096529

40. ISO17511:2020 IO for Standardization. In vitro diagnostic medical devices — Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples. Accessed November 28, 2025. https://cdn.standards.iteh.ai/samples/69984/6f1630afed4349189ad6fc31b46b8444/ISO-17511-2020.pdf

**Tables**

**Table 1: Overview of patient and sample numbers by histopathology**

| | N Patients | Overall Sample Frequency | Sample Frequency in Biopsies | Group |
|---|---|---|---|---|
| ***ABMR*** | 60 | 13.5% (10.8–16.5) | 20.0% (16.0–24.3) | Rejection |
| ***TCMR*** | 22 | 5.0% (3.3–6.9) | 7.4% (4.9–10.3) | Rejection |
| ***Mixed Rej*** | 12 | 2.7% (1.5–4.2) | 4.0% (2.3–6.2) | Rejection |
| ***DSA-MVI*** | 12 | 2.7% (1.5–4.2) | 4.0% (2.3–6.2) | Rejection |
| ***Neg. Biopsy*** | 25 | 4.7% (3.1–6.7) | 8.4% (5.6–11.8) | Non-Rejection |
| ***CNI Tox*** | 21 | 4.7% (3.1–6.7) | 7.1% (4.7–9.9) | Non-Rejection |
| ***IFTA*** | 22 | 5.0% (3.3–6.9) | 7.4% (4.9–10.3) | Non-Rejection |
| ***ATI*** | 16 | 3.6% (2.2–5.3) | 5.4% (3.3–7.9) | Non-Rejection |
| ***GN*** | 20 | 4.5% (2.9–6.4) | 6.7% (4.4–9.5) | Non-Rejection |
| ***BKVAN**** | 45 | 10.2% (7.8–12.8) | 15.2% (11.6–19.0) | Non-Rejection |
| ***UTI*** | 6 | 1.4% (0.6–2.5) | 2.0% (0.8–3.6) | Non-Rejection |
| ***other*** | 7 | 1.6% (0.7–2.8) | 2.4% (1.1–4.1) | Non-Rejection |
| ***Clin. Stable*** | 86 (146) | 33.0% (29.0–36.8) | | Non-Rejection |
| ***Borderline*** | 29 | 6.5% (4.6–8.7) | 9.8% (6.9–13.0) | Excluded |
| ***Total*** | 383 (443) | | | |

Numbers in brackets are the 95$^{th}$ confidence intervals (CI). ABMR: antibody mediated rejection; TCMR: T-cell mediated rejection; Mixed Rej.: Mixed rejection, DSA-MVI: donor-specific antibodies (DSA) negative C4d negative microvascular inflammation; Borderline: Suspicious TCMR; Neg. Biopsy: unremarkable histology; CNI Tox: Calcineurin Inhibitor Toxicity; IFTA: Interstitial fibrosis/Tubular atrophy; ATI: Acute tubular injury; GN: Glomerulonephritis; BKVAN: BK-virus associated nephropathy *(including 32 patient samples with high viremia samples): UTI: Urinary tract infection; Others: Amyloidosis(2), Recurrent disease(4), Tubulointerstitial nephritis(1); Clin. Stable: Clinically stable. Numbers in brackets is number of samples.

**Table 2: Summary of statistical significances**

| Group | N | Median % | Median cp/mL | Median CM | U-Test vs. Controls | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Controls | 25 | 0.25 | 23.0 | -0.38 | % | cp/mL | CM-Score | | | |
| ABMR | 60 | 1.24 | 73.5 | 0.33 | 6.7E-09 | 1.82E-08 | 1.4E-10 | | | |
| TCMR | 22 | 0.55 | 47.7 | -0.06 | 4.7E-03 | 4.47E-02 | 1.4E-03 | | | |
| Mixed Rej | 12 | 0.92 | 83.5 | 0.34 | 2.5E-05 | 3.29E-05 | 1.6E-05 | | | |
| DSA- MVI | 12 | 1.63 | 100.0 | 0.38 | 1.9E-05 | 8.38E-05 | 1.4E-02 | U-Test vs.TCMR | | |
| Borderline | 29 | 0.48 | 31.8 | -0.22 | 0.013 | 0.069 | 0.014 | % | cp/mL | CM-Score |
| CNI Tox | 21 | 0.32 | 12.0 | -0.48 | 0.14 | 0.050 | 0.310 | 0.178 | 0.000 | 0.003 |
| IFTA | 22 | 0.29 | 16.0 | -0.44 | 0.27 | 0.407 | 0.391 | 0.021 | 0.006 | 0.009 |
| ATI | 16 | 0.23 | 13.0 | -0.46 | 0.14 | 0.426 | 0.235 | 0.002 | 0.008 | 0.003 |
| GN | 20 | 0.31 | 11.5 | -0.49 | 0.37 | 0.045 | 0.085 | 0.012 | 0.000 | 0.000 |
| BKVAN | 45 | 0.27 | 25.0 | -0.35 | 0.41 | 0.104 | 0.244 | 0.014 | 0.046 | 0.018 |
| UTI | 6 | 0.38 | 13.5 | -0.44 | *0.05* | *0.136* | *0.460* | *0.2508* | *0.0069* | *0.0323* |
| other | 7 | 0.30 | 18.0 | -0.36 | *0.49* | *0.419* | *0.384* | *0.0463* | *0.0207* | *0.0234* |

Statistical significances of cp/mL, percentage and CM-Score against biopsy proven normal (Controls) and against TCMR. Values for UTI and other are only shown for reference (low numbers).

**Table 3: Summary of the performance parameters of the CM-score, Percentage and cp/mL derived from the Validation Cohort**

| Variable | Group | No Rej. | Rej. | Specificity ($CI^{95}$) | Sensitivity ($CI^{95}$) | AUC ($CI^{95}$) | NPV ($CI^{95}$) | PPV ($CI^{95}$) |
|---|---|---|---|---|---|---|---|---|
| CM-Score | All Rej., No Rej. & Stable | 279 | 75 | 94% (91–96) | 72% (63–81) | 90% (83–96) | 91% (89–93) | 80% (72–86) |
| Percent | All Rej., No Rej. & Stable | 279 | 75 | 82% (79–87) | 77% (68–85) | 86% (77–93) | 92% (89–94) | 59% (53–66) |
| cp/mL | All Rej., No Rej. & Stable | 279 | 75 | 89% (85–92) | 72% (63–81) | 89% (82–95) | 90% (88–93) | 68% (61–75) |

Borderline rejection excluded; NPV and PPV were calculated at 25% prevalence

**Table 4: Clinical Interpretation guide for CM-Score**

| Pre↓\|CM→ | 0.90 | 0.80 | 0.70 | 0.60 | 0.50 | 0.40 | 0.30 | 0.20 | 0.10 | 0.00 | -0.10 | -0.20 | -0.30 | -0.40 | -0.50 | -0.60 | -0.70 | -0.80 | -0.90 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5% | 47% | 46% | 45% | 43% | 41% | 38% | 35% | 30% | 25% | 16% | 10.4% | 5.9% | 3.0% | 1.7% | 0.9% | 0.5% | 0.2% | 0.2% | 0.2% |
| 10% | 65% | 64% | 63% | 62% | 60% | 56% | 54% | 47% | 41% | 29% | 20% | 12% | 6.1% | 3.5% | 1.9% | 1.0% | 0.4% | 0.4% | 0.3% |
| 15% | 75% | 74% | 73% | 72% | 70% | 67% | 65% | 59% | 52% | 40% | 28% | 17% | 9.4% | 5.5% | 3.0% | 1.5% | 0.6% | 0.6% | 0.6% |
| 20% | 81% | 80% | 79% | 78% | 77% | 74% | 72% | 67% | 61% | 48% | 35% | 23% | 13% | 7.6% | 4.2% | 2.1% | 0.9% | 0.8% | 0.8% |
| 25% | 85% | 84% | 84% | 83% | 82% | 79% | 78% | 73% | 67% | 55% | 42% | 28% | 16% | 9.8% | 5.6% | 2.8% | 1.2% | 1.1% | 1.0% |
| 30% | 88% | 87% | 87% | 86% | 85% | 83% | 82% | 78% | 73% | 61% | 49% | 34% | 20% | 12% | 7.1% | 3.6% | 1.5% | 1.4% | 1.3% |
| 35% | 90% | 90% | 89% | 89% | 88% | 86% | 85% | 81% | 77% | 67% | 54% | 39% | 24% | 15% | 8.7% | 4.5% | 1.9% | 1.7% | 1.7% |
| 40% | 92% | 91% | 91% | 91% | 90% | 88% | 87% | 84% | 80% | 71% | 59% | 44% | 28% | 18% | 11% | 5.5% | 2.3% | 2.1% | 2.0% |
| 45% | 93% | 93% | 93% | 92% | 92% | 90% | 90% | 87% | 83% | 75% | 64% | 49% | 32% | 21% | 13% | 6.7% | 2.8% | 2.6% | 2.5% |
| 50% | 94% | 94% | 94% | 94% | 93% | 92% | 91% | 89% | 86% | 79% | 69% | 54% | 37% | 25% | 15% | 8.1% | 3.4% | 3.2% | 3.0% |

Combined post-test probability at each CM-Score-value (CM), where the likelihood ratio is defined as $LR^{+}$ multiplied by $LR^{-}$ to yield the likelihood at the given CM-Score-value. (LR+ and LR- are given in Supplementary Table S6).

**Figure Legends**

**Figure 1:** Boxplots of dd-cfDNA in percent (A), cp/mL (B), and the CM-Score (C) segregated by biopsy group

Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; data points are plotted as circles.

**Figure 2:** Comparison of the ROC curves obtained from the Validation group (All Biopsy-proven Rejections vs. All Non-Rejections plus Clinically Stable Patients) of the CM-Score, dd-cfDNA percentage, and dd-cfDNA in cp/mL.

The insert shows the gain by the CM-Score at the important threshold (depicted by the arrow vs % and cp/mL). The ROC curve for eGFR is shown for reference AUC: 0.62 (0.49–0.73).

**Figure 3:** Comparison of performance characteristics.

The PPV at 25% prevalence shows a significant improvement to both single values ($P<0.001$) with retained high NPV (not significantly different). This is due to the significant improvement of the specificity with the CM-Score.

**Figure 4:** Comparison of the combination of absolute and relative dd-cfDNA (CM-Score) with Data from 12 clinical cohorts from 11 published studies using either absolute or relative dd-cfDNA values.

Clinical performance characteristics values and standard deviations are given for published single values compared to the CM-Score in the validation group. Predictive values are calculated at 25% pre-test probability (prevalence) using the published sensitivities and specificities. Details are given in the Supplement.

Figure 1 (A)

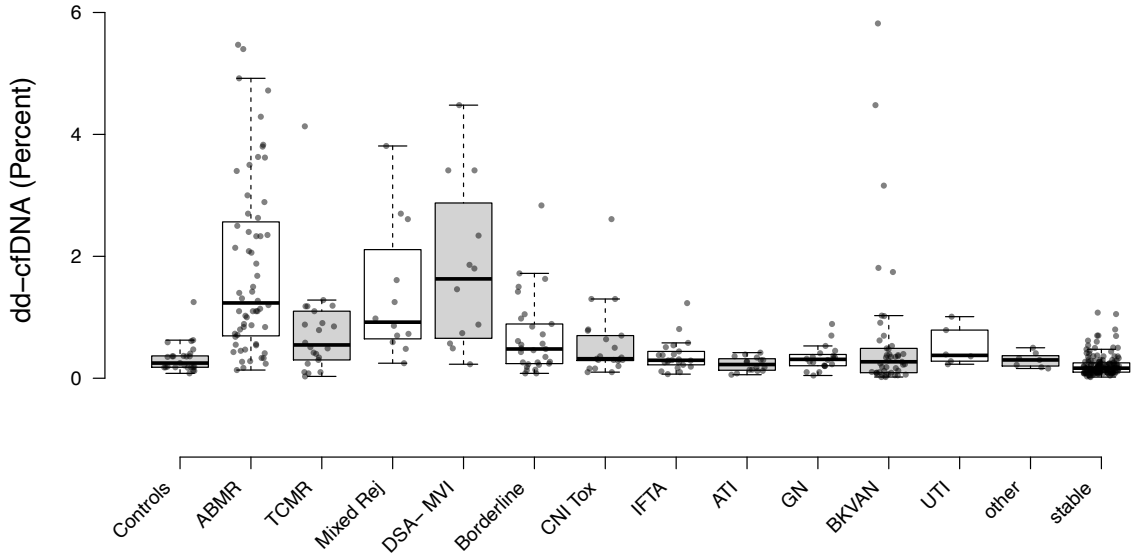


Figure 1(B)

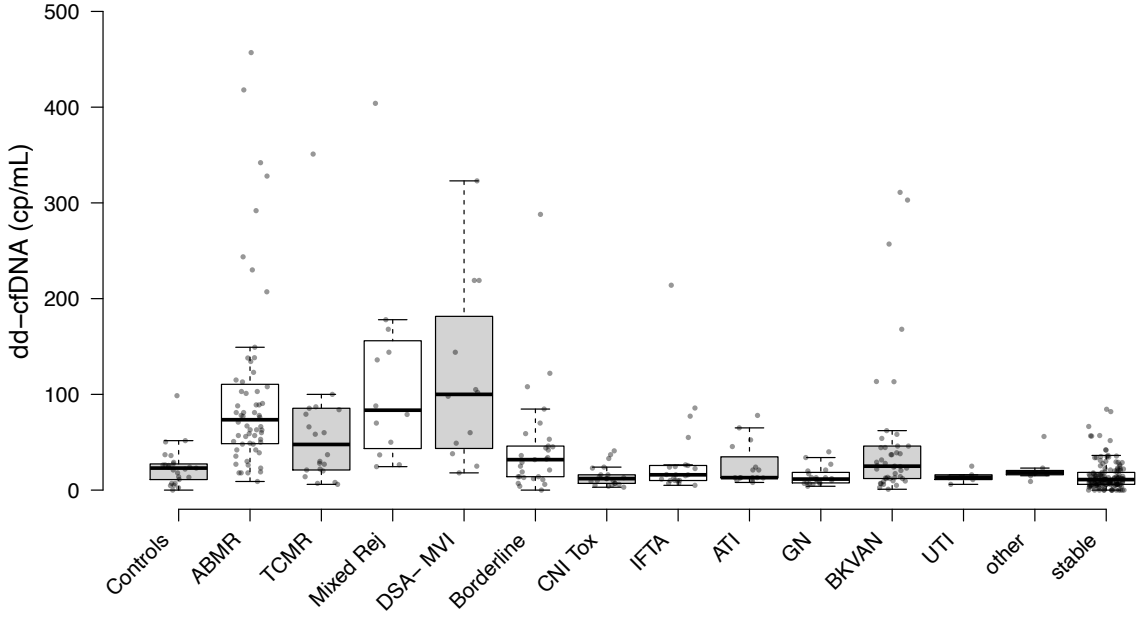


Figure 1 (C)

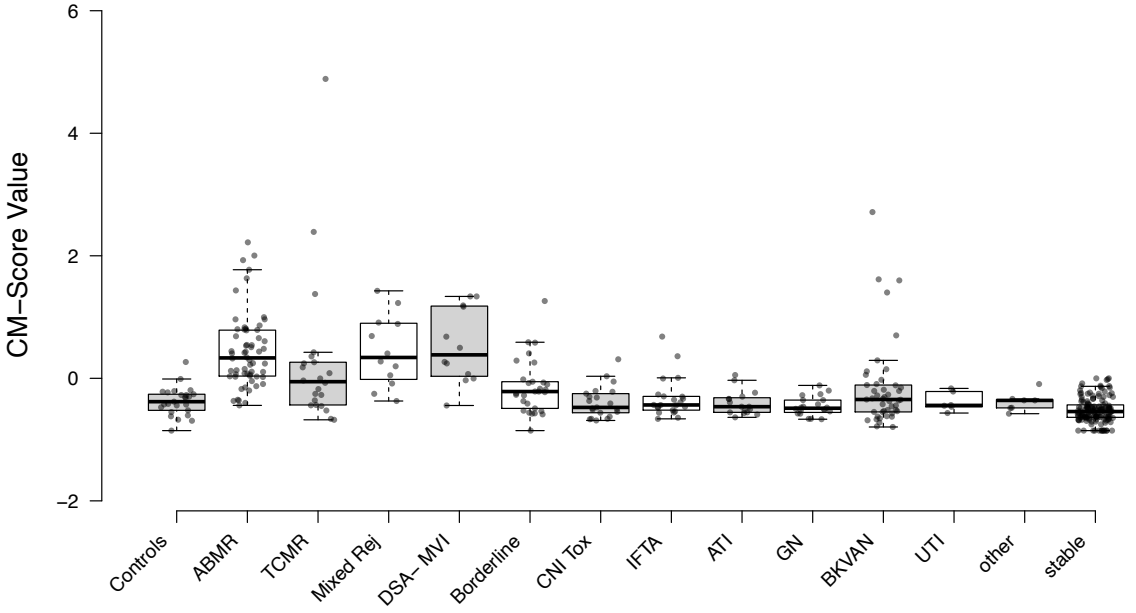

Figure 2: ROC curves for Percentage and cp/mL dd-cfDNA and CM-score from Validation cohort (n=279 non-rejections vs n=75 acute rejections)

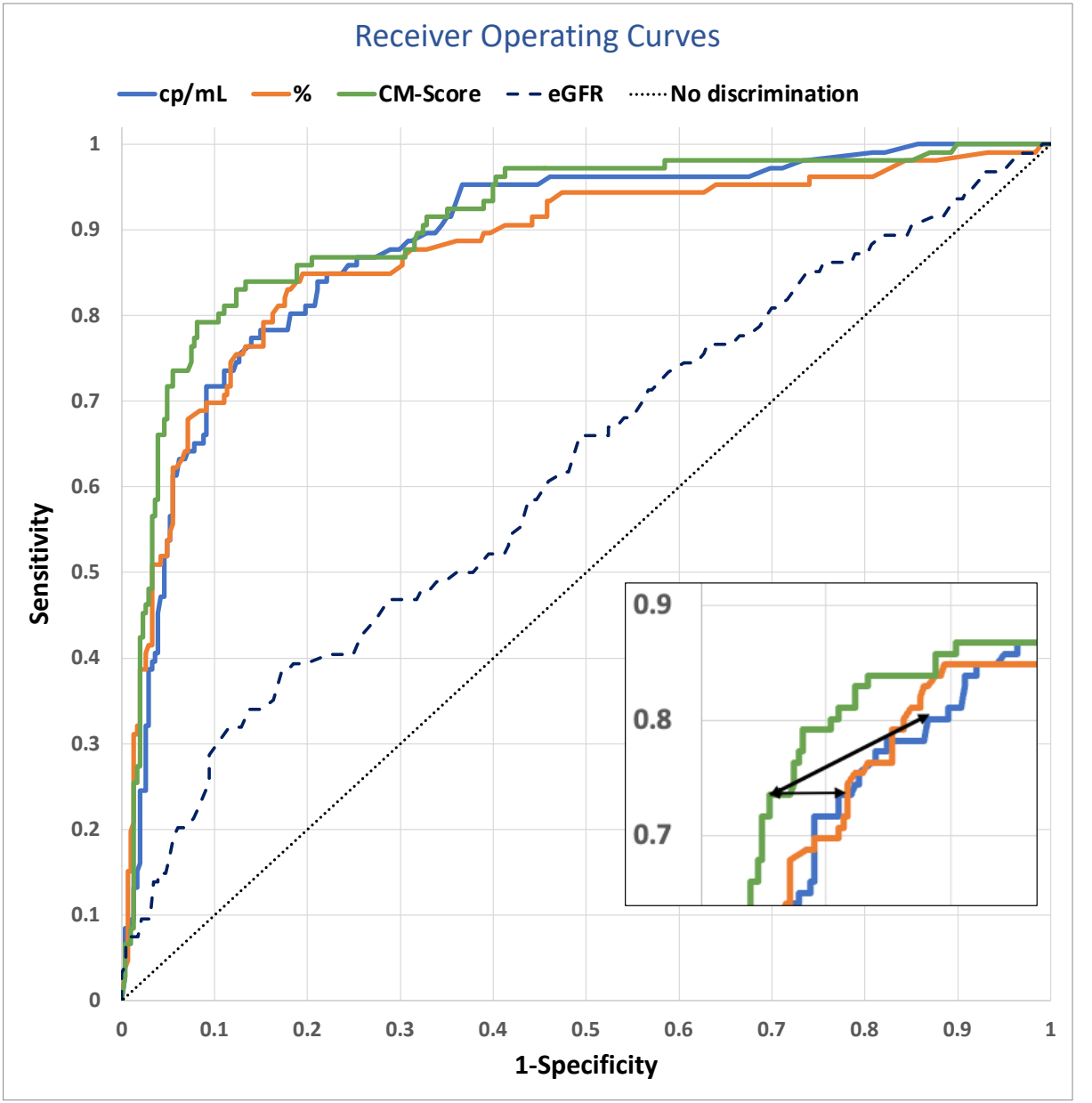

Figure 3: Comparison of Diagnostic Performance derived from the Validation Cohort

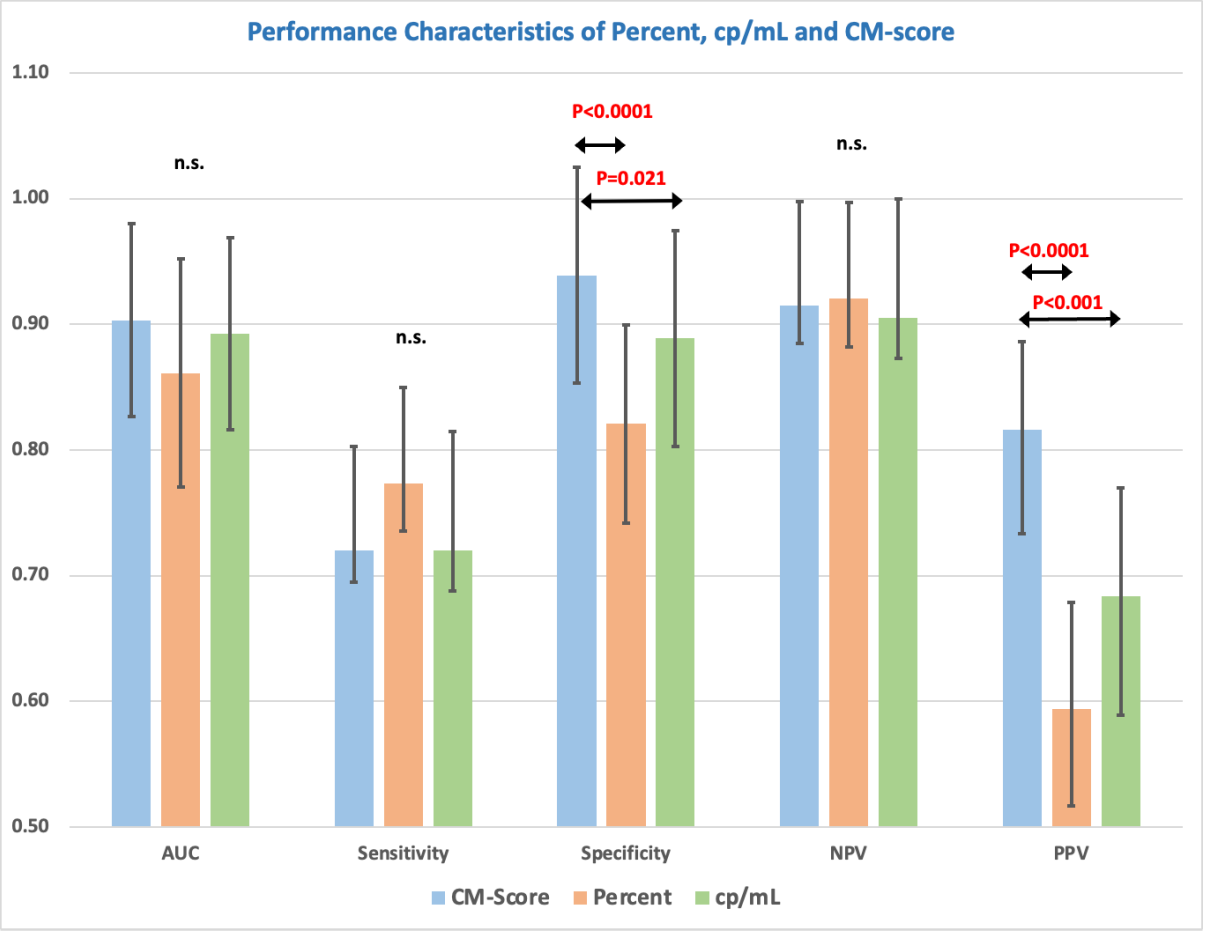
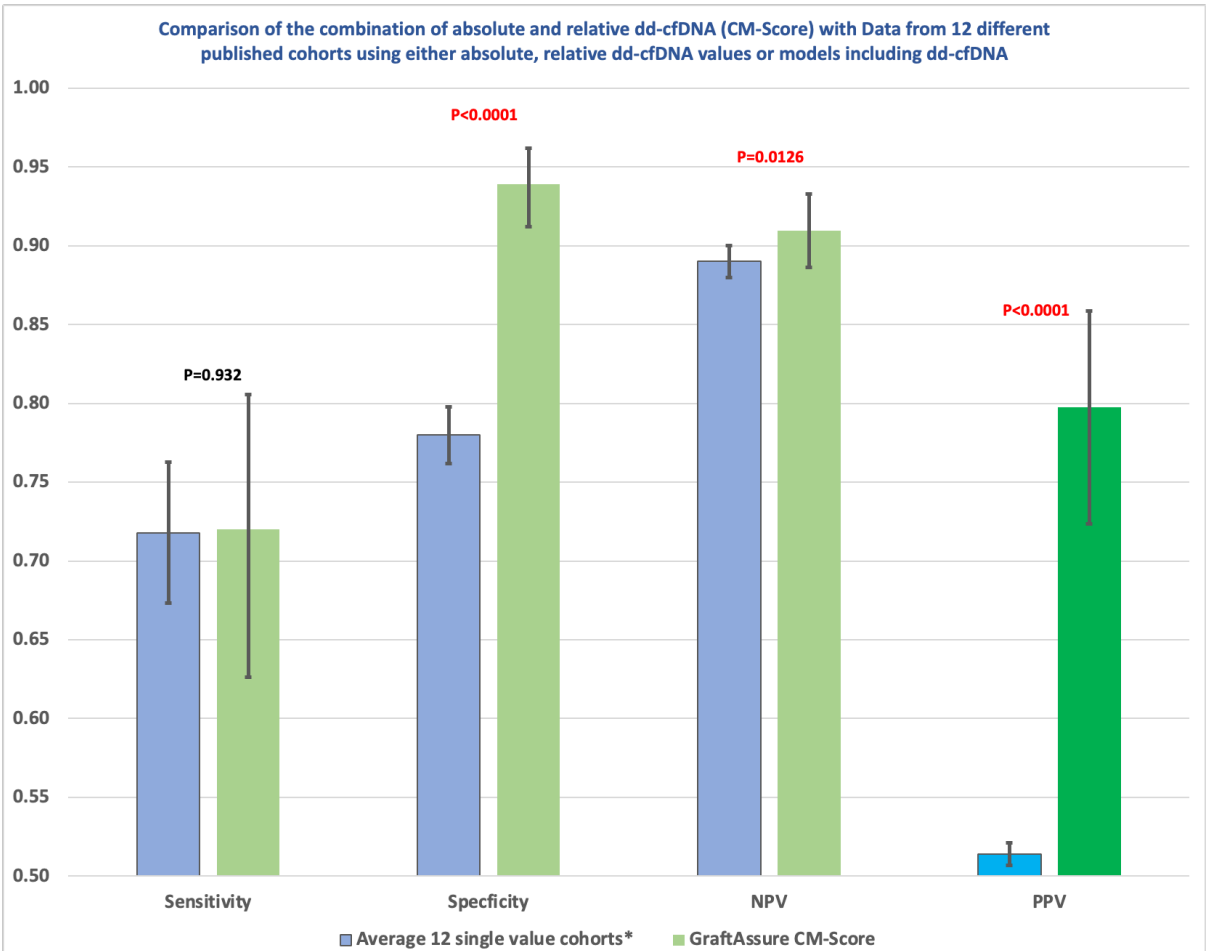
Figure 4 Comparison of values from the Validation cohort to published values