META-ANALYSIS

Xing et al.: A diagnostic marker for kidney-allograft rejection

Donor-derived cell-free DNA as a diagnostic marker for kidney-allograft rejection: A systematic review and meta-analysis

Yanbo Xing¹²#, Qiang Guo¹#, Cong Wang¹², Haoying Shi¹², Jiarui Zheng¹², Yijun Jia¹²,

Chengyong Li¹, and Chuan Hao¹*

¹Department of Urology, The Second Hospital of Shanxi Medical University, Taiyuan, Shanxi, China.

²Second Clinical Medical College, Shanxi Medical University, Taiyuan, Shanxi, China

*Corresponding author: Chuan Hao; E-mail: haochuan2003@sina.com

#These authors contributed equally to this manuscript.

DOI: https://doi.org/10.17305/bb.2024.10049

Submitted: 14 November 2024/ Accepted: 05 February 2024/ Published online: 20 February 2024

Conflicts of interest: Authors declare no conflicts of interest.

Funding: Authors received no specific funding for this work.
ABSTRACT

Donor-derived cell-free DNA (dd-cfDNA) has emerged as a promising biomarker for detecting graft rejection. This study aimed to evaluate the diagnostic accuracy and clinical value of applying it to kidney transplant rejection. Relevant literature on dd-cfDNA diagnostics in kidney transplant rejection was reviewed from PubMed, Embase, Cochrane Library, and Web of Science databases up to 2023. Data and study characteristics were extracted independently by two researchers, and disagreements were resolved through discussion. Diagnostic accuracy data for any rejection (AR) and antibody-mediated rejection (ABMR) were analyzed separately. Potential heterogeneity was analyzed by subgroup analysis or meta-regression. Funnel plots were used to clarify the presence or absence of publication bias. Nine publications provided data on dd-cfDNA accuracy in diagnosing patients with AR. The pooled sensitivity, specificity, and the area under the receiver operating characteristic (AUROC) curve with 95% confidence intervals (CI) were 0.59 (95% CI, 0.48-0.69), 0.83 (95% CI, 0.76-0.88), and 0.80 (95% CI, 0.76-0.83), respectively. Additionally, 12 studies focused on the diagnostic accuracy of dd-cfDNA for ABMR, showing pooled sensitivity, specificity, and the AUROC curve with 95% CI of 0.81 (95% CI, 0.72–0.88), 0.80 (95% CI, 0.73–0.86), and 0.87 (95% CI, 0.84-0.90), respectively. Study type, age group, and sample size contributed to heterogeneity. In summary, our findings indicate that while plasma dd-cfDNA accuracy in diagnosing patients with AR is limited by significant heterogeneity, it is a valuable biomarker for diagnosing ABMR.

KEYWORDS: Donor-derived cell-free DNA (dd-cfDNA), diagnosis, kidney, meta-analysis, rejection.
INTRODUCTION

Organ transplantation represents a pinnacle in surgical interventions for organ failure in patients. Among these, kidney transplantation has emerged as the optimal treatment for end-stage kidney disease, owing to its rapid and pioneering advancements, alongside its well-established technology and proficient management [1]. Kidney transplantation significantly enhances the patients’ survival rate and quality of life compared to dialysis [2]. However, transplant rejection is a major factor affecting the survival of solid transplanted organs [3]. Hence, it is crucial to promptly and precisely identify graft rejection and administer efficient therapy to prolong patient survival.

Currently, there are no reliable biomarkers for identifying graft rejection. Assessing transplant kidney rejection using traditional laboratory indices such as serum creatinine, urea nitrogen, and creatinine clearance offer limited value due to their low sensitivity and specificity [1]. Furthermore, a study has indicated that diagnosing rejection based on serum creatinine is often delayed, with approximately 50% of transplanted kidneys potentially suffering functional impairment by the time serum creatinine levels become abnormal [4]. While pathologic biopsy remains the preferred method for diagnosing rejection, and stands as the gold standard, it is invasive and carries potential complications [5,6].

Donor-derived cell-free DNA (dd-cfDNA) is a fragment of DNA released by cell death that exists in free form in the circulation and is not bound to cells. High levels of dd-cfDNA indicate graft damage and can be observed days or even weeks before the onset of acute rejection symptoms [1]. Several studies have reported the clinical validity of utilizing dd-cfDNA in kidney transplantation for the identification or exclusion of rejection or other graft injuries [7-
However, data from the study by Chang et al. [10] do not support its use as a biomarker of transplant rejection. Discrepancies between different studies exist, likely due to variations in research methodology, threshold setting, and study population. We conducted this meta-analysis to resolve existing controversies and determine dd-cfDNA’s role can as a diagnostic indicator for kidney transplant rejection.

MATERIALS AND METHODS

This systematic review was registered in International Prospective Register of Systematic Reviews (PROSPERO; CRD42023463407) and conducted following the PRISMA guidelines [11].

Search strategy

We systematically searched PubMed, Web of Science, Embase, Cochrane, and various other databases for studies evaluating the accuracy of dd-cfDNA in diagnosing rejection after kidney transplantation. The heading terms included “transplantation”, “dd-cfDNA” and “sensitivity or specificity”. Moreover, we assessed the bibliographies of selected research papers and contacted the authors when necessary. The search concluded on October 1, 2023.

Study selection

The included studies evaluated the diagnostic accuracy of blood dd-cfDNA in detecting graft rejection. Studies of urinary dd-cfDNA were not included. Eligible studies included cohort studies, case-control studies, and cross-sectional studies. Furthermore, the research provided comprehensive data for the construction of a 2x2 contingency table, including information on false positives, true positives, and negatives. We did not include animal tests, evaluations,
Data extraction

The data from all included studies were independently extracted by two investigators using the predefined data extraction form. Consultation with a third reviewer resolved any arising differences. Among the extracted information were the author’s name, publication year, study design, sample size, test method, cutoff thresholds, true positives (TPs), false negatives (FPs), false negative cases (FNs), true negative cases (TNs), as well as sensitivity and specificity values.

Quality assessment

Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was used to assess study quality [12]. Two authors independently evaluated each signaling question, assigning a score of “yes”, “no” or “unclear”. In case of disagreements, a third reviewer was involved for consensus.

Statistical analysis

The bivariate approach was employed to assess combined sensitivity, specificity, positive and negative likelihood ratios (PLR and NLR), and the diagnostic odds ratio (DOR). This analysis also produced the corresponding 95% confidence interval (CI) and generated the summary receiver operating characteristic curve (sROC).

We investigated the threshold phenomenon in this study using the Spearman correlation coefficient due to the varying cutoff values utilized in the included studies. The Chi-square test and $I^2$ test were employed to assess heterogeneity in the non-threshold effect. Significant heterogeneity was defined as $I^2 > 50\%$. Subgroup analyses and meta-regression were conducted
to explore sources of heterogeneity.

A Deeks' funnel plot was utilized to assess the potential presence of publication bias. Statistical analyses and meta-analyses were conducted using Stata version 15.1 (StataCorp, TX, USA).

RESULTS

We initially identified 526 articles from five databases. Following the elimination of duplicates and irrelevant studies based on title and abstract screening, a total of 83 articles were chosen for a comprehensive review of the full-text in order to extract pertinent information. After conducting a comprehensive analysis of the entire document, we incorporated a total of 22 articles, excluding reviews, case studies, and articles with unattainable or incomplete information (Figure 1) [7-10,13-30]. Additionally, Gielis et al.'s study was excluded due to its heavy reliance on unreliable sequido assay methods [31].

Characterization of the studies

The meta-analysis included the diagnostic criteria and distinctive features of every study, as displayed in Table 1. All studies were published within the last five years. Among these, 14 studies were conducted in the United States, 3 in China, 3 in Australia, 1 in Canada, and 1 in Germany. More than half of the studies were prospective in nature. The most widely employed method for measuring dd-cfDNA is Next-Generation Sequencing (NGS). In the majority of the included literature, AlloSure, developed and provided by CareDx, Inc., serves as a prominent NGS-based detection method for dd-cfDNA assessment. Additionally, ddPCR has also been successfully employed. The cutoff value of 1.00% was frequently employed across the majority of the studies. Diagnostic criteria for pathologic biopsies across all studies adhered to the Banff
Within the included literature, 9 articles are focused on identifying any rejection (AR), encompassing T cell-mediated rejection (TCMR), antibody-mediated rejection (ABMR), or mixed rejection (specific rejection types are detailed in Table S1). Additionally, 12 articles specifically assessed the identification of ABMR. Considering the distinct types of rejection, separate meta-analyses were conducted for AR and ABMR. Furthermore, a limited number of studies have individually explored the diagnostic performance of dd-cfDNA in distinguishing active rejection from acute rejection. However, due to the scarcity of literature on this specific aspect, we did not include them in the quantitative analysis.

**Quality assessment of included literature**

Out of the 22 papers analyzed, 11 studies demonstrated a low risk in terms of patient selection, while the remaining studies exhibited indication of a higher risk, primarily due to the inclusion of patients from non-randomized or non-consecutive sampling. Most studies lacked predetermined thresholds, leading to their classification as 'unclear risk' in the index test domain. None of the studies included in the reference standard domain were deemed 'high risk', as all of them utilized pathologic examination (kidney puncture biopsy) to determine whether a patient had experienced rejection. The majority of studies in the flow and timing domains were assessed as having a “low risk” for bias (Figure 2).

**Diagnostic accuracy of dd-cfDNA in AR**

Out of the studies included, 9 reported the diagnostic accuracy of dd-cfDNA for AR, with 1 specifically focusing on children. The meta-analysis showed that the pooled sensitivity and specificity were 0.59 (95% CI, 0.48-0.69) and 0.83 (95% CI, 0.76-0.88), correspondingly
(Figure 3A). The area under the curve (AUC) was 0.80 (Figure 3B), and the overall DOR was 7 (95% CI, 5-10). The pooled PLR was 3.5 (95% CI, 2.8-4.5), and the pooled NLR was 0.49 (95% CI, 0.40-0.61).

**Heterogeneity and subgroup analysis**

We found significant heterogeneity in sensitivity ($P < 0.01; I^2 = 81.8\%$) and specificity ($P < 0.01; I^2 = 90.8\%$) among the included studies. The Spearman correlation of 0.517 ($P = 0.15$) suggests that the threshold effect did not predominantly contribute to the heterogeneity.

We conducted subgroup analysis based on threshold levels. The subset that had a threshold of 1% exhibited greater combined specificity (0.85 compared to 0.66), while the subset with a threshold of 0.5% demonstrated higher sensitivity (0.78 compared to 0.57). Both subsets displayed similar PLR, NLR, DOR, and AUC values, as shown in Table S2.

Considering the differences between children and adults, we performed a subgroup analysis. In the adult group, our results indicate a combined sensitivity of 0.59 (95% CI, 0.46–0.70), a combined specificity of 0.83 (95% CI, 0.75–0.88), and an AUC of 0.79 (95% CI, 0.75–0.82) (Table S2). Due to the limited number of studies in the pediatric group, we refrained from conducting a quantitative analysis on it.

Upon reviewing the experimental designs in the literature, we observed that some subjects underwent protocol biopsies, while others underwent “clinically indicated” biopsies based on direct evidence of graft damage in the patient, such as elevated baseline serum creatinine and/or new-onset of proteinuria. The subgroup analyses showed that the protocol-contained group and the for-cause group had comparable pooled sensitivity and specificity when compared to the overall group (Table S2).
To delve deeper into the origins of heterogeneity, we performed meta-regression on the chosen covariates. The main covariates included dd-cfDNA detection methods, biopsy design, type of study, study size and age groups. The results showed that the study type \((P < 0.01)\), sample size \((P < 0.001)\), and age group \((P < 0.001)\) contributed to heterogeneity (Figure S1).

Ultimately, a Deeks’ funnel plot asymmetry test was executed to assess publication bias. Findings indicated that there is no publication bias \((P = 0.11)\) (Figure S4A).

**Diagnostic accuracy analysis of dd-cfDNA in ABMR**

The meta-analysis encompassed 12 diagnostic tests to evaluate the accuracy of diagnosing ABMR. Mayer's article reports on two different sets of data and is considered two separate studies. Forest plots illustrated the sensitivity and specificity in Figure 4A. The overall sensitivity was 0.81 (95% CI, 0.72–0.87), and specificity was 0.80 (95% CI, 0.73–0.85). The pooled estimates for PLR, NLR, and DOR were 4.0 (95% CI, 3.0–5.3), 0.24 (95% CI, 0.17–0.35), and 17 (95% CI, 10–28), respectively. The AUC value was 0.87, indicating a high diagnostic accuracy for ABMR using dd-cfDNA.

**Test of heterogeneity and subgroup analysis**

Variation among studies was observed regarding sensitivity \((P < 0.01; I^2 = 65.8\%)\) and specificity \((P < 0.01; I^2 = 79.7\%)\). The Spearman correlation of 0.098 \((P = 0.761)\) suggests that the threshold effect did not significantly contribute to the heterogeneity. Moreover, when considering only prospective studies, the combined sensitivity and specificity of the prospective study group were 0.76 (95% CI, 0.66–0.84) and 0.79 (95% CI, 0.74–0.83), respectively (Table S3). No heterogeneity was observed in sensitivity \((P = 0.23; I^2 = 28.0\%)\) and specificity \((P = 0.58; I^2 = 0)\) among prospective studies (Figure S2).
When limiting the analysis to literature on the for-cause group, the summary estimate showed a sensitivity of 0.83 (95% CI, 0.74–0.89) and a specificity of 0.80 (95% CI, 0.72–0.86), which were consistent with the overall findings (Table S3).

Similarly, our meta-regression results reveal that study type (P < 0.05), testing method (P < 0.01), and sample size (P < 0.001) significantly contribute to the observed heterogeneity (Figure S3). Additionally, Deeks' funnel plot analysis showed no substantial indication of publication bias (Figure S4B).

**DISCUSSION**

Over the past five years, several organizations have assessed the potential of dd-cfDNA as a biomarker for graft rejection in transplants, including the heart, liver, and kidney. The results of our meta-analysis suggest that dd-cfDNA can be a helpful marker for identifying ABMR. However, further validation is still required to assess its accuracy in recognizing AR. In addition, a few studies have investigated the diagnostic performance of dd-cfDNA in distinguishing acute rejection from active rejection. However, due to the limited number of relevant studies, a Meta-analysis was not conducted.

In comparison to dd-cfDNA, serum creatinine demonstrates neither specificity nor sensitivity in recognizing allograft injury. A previous study indicated that at the time of biopsy, serum creatinine failed to discriminate active rejection from no active rejection. The ROC curve for creatinine in distinguishing active rejection had an AUC of 0.54, indicating that at any cutoff level for creatinine, there were as many false positives as true positives[7]. In contrast, our meta-analysis found that dd-cfDNA exhibited moderate sensitivity and high specificity.
Moreover, it demonstrated a high level of diagnostic accuracy in predicting patients with AR, with a combined AUC of 0.80.

Currently, there is a lack of direct comparative research on the diagnostic performance of dd-cfDNA testing methods applied to pediatric and adult transplant groups. In our study, when the dd-cfDNA threshold was set at 1% to identify adult AR, the AUC was 0.79, with a specificity of 83% and a sensitivity of 59%. However, due to limited data on pediatric patients, we are unable to conduct quantitative analysis in this regard. Findings from Ranch et al.’s study in the pediatric group indicated that when dd-cfDNA>1%, the sensitivity for diagnosing AR is 66%, with a specificity of 96%[24].

Puliyanda posited that there are differences in the sensitivity and specificity of dd-cfDNA between pediatric and adult patients, and these variations may be attributed to the difference in size between the graft and the recipient's body[14]. The lower body weight of children, resulting in reduced background DNA, is also a contributing factor[14]. Additionally, a study highlighted that, due to the more significant kidney reserve provided by adult-sized grafts, serum creatinine in pediatric recipients remained within the normal range even during rejection[33]. In this context, the role of dd-cfDNA becomes particularly crucial in identifying rejection in children, given its high specificity, thereby significantly reducing the need for unnecessary biopsies.

The diversity of thresholds reflects the fact that centers are still in the exploratory phase, and currently, there are no uniform laboratory standards. When we used the threshold published in most studies (1%) as a criterion for the diagnosis of rejection, our study yielded an AUC of 0.81 and a diagnostic specificity of 0.85. In contrast, the sensitivity was lower, at 0.57, and was
prone to false-negative results. Notably, the subgroup with a threshold = 0.5% exhibited a higher sensitivity of 0.78, although the specificity was reduced to 0.66. Hence, additional research is required in the coming years to ascertain the most suitable threshold for dd-cfDNA. Although we performed separate subgroup analyses based on threshold and reason for biopsy, the results of these analyses showed small deviations in effect sizes for diagnostic metrics compared with the overall results. To investigate the origins of variability more deeply, we conducted a meta-regression, revealing that both study type and sample size contributed to the observed heterogeneity. Yet, due to limited validated data, further stratification of the included patients in the study was unattainable.

In this study, we observed that the combined sensitivity of dd-cfDNA in detecting ABMR patients is greater compared to its sensitivity in detecting patients with AR, despite having the same specificity. This pattern remained consistent when excluding studies with higher heterogeneity, specifically retrospective studies. The statistical difference in sensitivity could potentially be attributed to the inclusion of patients with TCMR in the study. Specifically, when ABMR occurs, it involves interactions between antibodies and the endothelium of allograft vessels, leading to the release of dd-cfDNA directly into the bloodstream. In contrast, the concentration of dd-cfDNA is correlated with the grade of TCMR, where borderline and grade 1A TCMR involve fewer vessels, resulting in mostly normal concentrations of dd-cfDNA. A further reason for false negative TCMR results may be attributed to the use of relatively long amplicons (100-130 bp) in some NGS methods[3,7,18]. In TCMR there may be a more extensive dd-cfDNA fragment degradation. Dauber et al.[15] demonstrated significantly higher dd-cfDNA test results when using smaller-size amplicons. A range of 66-103 bp has recently
been recommended[34]. However, due to limited availability of raw data, we were unable to conduct a separate quantitative analysis of TCMR patients. Further research is needed in the future to explore the diagnostic performance of NGS testing for detecting TCMR.

In an effort to further enhance the diagnostic performance of dd-cfDNA, researchers have made various attempts. Mayer's study suggests that, in comparison to serum creatinine, dd-cfDNA can be utilized for the early detection of ABMR in kidney transplant patients who are positive for donor-specific antibodies (DSA). Furthermore, the combined detection of dd-cfDNA and DSA enhances diagnostic performance[19]. Additionally, a few articles have explored the potential of quantifying absolute dd-cfDNA (copies/ml) as an alternative method. Two prior investigations yielded conflicting findings regarding the absolute values of dd-cfDNA[21,26]. Oellerich and colleagues highlighted the superiority of absolute dd-cfDNA quantification over dd-cfDNA fraction for distinguishing kidney allograft rejection. They asserted that absolute dd-cfDNA values remain unaffected by alterations in circulating recipient DNA, rendering them more dependable[21]. Nevertheless, these findings were not substantiated in the study conducted by Whitlam et al[26]. This discrepancy seems to be due to methodological differences, as Whitlam did not account for variation in cfDNA extraction efficiency and ddPCR amplification. These two variables account for 44-55% of the underestimation of the number of DNA copies per ml. In a study by Bunnapradist et al. [13], it was demonstrated that incorporating both the quantity and percentage of dd-cfDNA enhanced the test's sensitivity in kidney transplant patients while maintaining a high level of specificity compared to using only the dd-cfDNA percentage. Nevertheless, further studies are necessary to confirm the clinical significance of absolute dd-cfDNA levels due to the limited sample size [35].
There are some limitations to our study. Firstly, despite the implementation of stringent inclusion criteria, heterogeneity could not be entirely eliminated. Secondly, due to analytical challenges associated with some NGS methods for TCMR detection, the overall diagnostic value of dd-cfDNA seems not to be correctly represented by any AR in the presented systematic review. Thirdly, after a thorough review of the included literature, we found that, due to limited available information, specific diagnostic accuracy data for dd-cfDNA in distinguishing TCMR could not be obtained.

In conclusion, although the accuracy of Plasma dd-cfDNA in diagnosing patients with AR is not very reliable due to the observed large heterogeneity, our findings affirm the utility of Plasma dd-cfDNA levels as a diagnostic indicator for ABMR. However, there remains a necessity for future multicenter, prospective studies to investigate the ideal threshold of dd-cfDNA% for diagnosing rejection. Moreover, exploring the potential simultaneous use of dd-cfDNA quantification to enhance diagnostic accuracy is warranted.
REFERENCES


Table 1. Summary characteristics of included studies.

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Study type</th>
<th>Country</th>
<th>Sample type</th>
<th>Test method</th>
<th>Reference standard</th>
<th>Rejection type</th>
<th>Threshold (%)</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>SEN</th>
<th>SPE</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheng</td>
<td>2022</td>
<td>Retrospective</td>
<td>China</td>
<td>Plasma</td>
<td>ND</td>
<td>Banff, 2017</td>
<td>ABMR</td>
<td>1.11</td>
<td>16</td>
<td>17</td>
<td>2</td>
<td>25</td>
<td>88.9</td>
<td>59.5</td>
<td>0.76</td>
</tr>
<tr>
<td>Zhang</td>
<td>2020</td>
<td>Prospective</td>
<td>China</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2015</td>
<td>ABMR</td>
<td>1.00</td>
<td>16</td>
<td>5</td>
<td>2</td>
<td>14</td>
<td>88.9</td>
<td>73.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Jordan</td>
<td>2018</td>
<td>Prospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2013</td>
<td>ABMR</td>
<td>1.00</td>
<td>13</td>
<td>13</td>
<td>3</td>
<td>61</td>
<td>81.3</td>
<td>82.4</td>
<td>0.86</td>
</tr>
<tr>
<td>Dauber</td>
<td>2020</td>
<td>Prospective</td>
<td>Austria</td>
<td>Plasma</td>
<td>qPCR</td>
<td>Banff, 2015</td>
<td>Acute rejection</td>
<td>2.70</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>17</td>
<td>87.5</td>
<td>81.0</td>
<td>0.84</td>
</tr>
<tr>
<td>Huang</td>
<td>2019</td>
<td>Retrospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2013</td>
<td>Any rejection</td>
<td>1.00</td>
<td>23</td>
<td>8</td>
<td>11</td>
<td>21</td>
<td>67.6</td>
<td>72.4</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ABMR</td>
<td>1.00</td>
<td>20</td>
<td>11</td>
<td>4</td>
<td>28</td>
<td>83.3</td>
<td>71.8</td>
<td>ND</td>
</tr>
<tr>
<td>Sigdel</td>
<td>2018</td>
<td>Retrospective</td>
<td>United States</td>
<td>Plasma</td>
<td>mmPCR-NGS</td>
<td>Banff, 2017</td>
<td>Active rejection</td>
<td>1.00</td>
<td>33</td>
<td>50</td>
<td>5</td>
<td>129</td>
<td>86.8</td>
<td>72.1</td>
<td>0.87</td>
</tr>
<tr>
<td>Oellerich</td>
<td>2019</td>
<td>Prospective</td>
<td>Germany</td>
<td>Plasma</td>
<td>ddPCR</td>
<td>Banff, 2017</td>
<td>Any rejection</td>
<td>0.43</td>
<td>16</td>
<td>122</td>
<td>6</td>
<td>273</td>
<td>72.7</td>
<td>69.1</td>
<td>0.73</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Study Design</td>
<td>Country</td>
<td>Sample Type</td>
<td>Analyte</td>
<td>Location</td>
<td>Assay</td>
<td>T0</td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td>T4</td>
<td>T5</td>
<td>T6</td>
<td>T7</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>--------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
<td>----------</td>
<td>-------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Whitlam</td>
<td>2019</td>
<td>Prospective</td>
<td>Australia</td>
<td>Plasma</td>
<td>ddPCR</td>
<td>Banff, 2013</td>
<td>ABMR</td>
<td>0.75</td>
<td>11</td>
<td>12</td>
<td>2</td>
<td>36</td>
<td>84.6</td>
<td>75.0</td>
<td>0.91</td>
</tr>
<tr>
<td>Bloom</td>
<td>2017</td>
<td>Prospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2013</td>
<td>Active rejection</td>
<td>1.00</td>
<td>16</td>
<td>10</td>
<td>11</td>
<td>70</td>
<td>59.3</td>
<td>87.5</td>
<td>0.74</td>
</tr>
<tr>
<td>Bunnapradist</td>
<td>2021</td>
<td>ND</td>
<td>United States</td>
<td>Plasma</td>
<td>mmPCR-NGS</td>
<td>Banff, 2017</td>
<td>Active rejection</td>
<td>1.00</td>
<td>13</td>
<td>15</td>
<td>3</td>
<td>76</td>
<td>81.3</td>
<td>83.5</td>
<td>0.87</td>
</tr>
<tr>
<td>Park</td>
<td>2021</td>
<td>Prospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2019</td>
<td>Acute rejection</td>
<td>1.00</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>29</td>
<td>77.8</td>
<td>90.6</td>
<td>ND</td>
</tr>
<tr>
<td>Mayer (1)</td>
<td>2021</td>
<td>Retrospective</td>
<td>Austria</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2017</td>
<td>ABMR</td>
<td>0.78</td>
<td>20</td>
<td>4</td>
<td>5</td>
<td>16</td>
<td>80.0</td>
<td>80.0</td>
<td>0.89</td>
</tr>
<tr>
<td>Mayer (2)</td>
<td>2021</td>
<td>Retrospective</td>
<td>Austria</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2017</td>
<td>ABMR</td>
<td>0.62</td>
<td>15</td>
<td>6</td>
<td>2</td>
<td>7</td>
<td>88.2</td>
<td>53.8</td>
<td>0.69</td>
</tr>
<tr>
<td>Halloran</td>
<td>2022</td>
<td>Prospective</td>
<td>Canada</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2019</td>
<td>Any rejection</td>
<td>1.00</td>
<td>40</td>
<td>20</td>
<td>14</td>
<td>75</td>
<td>74.1</td>
<td>78.9</td>
<td>ND</td>
</tr>
<tr>
<td>Puliyanda</td>
<td>2022</td>
<td>Prospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2017</td>
<td>Acute rejection</td>
<td>1.00</td>
<td>18</td>
<td>0</td>
<td>3</td>
<td>27</td>
<td>86.0</td>
<td>100.0</td>
<td>0.996</td>
</tr>
<tr>
<td>Obrișcă</td>
<td>2022</td>
<td>Retrospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff, 2017</td>
<td>Any rejection</td>
<td>1.00</td>
<td>19</td>
<td>2</td>
<td>11</td>
<td>22</td>
<td>63.3</td>
<td>91.7</td>
<td>ND</td>
</tr>
<tr>
<td>Chang</td>
<td>2022</td>
<td>Retrospective</td>
<td>United States</td>
<td>Plasma</td>
<td>ND</td>
<td>Banff, 2019</td>
<td>ABMR</td>
<td>1.00</td>
<td>16</td>
<td>9</td>
<td>51</td>
<td>160</td>
<td>24.0</td>
<td>95.0</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ABMR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A-ABMR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A-Any rejection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A-ABMR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Last Name</td>
<td>Year</td>
<td>Study Type</td>
<td>Country</td>
<td>Panel Type</td>
<td>Criteria</td>
<td>ABMR</td>
<td>SEN</td>
<td>SPE</td>
<td>AUC</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>------------</td>
<td>---------</td>
<td>------------</td>
<td>----------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cheng</td>
<td>2021</td>
<td>Retrospective</td>
<td>China</td>
<td>Plasma</td>
<td>ND</td>
<td>Banff, 2017</td>
<td>0.96</td>
<td>19</td>
<td>1</td>
<td>2</td>
<td>28</td>
<td>90.5</td>
<td>96.6</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Ranch</td>
<td>2023</td>
<td>Prospective</td>
<td>United States</td>
<td>Plasma</td>
<td>ND</td>
<td>Banff criteria</td>
<td>Any rejection</td>
<td>1.00</td>
<td>23</td>
<td>1</td>
<td>12</td>
<td>25</td>
<td>65.7</td>
<td>96.2</td>
<td>ND</td>
</tr>
<tr>
<td>Dandamudi</td>
<td>2022</td>
<td>Prospective</td>
<td>United States</td>
<td>Plasma</td>
<td>NGS</td>
<td>Banff criteria</td>
<td>Acute rejection</td>
<td>1.00</td>
<td>9</td>
<td>3</td>
<td>18</td>
<td>79</td>
<td>33.3</td>
<td>96.3</td>
<td>0.82</td>
</tr>
<tr>
<td>Rizvi</td>
<td>2023</td>
<td>Prospective</td>
<td>United States</td>
<td>ND</td>
<td>ND</td>
<td>Banff, 2017</td>
<td>Any rejection</td>
<td>1.00</td>
<td>32</td>
<td>8</td>
<td>20</td>
<td>29</td>
<td>61.5</td>
<td>78.4</td>
<td>ND</td>
</tr>
<tr>
<td>Gupta</td>
<td>2022</td>
<td>Prospective</td>
<td>United States</td>
<td>ND</td>
<td>ND</td>
<td>Banff, 2017</td>
<td>Any rejection</td>
<td>1.00</td>
<td>20</td>
<td>8</td>
<td>8</td>
<td>29</td>
<td>71.4</td>
<td>78.4</td>
<td>ND</td>
</tr>
<tr>
<td>Bu</td>
<td>2022</td>
<td>Prospective</td>
<td>United States</td>
<td>ND</td>
<td>NGS</td>
<td>Banff, 2019</td>
<td>Any rejection</td>
<td>1.00</td>
<td>65</td>
<td>19</td>
<td>48</td>
<td>87</td>
<td>57.5</td>
<td>82.1</td>
<td>ND</td>
</tr>
</tbody>
</table>

TP: True positive; FP: False positive; FN: False negative; TN: True negative; SEN: Sensitivity; SPE: Specificity; AUC: Area under the curve; ABMR: Antibody-mediated rejection; ddPCR: Digital droplet polymerase chain reaction; NGS: Next-generation sequencing; qPCR: Quantitative polymerase chain reaction; ND: No data.
Figure 1. The flow diagram of the search and selection process.
Figure 2. Summary of the risk of bias of the included studies. QUADAS-2: Quality assessment of the diagnostic accuracy studies-2.

Figure 3. Summary of the diagnostic accuracy of dd-cfDNA for diagnosis of any rejection. (A) Pooled sensitivity and specificity; (B) Summary receiver operating characteristic curve.
**Figure 4.** Summary of the diagnostic accuracy of dd-cfDNA for diagnosis of ABMR. (A) Pooled sensitivity and specificity; (B) Summary receiver operating characteristic (sROC) curve.

**SUPPLEMENTAL DATA**

Supplementary data can be found at the following link: