Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection

Author 1

August 15, 2025

Abstract

The ability to monitor kidney allografts is limited by traditional forms of assessment like invasive biopsy, and also delayed biomarkers, such as serum creatinine (Rizvi et al., 2023a). This systematic review was performed to assess the clinical usefulness of donor-derived cell-free DNA (dd-cfDNA) as a non-invasive biomarker of allograft injury. An extensive search identified 83 studies, each assessed for methodological quality. Results indicate a rapidly maturing evidence base, including an increase in publications and a corresponding increase in sample size over time, reflected most notably in a landmark study on nearly 3,000 patients (Aubert et al., n.d.). The literature was generally of good quality, as most studies scored 3 or more out of 5 criteria created by the author. A key finding of the review was the consistent and generally good diagnostic accuracy of dd-cfDNA, with most studies indicating an Area Under the Curve (AUC) greater than 0.80 (Mantios et al., 2023). There is however, a stark geographical disparity in the inclusion of diversified ancestry; the available evidence for Black and Hispanic ancestry is largely based on North American evidence (Bu et al., 2022). This systematic review provides confirmatory evidence that dd-cfDNA is a strong and usable non-invasive biomarker (Kim et al., 2024a), but requires future work with a particular focus on the design of well-constructed large, multicenter studies, focusing on diversity to ensure equitable and optimal application of this promising tool in clinical practice (Botella et al., 2024a).

Introduction

Kidney transplantation is a vital therapy for patients with end-stage kidney disease, providing a better overall longevity and improved quality of life compared to long-term dialysis (Dreige et al., 2022), which is the non-surgical alternative to

transplantation. Despite its advantages and evidence for superior long-term outcomes, transplantation faces the ongoing challenge of immune-mediated rejection, which remains a prominent and often principal cause of patient graft-loss (Lakhani et al., 2021). The challenge is exacerbated by a poor risk management framework that relies on conventional surveillance practices that often are complicated and inaccurate to true quantitative values (Wei et al., 2024). Traditionally, markers for the identification and management of rejection have been insufficient due to the slow response of these traditional markers to tissue injury (Rizvi et al., 2023a). In short, these significant barriers represent the unfulfilled need space, or unmet medical need, in transplant medicine, because they do not allow clinicians to intervene prior to true tissue/organ damage (Nie et al., 2025).

In recent years, a new solution has been developed to solve these deficiencies: the measurement of dd-cfDNA (Zhang et al., 2020a), small pieces of DNA from the donor that are released from the transplanted organ into the circulation, especially when the organ is activated, or inflamed, or injured through histocompatibility, or rejection (Zhang et al., 2020b). When a person has an organ transplant, they also are a recipient of a donor's genetic information at the time of the transplant (Kumar et al., 2023). Quantifying the amount of dd-cfDNA in a blood sample can now provide a direct, non-invasive and organ-specific signal of graft injury (Mayer et al., 2021). This has been depicted as a pioneering biomarker with much potential to detect graft damage early (Botella et al., 2024a).

Though transplant surgeons are currently utilizing dd-cfDNA in clinical practice, the development of the literature is so fast-paced that a comprehensive review of the evidence is needed (Chen et al., 2022). It is no longer sufficient to say that dd-cfDNA is "effective" and the key question to focus on is how effective it is in specific clinical contexts. The value of this systematic review is in synthesizing all studies done to-date in a rigorous manner, meeting the inclusion criteria, which will allow us grasp a clear understanding on the subject and summarize the topic for others (Sharma et al., 2022); it is guided by the specific hypothesis that donor-derived cell-free DNA (dd-cfDNA) can result in increased diagnostic accuracy in detecting transplant rejection and graft injury versus current standard of care biomarkers, through earlier and non-invasive methods . The focus of this review is to (One) specify and quantify overall diagnostic accuracy of dd-cfDNA for various types of rejection, (Two) create and analyze data visualizations to synthesize

the characteristics and findings of the included studies, (Three) identify trends in performance and address key knowledge gaps, such as the biomarker's utility in diverse, underrepresented patient populations (Benning et al., 2022).

The findings of this systematic review confirm that dd-cfDNA is a useful measure of allograft health (Kim et al., 2024a) and that evidence confirms that dd-cfDNA is a useful biomarker for early allograft injury (Tian et al., n.d.), and that dd-cfDNA can be used as a non-invasive biomarker (Nie et al., 2025). Moreover, the data show that dd-cfDNA has an indication of prognostic value for long-term allograft outcomes (Loupy et al., 2024). However, this systematic review also identified that some of the evidence base is derived from single-center cohorts or those with limited ethnic diversity (Nguyen et al., 2025). The most significant findings of this review include the quantification of dd-cfDNA's high diagnostic accuracy, with AUC values consistently exceeding 0.80 (Rizvi et al., 2023a), and its ability to detect rejection months before clinical presentation, such as reducing the time to AMR diagnosis to just 2.8 months (Akifova et al., 2025). This review revealed the biomarker's significant prognostic value, with early elevations of dd-cfDNA (≥1%) being associated with a more than five-fold greater rate of future eGFR—estimated Glomerular Filtration Rate, a key measure of how well the kidneys are filtering waste from the blood decline,—(21.4% vs. 4.1%) (Sawinski et al., 2021a). Going forward, it is recommended that future research focuses on head-to-head comparisons of commercial assays to standardize protocols and on conducting large, prospective trials in more diverse patient populations to ensure the equitable and optimal application of this biomarker in clinical practice.

Background

The transplantation field is actively looking for new, non-invasive, and accurate ways to monitor graft health, given the many challenges with established surveillance modalities (Rizvi et al., 2023b). The gold standard now, kidney biopsy, is invasive, with procedural and other risk and discomfort for the patient. For the transplant physician, biopsy has limitations as it is a reactive measure that does not suit the frequent monitoring needed to prevent damage (Wei et al., 2024). Serum creatinine, the most widely used non-invasive biomarker, is a surrogate, vague and non-specific biomarker. Thus, when creatinine rises, there can be a significant degree of already irreversible graft damage (Parajuli et al., 2024a).

Recent advancements have uncovered the study of donor-derived cell-free DNA (dd-cfDNA) as an alternative solution that directly counters these flaws by offering a non-invasive and accelerated method for assessing transplant status, potentially transforming post-transplant care (Botella et al., 2024a). The following results will synthesize the evidence to quantify the biomarker's accuracy and its timelines for early detection, thereby assessing its potential to become a new standard in post-transplant surveillance (Bu et al., 2022).

Kidney Transplantation and Rejection

Organ transplantation is a life-saving medical procedure by which a surgically incompetent organ from the recipient is removed and surgically replaced with a healthy organ obtained from a healthy donor (Orandi et al., 2016). Organ transplantation is important because it improves the quality of life for an organ recipient and increases life expectancy (Montgomery & Gelb, n.d.). Kidney transplantation is the most common organ transplantation because its underlying disease, end-stage renal disease (ESRD), is also one of the most common diseases (affecting >800,000 persons in the United States) (*Kidney Disease Statistics for the United States - NIDDK*, n.d.). Conditions like diabetes and high blood pressure are the leading causes of kidney failure, affecting millions and contributing to the high demand for transplants (*Quick Kidney Disease Facts and Stats | American Kidney Fund*, 2024).

When an organ is transplanted, it is important to ensure the recipient's immune system will accept it and not recognize it as foreign; otherwise, an organ can be rejected (Nguyen et al., 2025). The most common immune responses are described as acute rejection, which happens shortly after the transplant has taken place, and chronic rejection, which occurs over time (Rizvi et al., 2023b). Ultimately, failure associated with acute and chronic rejection can be caused through progressive permanent damage in the transplanted organ with long-term changes in allograft function (Wei et al., 2024). The health consequences of kidney rejection for recipients are devastating with damage culminating in graft dysfunction such as a reduction in eGFR or complete graft failure (Sawinski et al., 2021b).

Currently, the monitoring of kidney transplant physiology is performed using a combination of clinical and biochemical monitoring protocols (Oellerich et al., 2021). Clinicians often take note of potential clinical signs or symptoms related to

possible kidney dysfunction, such as decreased urinary output or developing edema in the legs and ankles from fluid overload (Kim et al., 2024a). However, to truly assess graft dysfunction, clinicians must perform invasive procedures to obtain tissue or organ biopsies, which allows for definitive diagnosis of the pertinent histopathological process (Moein et al., 2024).

The current monitoring modalities have multiple problems that can greatly influence patient care (Rizvi et al., 2023b). First, serum creatinine is a lagging and suboptimal marker; in fact, serum creatinine will only begin to elevate when potentially irreversible injury to the donated kidney has already occurred (Rizvi et al., 2023a). Surgery is the acceptable gold standard for diagnosis of kidney disease; however, the kidney biopsy is invasive and carries considerable risk of bleeding, pain, and infection to the patient, netting small benefit to the patient (Cheng et al., 2021). Further, there is a risk of failing to recognize subtle or early signs of rejection and between clinical assessments/reporting, the potential chronic injury impacting long-term graft survival may occur with progressive injury (Bu et al., 2022).

Both the invasiveness of biopsies and the lagging sensitivity of serum creatinine are limitations in post-transplant management (Aubert et al., n.d.). There is a clear need for a noninvasive and improved biomarker for the kidney transplant rejection technique that can provide an early and accurate evaluation of graft health (Akalin et al., 2021).

Cell-free DNA use in transplantation

cfDNA, which consists of the short fragments of DNA that are free circulating in the bloodstream, shows promise as a biomarker in transplantation research (Zhang et al., 2020b). Cell free DNA originates when cells in the human body experience apoptosis, programmed cell death or necrosis, uncontrolled cell death (Oellerich et al., 2021). When cells die and degrade, the genetic information in the cells is released into the biological fluid surrounding the cells including plasma blood, urine, and cerebrospinal fluid (Nguyen et al., 2025).

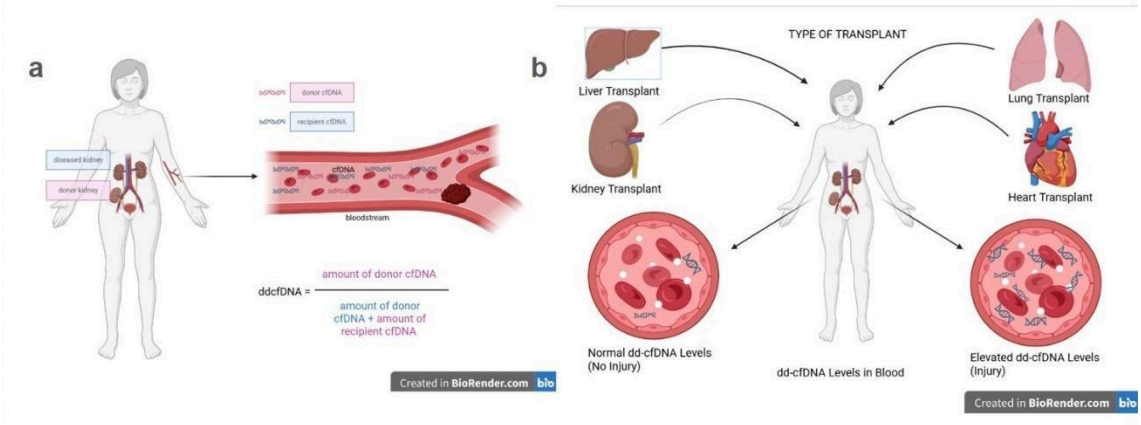


Figure 1: The Biological Principle of Donor-Derived Cell-Free DNA (dd-cfDNA) as a Biomarker. (a) This diagram illustrates that following a kidney transplant, both recipient and donor cfDNA circulate in the bloodstream. The fractional abundance of dd-cfDNA is calculated as the proportion of donor cfDNA relative to the total amount of cfDNA. (b) This diagram shows that dd-cfDNA is a general biomarker for all solid organ transplants, including liver, kidney, lung, and heart. In a healthy, uninjured state, dd-cfDNA levels are low; however, during an injury event like rejection, damaged cells from the transplanted organ release a higher amount of dd-cfDNA into the blood.

cfDNA comes from blood plasma, the fluid part of blood (Wolf-Doty et al., 2021). Several molecular methods are used currently to assess cfDNA quantity (Halloran et al., 2022). Both quantitative polymerase chain reaction (qPCR) and next-generation sequencing (NGS) can be viewed as modern methods for measuring cell-free DNA (cfDNA) (Akalin et al., 2021). qPCR does work similarly to a targeted measure, where specific primers are used to locate and amplify a pre-established and a small number of unique donor sequences (Nie et al., 2025). For NGS, millions of cfDNA fragments can be read concurrently across the genome (Kumar et al., 2023b). In this way, cfDNA can be assessed to identify the donor if they donated an organ (Wei et al., 2024).

Figure 1a provides biological context for utilizing dd-cfDNA as a biomarker. After a solid organ transplant (e.g. kidney), the recipient's blood will contain their own cfDNA in addition to a small amount of cfDNA from the donor organ (Figure 1a) (Oellerich et al., 2021). Under normal circumstances and with a healthy transplant, the amount of dd-cfDNA circulating in the blood is in a baseline state of low concentrations (Bu et al., 2022a). However, when an injury event causes a greater rate of cell death in the allograft, such as rejection, inflammation, or other forms of

graft stress, the allograft will release dd-cfDNA into the recipient's circulation (Figure 1b) (Rizvi et al., 2023b). These real-time changes in dd-cfDNA concentrations can reflect ongoing allograft injury (Nguyen et al., 2025).

cfDNA is a dynamic being in that the half-life of cfDNA is approximately 30 minutes in the bloodstream (Wolf-Doty et al., 2021). Changes in specific cfDNA characteristics, most notably an extraordinary increase in levels of cfDNA, are indications of tissue injury, inflammation, or subsequently, a disease such as rejection, which validate the value of cfDNA as a non-invasive biomarker (Rizvi et al., 2023). This indicates that cfDNA is a valid source of non-invasive biomarker when it is as simple as a blood draw to obtain cfDNA anyway (Wei et al., 2024). The drive and excitement around a real-time multi-level signal for more precise management is the appeal of this biomarker that affirms its usefulness (Aubert et al., 2024). dd-cfDNA has become such a strong force and continues to be a focus of transplant research (Halloran et al., 2022).

Methods

Search Strategy

A comprehensive literature search was conducted using the Publish or Perish application (*Publish or Perish*, n.d.) to identify relevant articles. Searches were performed using the Google Scholar and PubMed databases. The search focused on publications related to donor-derived cell-free DNA (dd-cfDNA) and transplantation, utilizing the following keywords: "dd-cfDNA", "donor-derived cell-free DNA", "cell-free DNA renal transplant", "allograft surveillance biomarkers", "kidney transplant rejection", "biomarker acute rejection", "Kidney Transplantation" (MeSH), "Graft Rejection" (MeSH), "Cell-Free Nucleic Acids" (MeSH), "Biomarkers" (MeSH), "Sensitivity and Specificity" (MeSH), "Comparative Effectiveness Research" (MeSH), "Diagnostic Imaging" (MeSH), "Creatinine/blood" (MeSH), "Monitoring, Immunologic" (MeSH), "diagnostic tool", "diagnostic performance", "surveillance method", "human subjects", "adult", "renal", "renal transplant", "kidney allograft", "diagnosis", "diagnostic", "AUC", "predictive value", "rejection", "sensitivity", "specificity", "performance", "comparison", "receiver operating characteristic curve", "serum creatinine", "biopsy", "imaging", "donor-specific antibodies", "DSA", "Racial Groups" (MeSH), "Ethnic Groups" (MeSH), "rejection monitoring", "allograft", "kidney".

These keywords were categorized to capture key aspects of the research question:

- **Population:** Keywords of “adult," "human subjects," "kidney transplant," "renal transplant," and "Kidney Transplantation" (MeSH) (including MeSH terms) were used to specify the patient group. "Racial Groups" (MeSH) and "Ethnic Groups" (MeSH) were included to address the "diverse ancestries" element.
- **Intervention:** Phrases such as "dd-cfDNA," "donor-derived cell-free DNA," "cell-free DNA renal transplant," and "Cell-Free Nucleic Acids" (MeSH) specifically targeted the primary diagnostic tool of interest.
- **Outcome:** Terms like "acute rejection," "graft rejection," "allograft rejection," and "Graft Rejection" (MeSH) directly identified the clinical event under investigation.
- **Comparison:** To encompass "standard biochemical and imaging surveillance," keywords like "serum creatinine," "biopsy," "imaging," "donor-specific antibodies," and related MeSH terms were utilized.
- **Diagnostic Role/Timing:** Phrases such as "diagnostic tool," "diagnostic performance," "surveillance method," "sensitivity," "specificity," "AUC," and "predictive value" were crucial for identifying studies assessing the diagnostic utility, accuracy, and particularly the comparative timing of dd-cfDNA.

The search process was restricted to articles published between January 1st, 2020, and May 31st, 2025.

Inclusion Criteria

Inclusion criteria were designed to bring in articles directly aligned with the research question:

English-language studies; Human subjects, specifically adult kidney transplant recipients (age ≥18 years); Use of donor-derived cell-free DNA (dd-cfDNA) as a primary or comparative diagnostic tool for acute rejection; Includes comparison to at least one standard surveillance method (e.g., serum creatinine, biopsy, imaging, donor-specific antibodies); Studies assessing the diagnostic role of dd-cfDNA in acute rejection, ideally reporting metrics (e.g., sensitivity, AUC) or qualitatively describing performance; Studies involving patients from diverse racial/ethnic backgrounds (where available); Studies published between January 2000 and January 2025

Articles not meeting the inclusion criteria were excluded.

Selection Process

The resulting set of articles were exported into a Google Sheets document, which served as the main data organization tool. Following this, a preliminary screening of manual reading of abstracts and titles was conducted to decide whether to include or exclude, the screening included: Reasoning as to why the article was chosen, Full Text Accessibility (only for texts included), Number of Citations, Authors, Title of Article, Year Published, Source, Publisher, Article URL, Cites URL, GS Rank, Query Date, Type of Article, DOI, ISSN, Citation URL, Volume, Issue, Start Page, End Page, ECC, Cites Per Year, Cites Per Author, Author Count, Age, and Abstract.

Data Extraction and Synthesis

The initial data extraction included standard article metadata such as GS Rank, Query Date, DOI, ISSN, Citations, Age of publication, and Author Count. For included articles, specific variables relevant to the research question were gathered. Gemini (Google LLC, 2025) was used to assist in extracting these variables, including: Title, Journal, Publication Year, Country, Study (single or multicenter), Study Type, Sample Size, Average Age, Proportion Male, Patient Race/Ethnicity, Type of Rejection, Type of cf-DNA, and AUC.

Study Scoring and Quality Assessment

A scoring system was applied to assess the relevance and methodological quality of each included article. Articles were awarded one point (1) if the following criteria were met, and zero points (0) if not. This included: Measured Blood? Analyzes Rejection Probability? Temporal Precedence Explicitly Stated? Longitudinal Monitoring for Timing? Compared to Traditional Biomarkers?

These scoring criteria were specifically chosen to assess each study's direct relevance and methodological rigor in answering the research question: "Among adult kidney transplant recipients across diverse ancestries, how much earlier can dd-cfDNA identify acute rejection compared to standard biochemical and imaging surveillance?" "Measured Blood?" confirmed the appropriate sample for dd-cfDNA, while "Analyzes Rejection Probability?" ensured focus on the primary outcome. "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for

Timing?" were crucial for evaluating the "earlier identification" aspect. Finally, "Compared to Traditional Biomarkers?" verified the study included the necessary comparative analysis against standard surveillance methods.

Data Synthesis

Articles receiving the high scores (3-5 points) were selected for in-depth analysis. Preliminary conclusions were written from these high-scoring articles after reading them entirely and compiled into a summary document with a brief summary and overview of each article for future data synthesis and reference.

Data Analysis

Data analysis involved a narrative synthesis of findings from high-scoring articles. Extracted data on diagnostic timing, performance metrics, such as AUC and sensitivity, and patient demographics (including ancestry) were noted for future reference and figures.

Data Visualization

All figures presented in the Results section were generated using Google Colab, a cloud-based computational environment. Data extracted from the included studies were compiled into a structured dataset and imported into the Colab environment for analysis. The Python programming language, along with the data manipulation library Pandas and the data visualization libraries Matplotlib and Seaborn, were used to create all plots, including histograms, box plots, and scatter plots, to visually synthesize the findings of this systematic review.

Literature Summary

Overview

The articles in this systematic review generally intended to assess the clinical potential and utility of dd-cfDNA as a viable, non-invasive, and convenient (it requires only a standard blood draw) quantifiable biomarker the monitor the health and overall condition of kidney allografts (Wei et al., 2024). The relevance of

dd-cfDNA can be assessed comparatively to other standard healthcare metrics such as serum creatinine and proteinuria (Bu et al., 2022). Most of the articles focused on and assessed dd-cfDNA's ability to identify allograft injury sooner (Bromberg et al, 2024), with many articles constructed to quantify exactly how much sooner dd-cfDNA may identify rejection in comparison to other biomarkers (Parajuli et al., 2024b) or changes in lagging biomarkers like serum creatinine (Schenk et al., 2024). In order to do this, articles evaluated dd-cfDNA's ability to identify both acute and other subclinical forms of rejection (Bu et al., 2022) alongside its ability to discriminate between antibody-mediated rejection (AMBR) and T-cell mediated rejection (TCMR) (Bromberg el al, 2024). Additionally, a number of articles investigated dd-cfDNA in the longitudinal monitoring of patient outcomes, not only for pre-rejection surveillance (Mantios et al., 2023) but also for assessing their response to treatment (Wolf-Doty et al., 2021).

Evaluating the efficacy of dd-cfDNA in estimating rejection

The first goal of the surveyed research was to quantify how dd-cfDNA was distinguished between rejection states (Mantios et al., 2023), thereby guiding clinical decisions by providing a non-invasive tool that, due to its high specificity, could help avoid unnecessary invasive biopsies (Wei et al., 2024). The primary measure used to evaluate dd-cfDNA's diagnostic performance is the Area Under the Receiver Operating Characteristic (AUROC) curve. The AUROC is a valuable statistical measure that summarizes a test's ability to correctly classify patients with and without a condition across all possible thresholds; a value of 1.0 represents a perfect test, while a value of 0.5 represents a test with no discriminatory ability.

Early Detection with dd-cfDNA

Additional research focused on the ability of dd-cfDNA to detect graft injury significantly earlier than clinical presentation (Bromberg et al., 2024) or elevations in standard biochemical markers like serum creatinine and proteinuria, which are known to be lagging (Kim et al., 2024b). Dd-cfDNA’s 30 minute half-life allows for the dynamic, consistent monitoring of ongoing damage and recovery (Wolf-Doty et al., 2021).

Identification of clinically relevant subtypes

Other publications were focused on building models to differentiate between AMBR

and TCMR (Botella et al., 2024b) . The key finding was that dd-cfDNA was a strong potential biomarker for ABMR in patients who are clinically stable and have normal renal function (as measured by serum creatinine) (Botella et al., 2024a) . This is a huge advantage, since rejection at the subclinical level of ABMR can cause silent cumulative damage to the graft. Another published study described two patients who developed biopsy proven ABMR within 60 days of a significantly elevated dd-cfDNA, despite both patients demonstrating normal renal function (Rizvi et al., 2023b) .

Dd-cfDNA is particularly useful in certain clinical settings and high-risk populations. The data suggests that dd-cfDNA is superior to creatinine for quantifying which patients with dnDSA (de novo donor-specific antibodies)—which are harmful immune proteins developed by the recipient that specifically attack the new organ—were definitively rejecting (Botella et al., 2024b), and needs to be considered for treatment decisions and guidance especially 15 days after transplantation (Botella et al., 2024a).

Performance in Multiple Ancestries

To ensure dd-cfDNA to be used as a valid biomarker, there needs to be evidence from studies with racially and ethnically diverse cohorts of patients (Stites et al., 2020). This race and ethnicity diverse cohorts of patients (Stites et al., 2020). This racial and ethnic diversity is important in acquiring generalizable evidence for the wider population of kidney transplant recipients. A number of studies reviewed did include or reported minority populations as part of their cohorts (Sawinski et al., 2021c).

Analysis

In order to fulfill the objectives of this systematic review, the results were synthesized by firstly, describing the studies included, and then quantifying and gaining a general idea about the diagnostic accuracy of dd-cfDNA, and lastly, identifying some important performance trends and knowledge gaps.

First, to synthesize the characteristics and findings of the included studies (Goal 2), the compiled list of 83 studies that met the inclusion criteria was examined. We then examined the distribution of studies per publication year (Fig. 1a). We found

that publications have been steadily increasing over time, peaking in 2024 with approximately 22 studies, which suggests a rapid maturation of the evidence base. Furthermore, we wanted to understand how sample size changes with time in the field of kidney transplant dd-cfDNA research (Fig. 1b). Similar to the number of publications, we found that the median sample size has been increasing, especially in 2024, when a large study of nearly 3,000 patients (Aubert et al., 2024) was published, representing the largest dd-cfDNA study in kidney transplantation to date. The evidence is predominantly composed of high-quality study designs, those following participants forward in time and observing their outcomes as they happen, thereby minimizing bias, with prospective studies forming the largest single category (Fig. 1e).

To address the third goal of identifying trends and knowledge gaps, particularly regarding the biomarker's utility in diverse, underrepresented patient populations, the geographic and demographic distribution of the studies was analyzed. An examination of participant characteristics confirmed that the study focuses on a middle aged cohort, with an average age of 52 years (Fig. 1c), which is both a useful finding and a commonly observed finding in clinical practice since the average age of transplantation is about 52.9. The analysis also highlighted that study cohorts tended to be predominantly male with a very prominent peak, that was a proportion of male participants up to a significant 0.63 (Fig. 1d). To gain a better understanding of the generalizability and nature of the evidence in a global context a breakdown of the study types by geographic region was conducted. This allowed us to see how the research in North America varied in methods put into practice, contributing the most retrospective and observational research, whereas research in Europe and Asia contributed higher levels of prospective designs (Fig. 1f). The choice to go through the analysis this way was significant because this page had a demonstrated blend of complementary types of research to provide a wider and more diverse evidence base - the prospective studies provide higher quality-controlled data while the retrospective and observational studies filled in the gap of real-word evidence, which in sum produced better conclusions for this review.

Finally, to specify and quantify the overall diagnostic accuracy of dd-cfDNA for various types of rejection (Goal 1), the performance metrics from the included studies were synthesized. An important goal of this study was to understand the

efficacy of cfDNA biomarkers in multiple ancestries, because it is important to have an equitable biomarker. To this end, we first examined the number of studies covering each self-reported race/ancestry group, which we defined as “Asian”, “Black”, “White”, and “Hispanic/Latino.” We found that the studies analyzing Black and Hispanic/Latino populations come almost exclusively from North American studies, while data on Asian populations is contributed by studies from both North America and Asia, primarily in Japan, China, Singapore and Thailand.

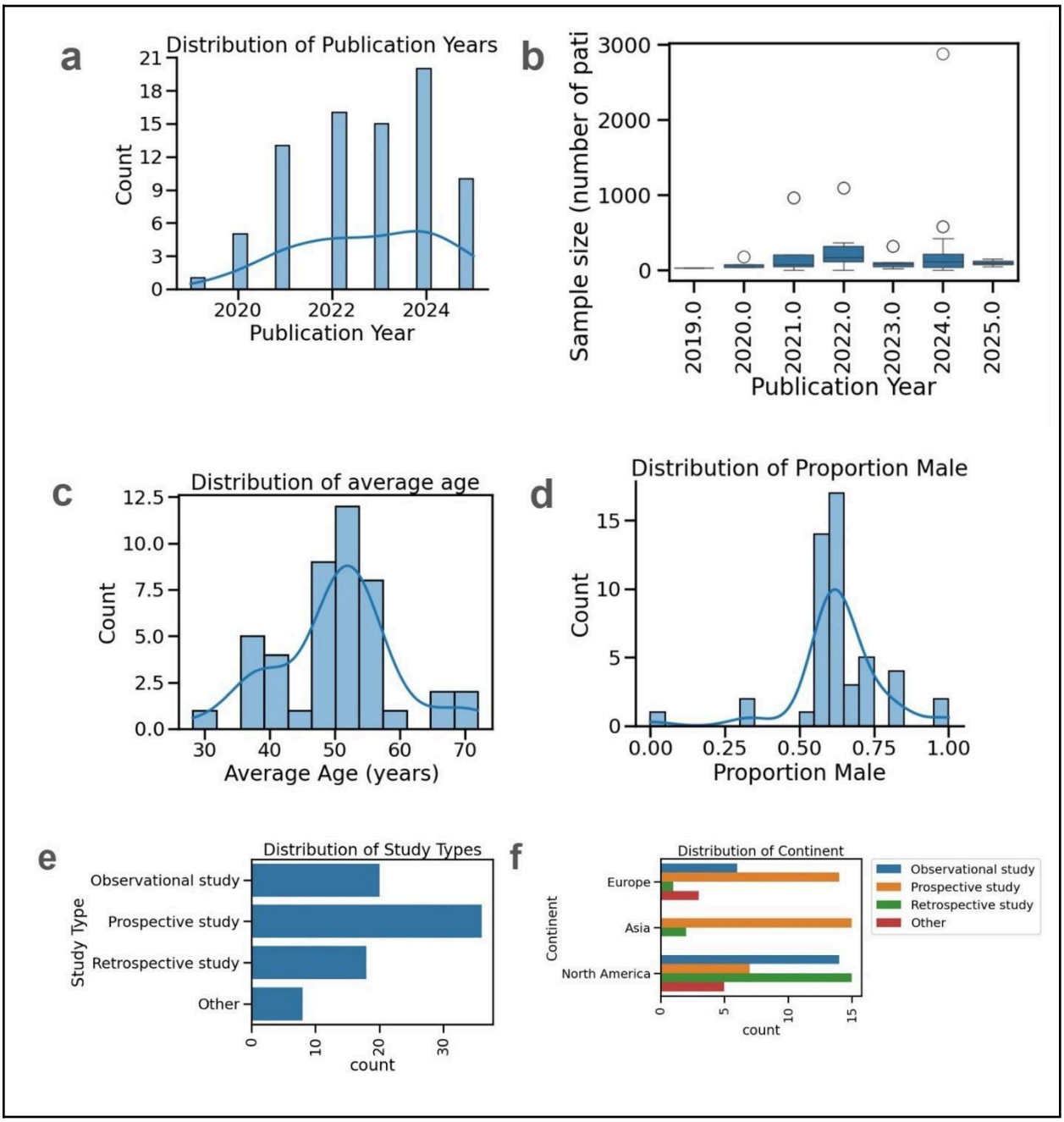
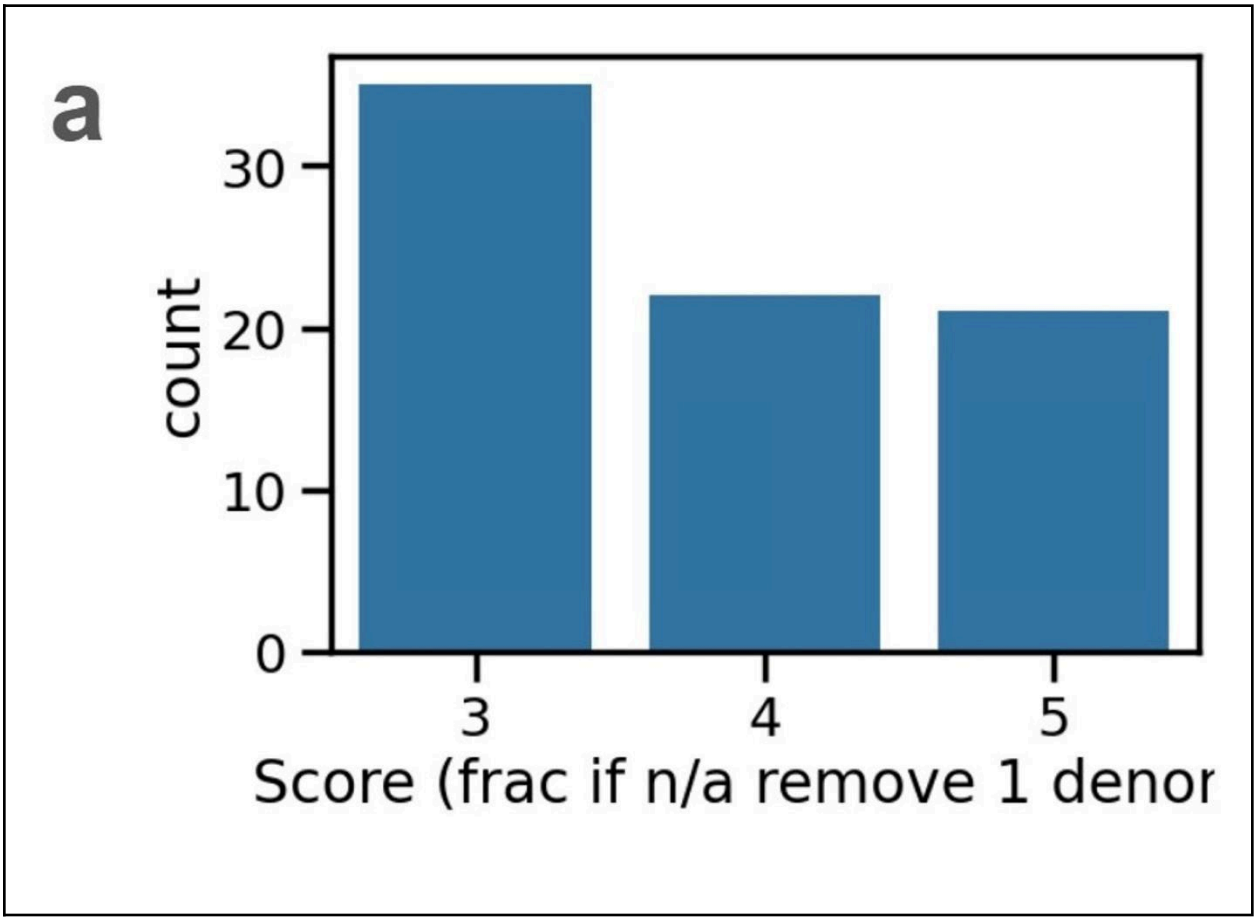


Figure 2: Publication Trends, Demographics, and Methodological Characteristics of Included Studies. (a) A histogram showing the number of included studies published per year, illustrating a rising trend in research volume peaking in 2024. (b) A box plot visualizing the distribution of study sample sizes by publication year, highlighting an increase in the scale of research and several large-scale outliers. (c) A histogram displaying the distribution of the average age of participants, showing a concentration in the middle-aged demographic. (d) A histogram of the proportion of male participants in study cohorts, indicating a predominance of male-majority studies. (e) A horizontal bar chart quantifying the included studies by their primary methodological design, with prospective studies being the most common. (f) A grouped horizontal bar chart illustrating the distribution of study types across the primary geographic regions of North America, Asia, and Europe.

To assess the methodological soundness of the literature used in the study, we designed a 5-point quality scoring system. The 5-point quality scoring system was created to determine which sources most fit the hypothesis. We awarded each study one point for fulfilling each of the criteria: (1) Measured Blood?, (2) Analyzes probability of rejection?, (3) Temporal Precedence Explicitly Stated?, (4) longitudinal monitoring for time?, (5) compares to traditional biomarkers?. These scoring criteria were chosen to evaluate the relevance of each study to the specific research hypothesis. The criteria "Measured Blood?" evaluated whether the study utilized the appropriate sample type to analyze dd-cfDNA (Sawinski et al., 2021a). The criteria "Analyzes rejection probability?" ensured that the study focused on the primary clinical outcome, i.e. rejection (Bu et al., 2022b). "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for Timing?" allowed for the evaluation of the hypothesis, specifically the earlier detection factor which required identifying studies that monitored the biomarker over time (Parajuli et al., 2024b). Finally, "Compares to Traditional Biomarkers?" validated that the study conducted the comparative clinical analysis of the complete framework, thereby determining the "superior performance" of dd-cfDNA analysis compared to traditional clinical surveillance methods such as serum creatinine (Mantios et al., 2023). The overall frequency distribution of quality scores (Figure 2a) indicates that the overall evidence base is generally high quality. Most studies reporting their scores are clustered in the upper range of the scale (e.g. 35 studies scored a 3; 22 studies scored a 4; and 21 studies scored a 5).

To add more specificity to the study characteristics that were contributing to high

quality evidence, we further broke down what criteria were met most frequently across the studies (Figure 2b). The breakdown shows that foundational criteria were met across studies, particularly utilizing blood as a sample source (n=79) and comparing dd-cfDNA to traditional biomarkers (n≈78). Despite this, there appears to be a large methodological gap in the literature regarding timing of rejection. While a reasonably sized number of studies (n=45) utilized some form of longitudinal monitoring for timing, a much smaller, and specific subset (n≈29) stated the temporal precedence of dd-cfDNA elevation before a rejection diagnosis was confirmed. This indicates a clear opportunity for future research to focus more rigorously on study designs that can quantify this early detection window.



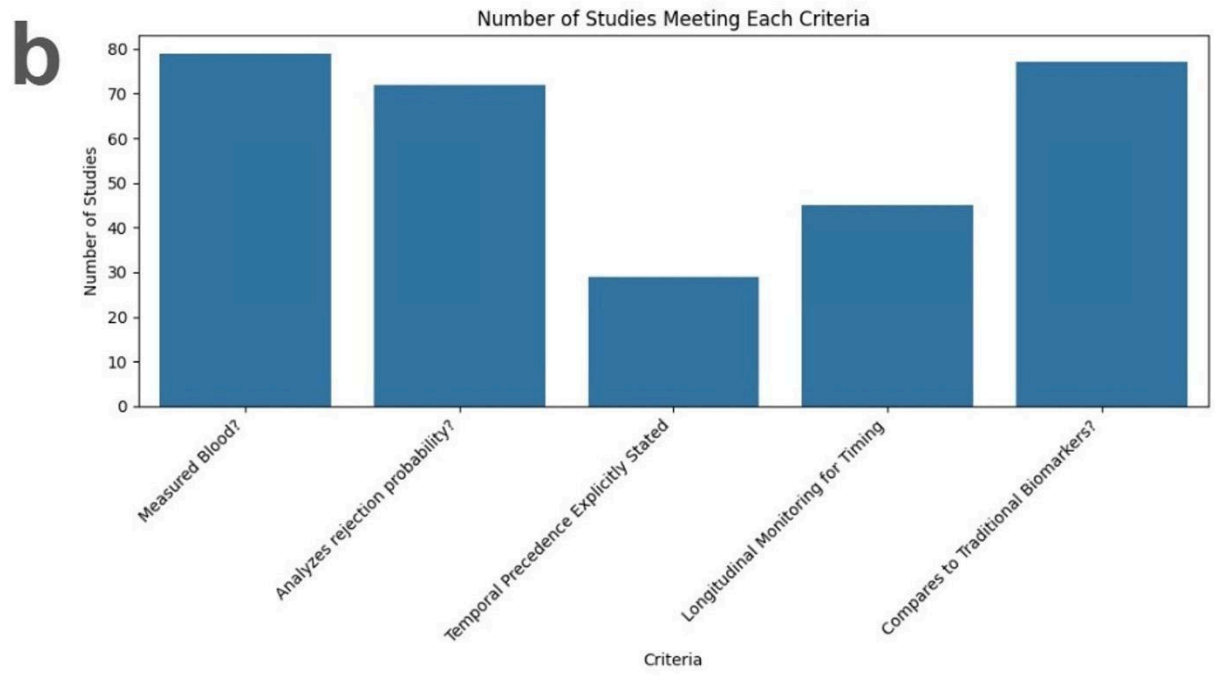


Figure 3: Distribution of Study Quality Scores and Fulfillment of Methodological Criteria. (a) A histogram showing the distribution of the overall 5-point quality scores across all included studies. (b) A bar chart quantifying the number of studies that met each of the five specific methodological quality criteria.

In order for us to understand the efficacy of dd-cfDNA in multiple ancestries, which is an important goal for comprehensive utility of ddcfDNA in kidney transplant settings, the inclusion of diverse racial and ethnic groups across the primary research continents was examined. The analysis reveals that North American research provides the most diverse evidence base. North American studies were the exclusive source of data for the inclusion of Black participants (n≈24) (Fig. 3a) and Hispanic/Latino participants (n≈18) (Fig. 3c). Furthermore, North American research also contributed the largest number of studies that include Asian participants (n≈17) (Fig. 3b) and White participants (n≈22) (Fig. 3d).

In contrast, the European studies included in this review, while numerous, did not explicitly report the inclusion of Black, Asian, or Hispanic participants, with many not reporting race altogether, though a small number did report including White patients (n≈3). The Asian studies, as expected, contributed evidence on Asian populations (n≈4) but did not report inclusion of other groups. This distribution

highlights a significant gap in the global evidence, suggesting that the generalizability of findings from Europe and Asia to minority populations is limited, and underscores the critical role of North American studies in validating dd-cfDNA's utility across a broad spectrum of ancestries.

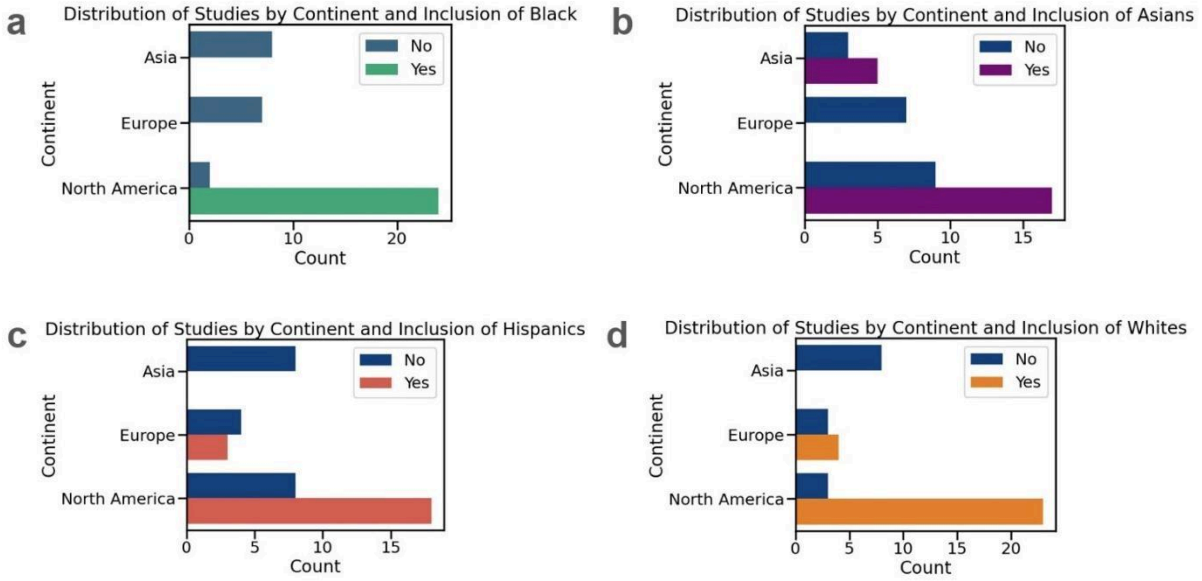


Figure 4: Geographic Distribution of Studies by Patient Ancestry. A series of horizontal bar charts showing the number of studies from North America, Europe, and Asia that did ("Yes") or did not ("No") report the inclusion of specific patient ancestries. Studies that did not report the race/ethnicity of their participants are not included. (a) Inclusion of Black participants. (b) Inclusion of Asian participants. (c) Inclusion of Hispanic/Latino participants. (d) Inclusion of White participants.

We had to determine how effective ddcfDNA was at identifying cases of kidney transplant rejection; therefore, AUC values were extracted from studied and analyzed. The diagnostic performance of dd-cfDNA was in terms of AUC was found to be high across the included studies, as summarized in the histogram of Area Under the Curve (AUC) values (Fig. 4a). The Receiver Operating Characteristic (ROC) curve is a representation of a diagnostic test's overall performance, as described by one number, the Area Under the Curve (AUC). The AUC takes a value between 0.5 (no better than chance) and 1.0 (perfect test). Values between 0.80 and 0.90 may be considered "good". For the purposes of this systematic review, the AUC indicates the extent to which dd-cfDNA separates patients with biopsy proven rejection (the cases) from patients without biopsy proven rejection (the controls); therefore, a

high AUC indicates that the biomarker is quite accurate in identifying patients that are actively rejecting their organ. The data reveals a consensus towards models using ddcfDNA having strong performance in the prediction of rejection, with a bimodal distribution showing two primary peaks where the evidence is concentrated. The largest cluster of studies reported an Area Under the Curve (AUC) with a peak centered approximately around AUC=0.80, while a second prominent peak centered even higher around 0.90. It is important to put this in perspective; for a standalone clinical diagnostic test, an AUC of 0.80 is generally considered poor as this indicates significant false positive and false negative rates that may cause indeterminate clinical decisions. A tempered conclusion is that models reporting dd-cfDNA levels show moderate to good potential as a surveillance test; a test that could assist with and not replace diagnosis per say. An interesting finding is that studies do not report mediocre performance and if this is potentially influenced by publication bias of studies with negative findings, it may at least suggest that when dd-cfDNA is investigated, it is consistently a good and effective biomarker for detection of allograft rejection.

Next, we examined whether the AUC, as a proxy for performance, has been changing as the literature matures. The diagnostic performance of dd-cfDNA was found to be reliably high in the studies that were included, as shown in the histogram of Area Under the Curve (AUC) values (Fig. 4a). There is a clear shared opinion towards high identifiability with a bimodal distribution presenting two peaks where the evidentiary support lies. The largest mass of studies report an AUC that is concentrated around 0.80 indicating a very good diagnostic test, whereas the second primary peak is concentrated even higher, around AUC 0.90, which demonstrates excellent diagnostic performance. The clear failure of studies to report poor or mediocre performance clearly indicates, in a particularly robust quantitative sense, that dd-cfDNA is a reliable and effective biomarker in allograft rejection detection.

We then performed a linear regression analysis to model the relationship of year of publication and AUC analyses in order to assess whether reported performance of the biomarker has improved as studies have accrued (a primary goal of this review) This does demonstrate an increasing association over time (slope ≈ +0.01 per year of AUC) (Fig. 4b). This positive slope suggests that more 'recent' studies, on average, have reported greater AUC than 'earlier' studies. This potential trend may in part

reflect field maturation, perhaps including improved laboratory assays (i.e., could be absolute quantification vs. fractional abundance) and more sophisticated study design (years and capabilities are reflected in the increasing median sample sizes). The evidence for dd-cfDNA has not weakened over time, it seems to have grown and solidified. An alternative explanation might include a growing publication pressure to report findings that best represent previous literature. This study's upward trend in AUC is accompanied by an upward trend in the sample sizes of studies over time may be another indication that there are larger statistically powerful trials contributing to stable and precise estimates of high performance.

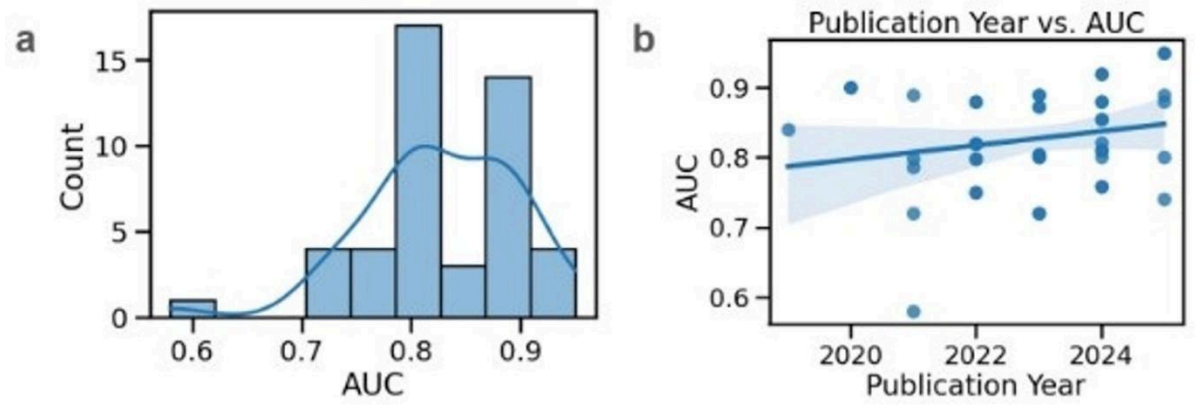


Figure 5: Summary of Diagnostic Accuracy and Performance Trends. (a) A histogram showing the bimodal distribution of reported Area Under the Curve (AUC) values from all included studies. (b) A regression plot illustrating the positive correlation between the publication year of the studies and their reported AUC values.

To further assess the generalizability of dd-cfDNA's performance, the Area Under the Curve (AUC) values were stratified by studies that reported the inclusion of different racial and ethnic groups. For this analysis, studies that did not report AUC or did not specify the racial demographics of their cohort were excluded, resulting in a subgroup of 45 studies. Overall, the data suggests that dd-cfDNA consistently has accuracy in identifying potential rejection across studies with different cohort demographics. The median AUC of studies that specifically included Black participants (N=24) was 0.81. This is an important observation because it supports the accuracy of this research for this patient population, which is specifically at higher risk for adverse transplant outcomes.

Another aspect we found was that the studies with Hispanic participants found a

similar accuracy to the Black patient population. The median AUC for Hispanic populations is 0.81, and like for the Black population, the interquartile range is similarly narrow and consistent, which suggests that the studies all reported similar results. The high consistency is meaningful since it reinforces that the biomarker is found to have reasonable performance in this ethnicity. This was consistent with large multisite studies that included Hispanics and reported strong positive results.

The studies measuring the performance of dd-cfDNA that reported Asian participation strengthen the evidence for applicability of dd-cfDNA across racial and ethnic groups. Interestingly, the analysis suggests that studies including Asian participants demonstrate a trend towards higher and more consistent accuracy, with a median Area Under the Curve (AUC) of approximately 0.82 and a notably narrow interquartile range. This enhanced consistency may be attributed to factors like more homogeneous patient cohorts or the use of optimized, population-specific assays, highlighting the potential for dd-cfDNA's performance to be further refined in specific clinical and demographic contexts.

Finally, the analysis of studies including White participants provides further evidence of biomarker's performance. The apparent better performance of studies that did not explicitly report including White participants is likely due to those studies being a heterogeneous mix of smaller or more homogeneous international cohorts where ethnicity is often not specified, whereas the "Yes" group consists primarily of large, diverse North American trials whose real-world complexity results in a more consistent, albeit slightly lower, median AUC.

In conclusion, the consistent high AUC values across all four of these distinct analyses strengthens the overall hypothesis that dd-cfDNA is a superior and broadly applicable biomarker for detecting allograft rejection in a wide spectrum of patient populations.

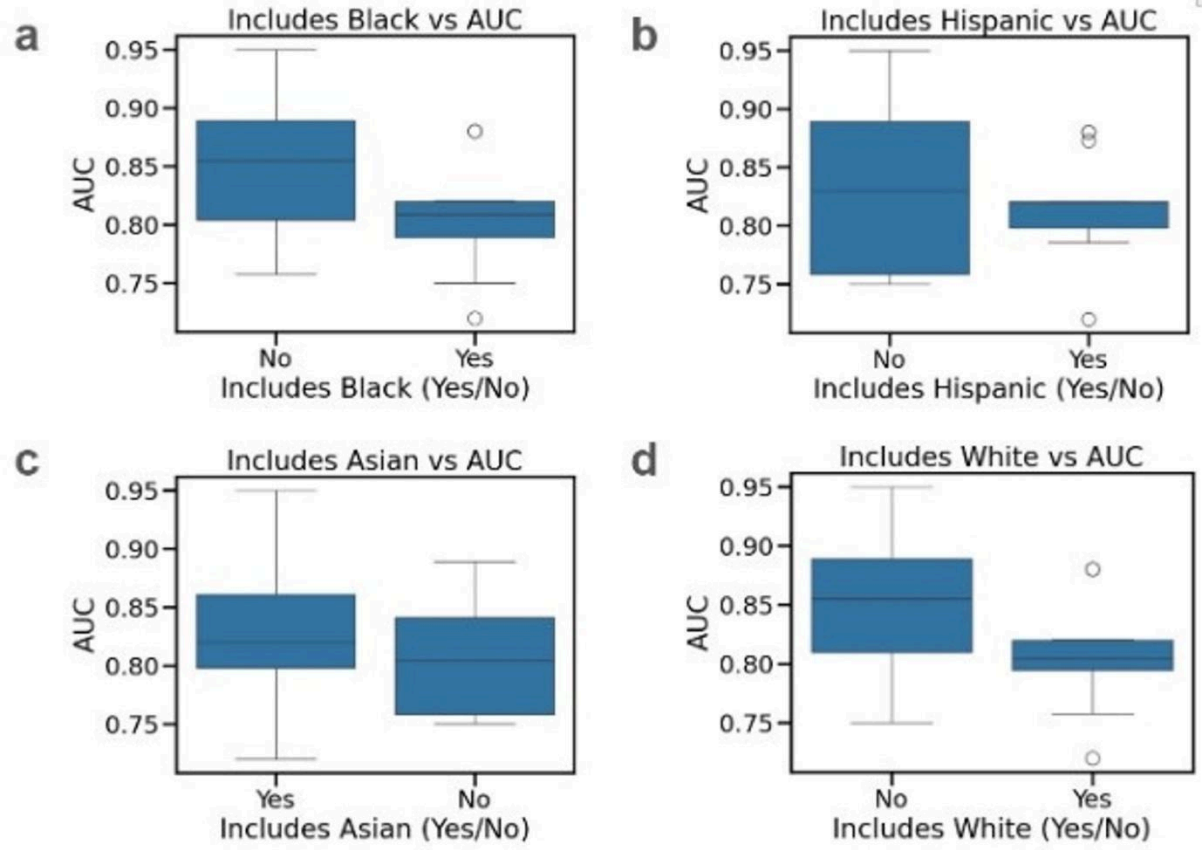


Figure 6: Diagnostic Accuracy (AUC) by Patient Ancestry Inclusion. A series of box plots showing the distribution of reported Area Under the Curve (AUC) values, comparing studies that did ("Yes") or did not ("No") report the inclusion of specific patient ancestries. (a) Distribution of AUC by inclusion of Black participants. (b) Distribution of AUC by inclusion of Hispanic participants. (c) Distribution of AUC by inclusion of Asian participants. (d) Distribution of AUC by inclusion of White participants.

Beyond the overall quality of the studies, it is also important to understand how the performance of dd-cfDNA varies across different study contexts, as detailed in Figure 6. An analysis of performance by geographic region reveals a consistently high level of diagnostic accuracy globally, with the median AUC for studies from Europe, Asia, and North America all above 0.80 (Fig. 6a). This suggests that the biomarker has promising performance across different healthcare systems and populations. Furthermore, an analysis of study scale shows no significant correlation between a study's sample size and its reported AUC (Fig. 6b); the nearly flat regression line indicates that high performance is not simply an artifact of small studies, a conclusion reinforced by the finding that the largest study with nearly

3,000 patients still reported a strong AUC.

The robustness of dd-cfDNA's performance is further supported by an analysis of demographic and methodological factors. The regression plot of average participant age versus AUC shows a slight downwards trend. (Fig. 6c). This indicates that in the age group between ages 30 to 60, the AUC decreased as age increased, thereby signifying the need for more studies to further validate this relationship. Finally, while both single-center and multicenter studies reported similarly high median AUCs, the violin plot reveals that the results from multicenter studies were more consistent, with a tighter distribution and less variability (Fig. 6d). One aspect to note about Fig. 6d is that a violin plot was utilized over a standard box plot because it visualizes the full probability density of the data.

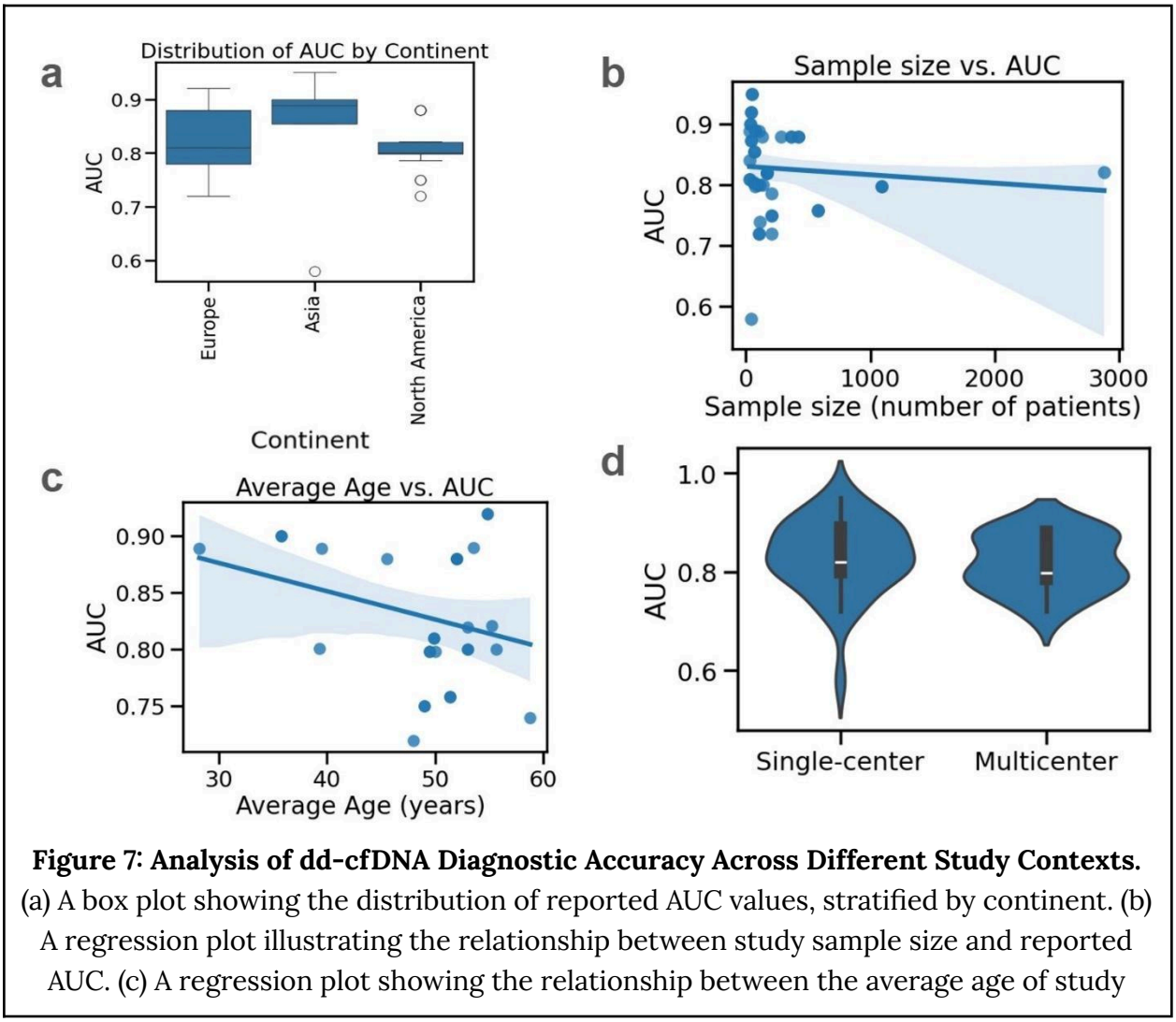


Figure 7: Analysis of dd-cfDNA Diagnostic Accuracy Across Different Study Contexts. (a) A box plot showing the distribution of reported AUC values, stratified by continent. (b) A regression plot illustrating the relationship between study sample size and reported AUC. (c) A regression plot showing the relationship between the average age of study

participants and reported AUC. (d) A violin plot comparing the distribution of reported AUC values between single-center and multicenter studies.

We also analyzed the relationship between the methodological quality of the included studies and their reported diagnostic accuracy (AUC). The box plot of overall quality scores reveals that studies with higher scores (3, 4, and 5) consistently reported high median AUCs above 0.80 (Fig. 7a). This indicates that the quality score of a study is not directly related to the reported AUC, as high diagnostic performance was observed across all quality tiers. A more granular analysis shows that studies incorporating more rigorous methodologies, such as longitudinal monitoring, not only reported a high median AUC but also showed more consistent results, evidenced by a tighter distribution of values compared to studies that did not meet this criterion (Fig. 7b). This suggests that the evidence supporting dd-cfDNA is strongest when derived from the studies that include a robust study design.

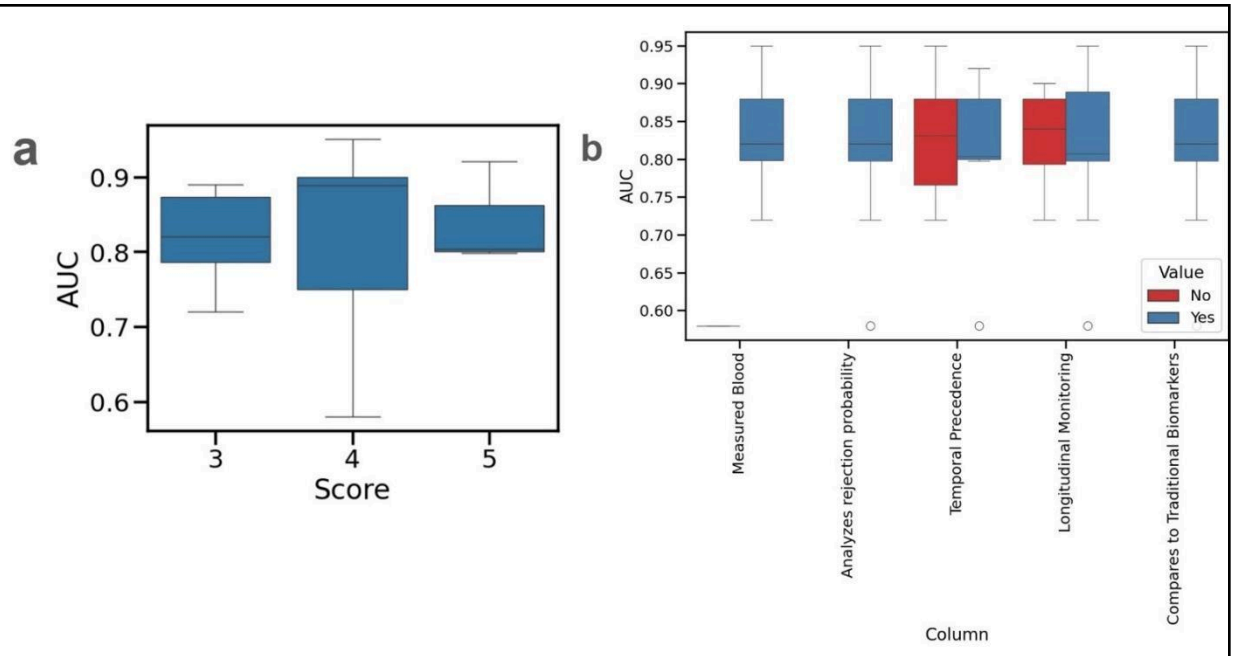


Figure 8: Relationship Between Study Quality and Diagnostic Accuracy (AUC). (a) A box plot showing the distribution of reported AUC values, stratified by the overall 5-point quality score of the studies. (b) A box plot comparing the distribution of reported AUC values for studies that did ("Yes") or did not ("No") meet each of the five specific methodological quality criteria.

Lastly, to understand the evidence base for different rejection pathways, the included studies were analyzed based on their inclusion of either antibody-mediated rejection (ABMR) or T-cell mediated rejection (TCMR). The analysis shows that the literature provides robust coverage of both primary rejection subtypes. A substantial majority of the included studies, approximately 48, investigated patient cohorts that included cases of ABMR (Fig. 8a), while a similarly large number of studies, approximately 45, included cases of TCMR (Fig. 8b), though many included both. This strong representation across both types of rejection provides a solid foundation for evaluating the biomarker's performance in different clinical scenarios.

Furthermore, an analysis of the reported diagnostic accuracy (AUC) shows that the high performance of dd-cfDNA is consistent regardless of a study's specific focus on rejection subtypes. The median AUC for studies that included ABMR was approximately 0.82, a value nearly identical to the median AUC for studies that did not specifically focus on ABMR (Fig. 8c). A similar pattern was observed for TCMR, where the median AUC was also consistent at approximately 0.82 in studies that investigated this rejection type (Fig. 8d)

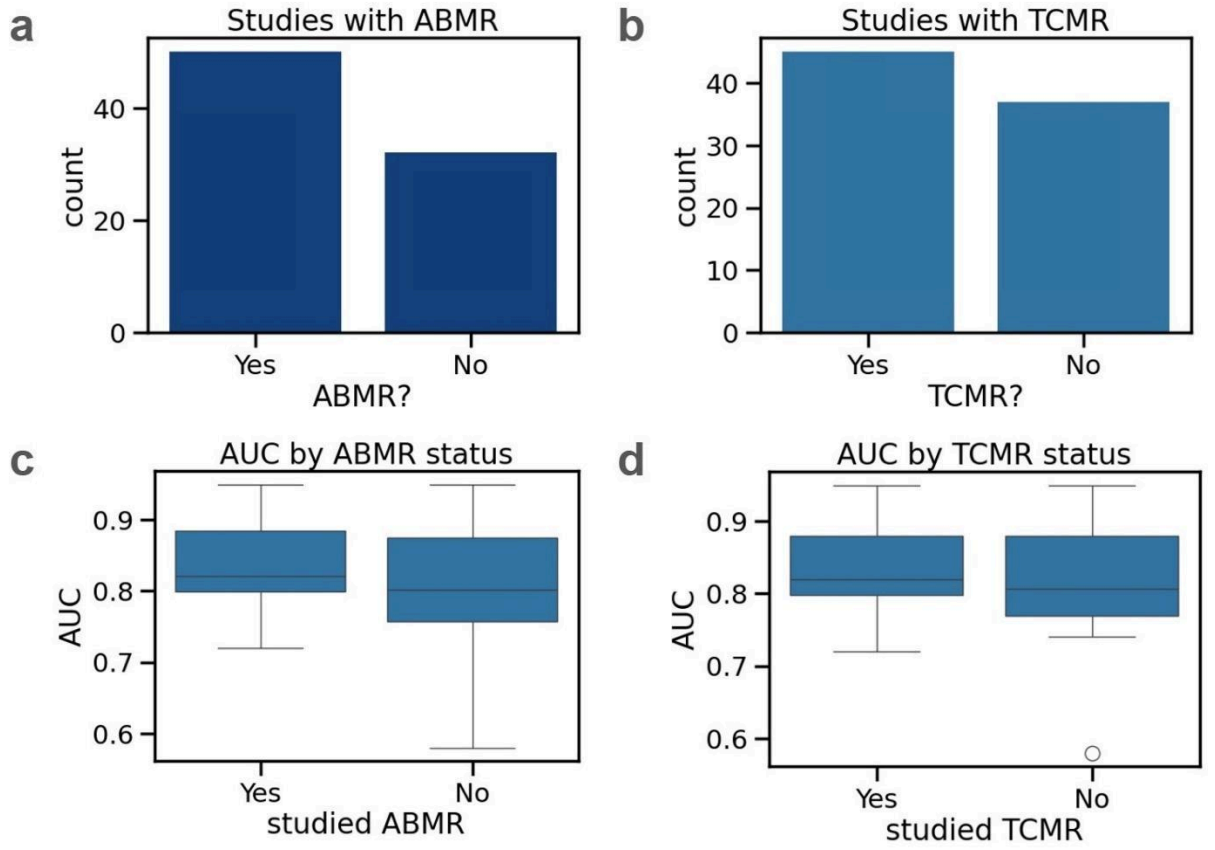


Figure 9: Research Focus on Rejection Subtypes and Corresponding Diagnostic Accuracy. (a) A bar chart showing the number of studies that did ("Yes") or did not ("No") investigate antibody-mediated rejection (ABMR). (b) A bar chart showing the number of studies that did ("Yes") or did not ("No") investigate T-cell mediated rejection (TCMR). (c) A box plot comparing the distribution of reported AUC values between studies that did or did not investigate ABMR. (d) A box plot comparing the distribution of reported AUC values between studies that did or did not investigate TCMR.

Limitations of dd-cfDNA in transplantation research

While the synthesized evidence provides strong support for the clinical utility of dd-cfDNA, this systematic review also produced some important limitations. The greatest limitation revealed in the analysis is insufficient ethnic and geographic diversity in the evidence base. The results show (Fig. 3) that the sample for Black and Hispanic/Latino populations is almost exclusively North American data (Ralph et al., 2025), inferring that it needs ongoing validation in global patient populations to confirm applicability for those populations.

The literature suggests dd-cfDNA is not a standalone diagnostic test, but a strong complementary tool for diagnostics, which is reflected by this review's findings on its variable diagnostics performance. The results demonstrated that there was variable diagnostic accuracy for dd-cfDNA in some rejection subtypes like T-cell mediated rejection (TCMR) and there was variability dependent upon assays (Halloran et al., 2022). These results informed that dd-cfDNA is most useful if used within a multi-modal approach with histology where dd-cfDNA has a supportive role in increasing surveillance and avoiding unnecessary invasive biopsy (Wei et al., 2024).

In summary, this review demonstrated that dd-cfDNA is a dynamic marker, that must be interpreted cautiously. The finding that the dynamics of dd-cfDNA may become "increasingly erratic" following kidney injuries further accentuates our argument that a single static measurement could be misleading (Nguyen et al., 2025). The assessment of the methodological quality of the studies included in this review brought attention to this issue, and when looking at the literature, we were able to identify a gap in the literature to explicitly assess the temporal precedence (Fig. 2b) of dd-cfDNA, and this further supports the conclusion that dd-cfDNA should be interpreted in a sequential manner over time as observations of trends in dd-cfDNA levels provides the best differentiation between true signals of injury versus transient fluctuations of observed dd-cfDNA levels (Stites et al., 2020).

Future Directions and Recommendations

To realize the clinical potential of donor derived cell-free DNA (dd-cfDNA) fully, the next steps need to focus on precision, standardization, and appropriate incorporation into existing clinical guidelines. This systematic review confirmed that the biomarker has high diagnostic accuracy (Fig. 4a), but as shown in this review, different commercial assays can have different rejection detection rates for the same type of rejection (Fig. 8d). To establish standardized diagnostic thresholds, robust, prospective trials are still needed (Benning et al., 2023b). The next phase should include head-to-head comparisons of the existing dd-cfDNA platforms to create standardized reporting metrics, as well as validate their performance across real-world populations (Loupy et al., 2024b). Also, there is a need for evidence-based guidelines regarding frequency of monitoring, such as testing

protocols of monthly monitoring for the first 6 months post-transplant, and quarterly thereafter, in order to determine the most cost-effective monitoring schedule for the early detection of subclinical rejection, which is still in an actionable state (Sharma et al., 2022).

The ultimate success of this biomarker will hinge upon its direct ability to translate into improvements in the longer-term outcomes of patients and allografts. As we established in this review of methodological quality, there were many studies (n≈45) that undertook longitudinal monitoring, and a much smaller number (n≈29) deliberately documented temporal precedence (Fig. 2b). This indicates an urgent need for intervention trials that utilize dd-cfDNA's early warning signal to direct pre-emptive treatment (i.e., if dd-cfDNA heightens, initiate increased immunosuppression) (Mirza et al., 2024). For example, a randomized controlled trial would directly compare a dd-cfDNA based strategy with the current standard of care (i.e., initiation of treatment when a clinical event has already occurred) (Tian et al., 2025). This would provide direct evidence as to whether or not these intervention strategies can improve the trajectory of allograft health (as currently we have only observational evidence that an increased renewable dd-cfDNA level considerably raises the risk of decline in the subsequent future graft health) (Obrișcă et al., 2022).

Although dd-cfDNA is now being utilized in clinical practice, this systematic review established several important limitations within the existing literature that need to be overcome in order to properly refine the field of dd-cfDNA use. The guidelines on dd-cfDNA use are evolving quickly, but the best way to incorporate dd-cfDNA into those guidelines is yet to be resolved (Kim et al., 2024c). The fact that all of the evidence for use in Black and Hispanic populations came from North American based studies (Fig. 3), there is an immediate need for future research on these underrepresented populations in Europe and Asia to allow for equitable and effective dd-cfDNA use (González-López et al., 2023). When we operationalize these knowledge gaps - focusing on performance in broader and diverse cohorts, as well as specific rejection sub-types like TCMR - the clinical community can improve the utility of dd-cfDNA and establish it as a new standard in personalized, non-invasive transplant care (Dauber et al., 2020).

Conclusions

This systematic review serves as a thorough synthesis of recent and robust research on donor-derived cell-free DNA that helps solidify its importance in the post-transplant period (Huang et al., 2023). The overall body of evidence supports that dd-cfDNA is consistently more accurate diagnostically than traditional biomarkers, such as serum creatinine, because it can detect rejection in those with stable creatinine levels (Gupta et al., 2022). The overall implication is that the use of dd-cfDNA represents a shift from reactive to proactive, evidence-based monitoring and screening that enables clinicians to detect allograft injury weeks or months before it presents clinically (Benning et al., 2023a). Although this review did systematically confirm that a number of studies reported a consistently high measure of diagnostic accuracy (eg, > 0.80 as the AUC; Mantios et al., 2023), it also highlighted an important knowledge gap around the biomarker as most of the data on black and Hispanic populations came exclusively from North American cohorts. One concrete recommendation based on this analysis is that future meta-analyses should synthesize other core metrics, such as Negative Predictive Value (NPV), in order to clearly define how dd-cfDNA should be used across different clinical scenarios (Aubert et al., n.d.). In summary, this systematic review concludes that dd-cfDNA is a milestone in transplant monitoring that provides a foundation for more personalized and non-invasive patient care (Loupy et al., 2024a).

Code and Data Availability

Data collected as part of this systematic review is freely available at https://docs.google.com/spreadsheets/d/1KAfEUQF-zXAJviuMSKyAEOpJUpMJbEMjBCOCbhbmgj4/edit?usp=sharing. The Google Colab notebook used for analysis and plot generation is freely available at https://colab.research.google.com/drive/1oq564UU3FZyur_mscmA8rdN5-H1EsuY1?usp=sharing.

Acknowledgements

I would like to express my sincere gratitude to my mentor, Christa Caggiano, whose insightful feedback and critical comments at every stage were invaluable in shaping the direction and clarity of this systematic review. I am especially grateful for her mentorship in data science, as she patiently taught me the fundamentals of

programming in Python to transform the raw data into the meaningful figures presented in the Results section. This project would not have been possible without her generous investment of time, knowledge, and unwavering support.

References

Akalin, E., Weir, M. R., Bunnapradist, S., Brennan, D. C., Delos Santos, R., Langone, A., Djamali, A., Xu, H., Jin, X., Dholakia, S., Woodward, R. N., & Bromberg, J. S. (2021). Clinical Validation of an Immune Quiescence Gene Expression Signature in Kidney Transplantation. *Kidney360*, 2(12), 1998. https://doi.org/10.34067/KID.0005062021

Akifova, A., Budde, K., Amann, K., Buettner-Herold, M., Choi, M., Oellerich, M., Beck, J., Bornemann-Kolatzki, K., Schütz, E., Bachmann, F., Halleck, F., von Hoerschelmann, E., Koch, N., Schrezenmeier, E., Seelow, E., Waiser, J., Zukunft, B., Eckardt, K.-U., Halbritter, J., … Osmanodja, B. (2025). Donor-derived cell-free DNA monitoring for early diagnosis of antibody-mediated rejection after kidney transplantation: A randomized trial. *Nephrology Dialysis Transplantation*, 40(7), 1384–1395. https://doi.org/10.1093/ndt/gfae282

Aubert, O., Ursule-Dufait, C., & Brousse, R. (n.d.). *Cell-free DNA for the detection of kidney allograft rejection | Nature Medicine*. Retrieved August 11, 2025, from https://www.nature.com/articles/s41591-024-03087-3

Aubert, O., Ursule-Dufait, C., Brousse, R., Gueguen, J., Racapé, M., Raynaud, M., Van Loon, E., Pagliazzi, A., Huang, E., Jordan, S. C., Chavin, K. D., Gupta, G., Kumar, D., Alhamad, T., Anand, S., Sanchez-Garcia, J., Abdalla, B. A., Hogan, J., Garro, R., … Loupy, A. (2024). Cell-free DNA for the detection of kidney allograft rejection. *Nature Medicine*, 30(8), 2320–2327. https://doi.org/10.1038/s41591-024-03087-3

Benning, L., Morath, C., Fink, A., Rudek, M., Speer, C., Kälble, F., Nusshag, C., Beimler, J., Schwab, C., Waldherr, R., Zeier, M., Süsal, C., & Tran, T. H. (2023). Donor-Derived Cell-Free DNA (dd-cfDNA) in Kidney Transplant Recipients With Indication Biopsy—Results of a Prospective Single-Center Trial. *Transplant International*, 36, 11899. https://doi.org/10.3389/ti.2023.11899

Benning, L., Morath, C., Kühn, T., Bartenschlager, M., Kim, H., Beimler, J., Buylaert, M., Nusshag, C., Kälble, F., Reineke, M., Töllner, M., Schaier, M., Klein, K., Blank, A., Schnitzler, P., Zeier, M., Süsal, C., Bartenschlager, R., Tran, T. H., & Speer, C. (2022). Humoral response to SARS-CoV-2 mRNA vaccination in previous non-responder kidney transplant recipients after short-term withdrawal of mycophenolic acid. *Frontiers in Medicine*, 9. https://doi.org/10.3389/fmed.2022.958293

Botella, C., Galián, J. A., Jiménez-Coll, V., Fernández-González, M., Morales, F.,

Martínez-Gómez, G., González-López, R., Alegría, M. J., Moya, M. R., Martinez-Banaclocha, H., Minguela, A., Legaz, I., Llorente, S., & Muro, M. (2024a). Early Monitoring of Donor-Derived Cell-Free DNA in Kidney Allograft Recipients Followed-Up for Two Years: Experience of One Center. *Life*, 14(11), 1491. https://doi.org/10.3390/life14111491

Botella, C., Galián, J. A., Jiménez-Coll, V., Fernández-González, M., Morales, F., Martínez-Gómez, G., González-López, R., Alegría, M. J., Moya, M. R., Martinez-Banaclocha, H., Minguela, A., Legaz, I., Llorente, S., & Muro, M. (2024b). Early Monitoring of Donor-Derived Cell-Free DNA in Kidney Allograft Recipients Followed-Up for Two Years: Experience of One Center. *Life*, 14(11), 1491. https://doi.org/10.3390/life14111491

Bromberg, J. S., Bunnapradist, S., Samaniego-Picota, M., Anand, S., Stites, E., Gauthier, P., Demko, Z., Prewett, A., Armer-Cabral, M., Marshall, K., Kaur, N., Bloom, M. S., Tabriziani, H., Bhorade, S., & Cooper, M. (2024). Elevation of Donor-derived Cell-free DNA Before Biopsy-proven Rejection in Kidney Transplant. *Transplantation*, 108(9), 1994–2004. https://doi.org/10.1097/TP.0000000000005007

Bu, L., Gupta, G., Pai, A., Anand, S., Stites, E., Moinuddin, I., Bowers, V., Jain, P., Axelrod, D. A., Weir, M. R., Wolf-Doty, T. K., Zeng, J., Tian, W., Qu, K., Woodward, R., Dholakia, S., Golovine, A. D., Bromberg, J. S., Murad, H., & Alhamad, T. (2022). Clinical outcomes from the Assessing Donor-derived cell-free DNA Monitoring Insights of kidney Allografts with Longitudinal surveillance (ADMIRAL) study. *Kidney International*, 101(4), 793–803. https://doi.org/10.1016/j.kint.2021.11.034

Chen, X.-T., Qiu, J., Wu, Z.-X., Zhang, H., Chen, T., Yang, S.-C., Zhao, G.-D., He, Y., Shen, X., Luo, J.-Q., Huang, Y., Wang, C.-X., Chen, L.-Z., Wu, C.-L., & Huang, G. (2022). Using Both Plasma and Urine Donor-Derived Cell-Free DNA to Identify Various Renal Allograft Injuries. *Clinical Chemistry*, 68(6), 814–825. https://doi.org/10.1093/clinchem/hvac053

Cheng, D., Liu, F., Xie, K., Zeng, C., Li, X., Ni, X., Ge, J., Shu, L., Zhou, Y., Shi, H., Liu, H., & Chen, J. (2021). Donor-derived cell-free DNA: An independent biomarker in kidney transplant patients with antibody-mediated rejection. *Transplant Immunology*, 69, 101404. https://doi.org/10.1016/j.trim.2021.101404

Cuadrado-Payán, E., González-Roca, E., Rodríguez-Espinosa, D., Molina-Andújar, A., Montagud-Marrahi, E., Arana, C., González-Rojas, A., Esforzado, N., Torregrosa, V., Ventura-Aguiar, P., Argudo, M., Serrano-Jorcano, D.,

Ramírez-Bajo, M. J., Bañón-Maneus, E., Casas, S., Broseta, J. J., Puig-Butille, J. A., Revuelta, I., Diekmann, F., & Cucchiari, D. (2025). Implementation of the in-house technique for the determination of donor-derived cell-free DNA in daily clinical practice: Experience from the Hospital Clinic of Barcelona. *Nefrología (English Edition)*, 45(4), 294–301. https://doi.org/10.1016/j.nefroe.2025.03.004

Dreige, Y., Lima, G., Cristelli, M., Nakamura, M., Foresto, R., Pierry, P., Glehn-Ponsirenas, R., Levi, J., Medina-Pestana, J., & Tedesco-Silva, H. (2022). 336.3: Clinical Utility of Donor-Derived Cell-Free DNA During the Period of Recovery of Renal Function After Kidney Transplantation. *Transplantation*, *106*(9S), S297. https://doi.org/10.1097/01.tp.0000886964.60815.ca

Halloran, P. F., Reeve, J., Madill-Thomsen, K. S., Kaur, N., Ahmed, E., Cantos, C., Al Haj Baddar, N., Demko, Z., Liang, N., Swenerton, R. K., Zimmermann, B. G., Van Hummelen, P., Prewett, A., Rabinowitz, M., Tabriziani, H., Gauthier, P., Billings, P., & Investigators*, the T. (2022). Combining Donor-derived Cell-free DNA Fraction and Quantity to Detect Kidney Transplant Rejection Using Molecular Diagnoses and Histology as Confirmation. *Transplantation*, *106*(12), 2435. https://doi.org/10.1097/TP.0000000000004212

Kidney Disease Statistics for the United States—NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved August 11, 2025, from https://www.niddk.nih.gov/health-information/health-statistics/kidney-disease

Kim, H. D., Bae, H., Kang, H., Lee, H., Eum, S. H., Yang, C. W., Choi, Y. J., Chung, B. H., & Oh, E.-J. (2024a). Donor-derived cell-free DNA predicted allograft rejection and severe microvascular inflammation in kidney transplant recipients. *Frontiers in Immunology*, *15*, 1433918. https://doi.org/10.3389/fimmu.2024.1433918

Kim, H. D., Bae, H., Kang, H., Lee, H., Eum, S. H., Yang, C. W., Choi, Y. J., Chung, B. H., & Oh, E.-J. (2024b). Donor-derived cell-free DNA predicted allograft rejection and severe microvascular inflammation in kidney transplant recipients. *Frontiers in Immunology*, *15*, 1433918. https://doi.org/10.3389/fimmu.2024.1433918

Kumar, N., Tandon, A., Rana, R., Rana, D. S., Bhalla, A. K., Gupta, A., Sachdeva, M. P., Huirem, R. S., Chauhan, K., Yashavarddhan, M. H., Basnal, A., Gupta, R., Mallick, P. K., & Ganguly, N. K. (2023). Donor-Derived Cell-Free DNA as a Non-Invasive Biomarker for Graft Rejection in Kidney Transplant Recipients:

A Prospective Study among the Indian Population. *Diagnostics*, 13(23), 3540. https://doi.org/10.3390/diagnostics13233540

Lakhani, L., Alasfar, S., Bhalla, A., Aala, A., Rosenberg, A., Ostrander, D., Schollenberger, M. D., Brennan, D. C., & Lipson, E. J. (2021). Utility of Serial Donor-derived Cell-free DNA Measurements for Detecting Allograft Rejection in a Kidney Transplant Recipient After PD-1 Checkpoint Inhibitor Administration. *Transplantation Direct*, 7(2), e656. https://doi.org/10.1097/TXD.0000000000001113

Loupy, A., Certain, A., Tangprasertchai, N. S., Racapé, M., Ursule-Dufait, C., Benbadi, K., Raynaud, M., Vaskova, E., Marchis, C., Casas, S., Hague, T., Bestard, O., Kervella, D., Lefaucheur, C., Viard, T., & Aubert, O. (2024). Evaluation of a Decentralized Donor-Derived Cell-Free DNA Assay for Kidney Allograft Rejection Monitoring. *Transplant International*, 37, 13919. https://doi.org/10.3389/ti.2024.13919

Mantios, E., Filiopoulos, V., Constantoulakis, P., Liapis, G., Vittoraki, A., Casas, S., Marinaki, S., & Boletis, J. N. (2023). Assessment of Donor Derived Cell Free DNA (dd-cfDNA) at Surveillance and at Clinical Suspicion of Acute Rejection in Renal Transplantation. *Transplant International*, 36, 11507. https://doi.org/10.3389/ti.2023.11507

Mayer, K. A., Doberer, K., Tillgren, A., Viard, T., Haindl, S., Krivanec, S., Reindl-Schwaighofer, R., Eder, M., Eskandary, F., Casas, S., Wahrmann, M., Regele, H., & Böhmig, G. A. (2021). Diagnostic value of donor-derived cell-free DNA to predict antibody-mediated rejection in donor-specific antibody-positive renal allograft recipients. *Transplant International*, 34(9), 1689–1702. https://doi.org/10.1111/tri.13970

Moein, M., Settineri, J. P., Suleiman, H., Sidhu, J., Papa, S., Coyle, S., Dvorai, R. H., Bahreini, A., Leggat, J., & Saidi, R. F. (2024). Application of Combined Donor-Derived Cell-Free DNA and Transcriptome in Diagnosis of Kidney Transplant Rejection. *Transplantation Proceedings*, 56(6), 1259–1263. https://doi.org/10.1016/j.transproceed.2024.02.023

Montgomery, R., & Gelb, B. (n.d.). *Four Things You May Not Know About Kidney Transplants*. NYU Langone News. Retrieved August 11, 2025, from https://nyulangone.org/news/four-things-you-may-not-know-about-kidney-transplants

Nguyen, P. T., Nakaoka, H., Mitsunaga, S., Aoyama, H., Kitamura, H., Saigo, K., & Inoue, I. (2025). *Applications of donor-derived cell-free DNA in kidney*

transplantation healthcare: View from a prospective single-center study. Research Square. https://doi.org/10.21203/rs.3.rs-6592444/v1

Nie, W., Wang, Y., Fu, Q., Wu, C., Deng, R., Yu, X., Ye, C., Liu, X., Xu, B., Sun, P., Liu, L., Li, J., Zhang, H., & Wang, C. (2025). Integrating Donor Derived Cell-Free DNA Fraction and Absolute Quantification for Enhanced Rejection Diagnosis in Kidney Transplant Recipients. *Diagnostics*, 15(3), 237. https://doi.org/10.3390/diagnostics15030237

Oellerich, M., Sherwood, K., Keown, P., Schütz, E., Beck, J., Stegbauer, J., Rump, L. C., & Walson, P. D. (2021). Liquid biopsies: Donor-derived cell-free DNA for the detection of kidney allograft injury. *Nature Reviews Nephrology*, 17(9), 591–603. https://doi.org/10.1038/s41581-021-00428-0

Orandi, B. J., Luo, X., Massie, A. B., Garonzik-Wang, J. M., Lonze, B. E., Ahmed, R., Van Arendonk, K. J., Stegall, M. D., Jordan, S. C., Oberholzer, J., Dunn, T. B., Ratner, L. E., Kapur, S., Pelletier, R. P., Roberts, J. P., Melcher, M. L., Singh, P., Sudan, D. L., Posner, M. P., … Segev, D. L. (2016). Survival Benefit with Kidney Transplants from HLA-Incompatible Live Donors. *New England Journal of Medicine*, 374(10), 940–950. https://doi.org/10.1056/NEJMoa1508380

Organ transplants exceeded 48,000 in 2024; a 3.3 percent increase from the transplants performed in 2023—OPTN. (n.d.). Retrieved August 11, 2025, from https://optn.transplant.hrsa.gov/news/organ-transplants-exceeded-48-000-in-2024-a-33-percent-increase-from-the-transplants-performed-in-2023/

Parajuli, S., Garg, N., Dodin, B., Breyer, I., Zona, E., Patel, S., Pinney, K., & Mandelbrot, D. (2024a). Changes in Donor-Derived Cell-Free DNA Before and After Rejection and De Novo DSA Detection in Primary and Repeat Kidney Transplant Recipients. *Clinical Transplantation*, 38(11), e70019. https://doi.org/10.1111/ctr.70019

Parajuli, S., Garg, N., Dodin, B., Breyer, I., Zona, E., Patel, S., Pinney, K., & Mandelbrot, D. (2024b). Changes in Donor-Derived Cell-Free DNA Before and After Rejection and De Novo DSA Detection in Primary and Repeat Kidney Transplant Recipients. *Clinical Transplantation*, 38(11), e70019. https://doi.org/10.1111/ctr.70019

Publish or Perish. (n.d.). Retrieved August 11, 2025, from https://harzing.com/resources/publish-or-perish

Quick kidney disease facts and stats | American Kidney Fund. (2024, April 19). https://www.kidneyfund.org/all-about-kidneys/quick-kidney-disease-facts-and-stats

Rizvi, A., Faiz, S., Thakkar, P. H., Hussain, S., Gamilla-Crudo, A. N., Kueht, M., & Mujtaba, M. A. (2023a). Kidney Allograft Monitoring by Combining Donor-Derived Cell-Free DNA and Molecular Gene Expression: A Clinical Management Perspective. *Journal of Personalized Medicine*, 13(8), 1205. https://doi.org/10.3390/jpm13081205

Rizvi, A., Faiz, S., Thakkar, P. H., Hussain, S., Gamilla-Crudo, A. N., Kueht, M., & Mujtaba, M. A. (2023b). Kidney Allograft Monitoring by Combining Donor-Derived Cell-Free DNA and Molecular Gene Expression: A Clinical Management Perspective. *Journal of Personalized Medicine*, 13(8), 1205. https://doi.org/10.3390/jpm13081205

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021a). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021b). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021c). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Schenk, K. M., Deutsch, J. S., Chandra, S., Davar, D., Eroglu, Z., Khushalani, N. I., Luke, J. J., Ott, P. A., Sosman, J. A., Aggarwal, V., Schollenberger, M. D., Sharfman, W. H., Bibee, K. P., Scott, J. F., Loss, M. J., Wang, H., Qi, H., Sharon, E., Streicher, H., … Lipson, E. J. (2024). Nivolumab + tacrolimus + prednisone +/− ipilimumab for kidney transplant recipients with advanced cutaneous cancers. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, 42(9), 1011–1020. https://doi.org/10.1200/JCO.23.01497

Sharma, R., Wolff, J., & Wilpula, E. (2022). Novel Technology Has Made It Possible To Diagnose A New Clinical Phenotype Of Rejection called Early Antibody Mediated Rejection, With Significant Therapeutic Implications For Renal Allograft Survival. *Experimental and Clinical Transplantation*, 20(10), 967–969.

https://doi.org/10.6002/ect.2021.0489

Shen, J., Guo, L., Yan, P., Zhou, J., Zhou, Q., Lei, W., Liu, H., Liu, G., Lv, J., Liu, F., Huang, H., Dong, W., Shu, L., Wang, H., Wu, J., Chen, J., & Wang, R. (2020). Prognostic value of the donor-derived cell-free DNA assay in acute renal rejection therapy: A prospective cohort study. *Clinical Transplantation*, 34(10), e14053. https://doi.org/10.1111/ctr.14053

Stites, E., Kumar, D., Olaitan, O., John Swanson, S., Leca, N., Weir, M., Bromberg, J., Melancon, J., Agha, I., Fattah, H., Alhamad, T., Qazi, Y., Wiseman, A., & Gupta, G. (2020). High levels of dd-cfDNA identify patients with TCMR 1A and borderline allograft rejection at elevated risk of graft injury. *American Journal of Transplantation*, *20*(9), 2491–2498. https://doi.org/10.1111/ajt.15822

Tian, Y., Frischknecht, L., Anna Mallone, Rossler, F., Schachtner, T., & Nilsson, J. (n.d.). *Evaluation of de novo donor specific antibodies after kidney transplantation in the era of donor-derived cell-free DNA - PMC*. Retrieved August 11, 2025, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11779610/

Wei, L., Zhao, Y., Deng, S., Wu, S., Wang, H., Luo, X., & Yang, H. (2024). Graft-derived cell free DNA: Used for assessment of early graft status and its implications for long-term kidney function. *Frontiers in Physiology*, *15*, 1440799. https://doi.org/10.3389/fphys.2024.1440799

Wolf-Doty, T. K., Mannon, R. B., Poggio, E. D., Hinojosa, R. J., Hiller, D., Bromberg, J. S., & Brennan, D. C. (2021). Dynamic Response of Donor-Derived Cell-Free DNA Following Treatment of Acute Rejection in Kidney Allografts. *Kidney360*, *2*(4), 729. https://doi.org/10.34067/KID.0000042021

Zhang, H., Zheng, C., Li, X., Fu, Q., Li, J., Su, Q., Zeng, L., Liu, Z., Wang, J., Huang, H., Xu, B., Ye, M., Liu, L., & Wang, C. (2020a). Diagnostic Performance of Donor-Derived Plasma Cell-Free DNA Fraction for Antibody-Mediated Rejection in Post Renal Transplant Recipients: A Prospective Observational Study. *Frontiers in Immunology*, *11*, 342. https://doi.org/10.3389/fimmu.2020.00342

Zhang, H., Zheng, C., Li, X., Fu, Q., Li, J., Su, Q., Zeng, L., Liu, Z., Wang, J., Huang, H., Xu, B., Ye, M., Liu, L., & Wang, C. (2020b). Diagnostic Performance of Donor-Derived Plasma Cell-Free DNA Fraction for Antibody-Mediated Rejection in Post Renal Transplant Recipients: A Prospective Observational Study. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.00342

Submission title: Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection

Date of Review: 8/27/25

Decision: Revise and resubmit (major revisions needed, acceptance not guaranteed)

Comments:

- Originality & Significance – Does the paper contribute new insights or perspectives to the field?
    - Yes. Very interesting points made and clear path of innovation to the field of renal transplant
- Clarity & Structure – Is the argument well-organized and easy to follow? Are ideas clearly presented?
    - Flow and presentation of information can be significantly improved.
- Use of Evidence & Research Methods – Are sources appropriately cited? Is their methodology sound and well-explained?
    - Yes. Very clear, excellent work
- Engagement with Literature – Does the paper demonstrate an understanding of relevant research in the field? Do they acknowledge known results and connect their findings well to them?
    - There is a clear understanding of relevant research but there is little connection of current research to the analysis performed in this review. For example, are there other literature reviews/meta-analyses that engage with a similar topic and how do their findings relate to the findings of this manuscript?
- Grammar & Language – Is the writing clear and professional? Minor grammatical and stylistic errors should be noted, but they should not be the main focus of the review.
    - Stylistic/grammar errors are noted throughout the manuscript
    - Figures are incorrectly referenced throughout the manuscript

Suggested revisions:

- Abstract
    - Usage of the word “good” throughout does not give descriptive information, consider using a more descriptive adjective
    - “criteria created by the author”
- Introduction
    - “In short, these significant barriers…” this sentence is wordy, try to make more concise. Consider choosing “unfulfilled need space” or “unmet medical need”
    - 4th paragraph of introduction is typically included in the discussion or conclusion sections. Avoid including results or analysis in the introduction
- Background
    - Background sections are not typically included in scientific/clinical manuscripts as the relevant background should be included early in the introduction
    - Avoid using paratheses out of usage for sources/citations
    - Background section does not discuss dd-cfDNA to the extent of the other non-invasive methods but was stated as the main objective of the literature review. Is cfDNA the same as dd-cfDNA? This is not clear but there is a section on general cfDNA. Consider shortening other sections to lengthen background of dd-cfDNA specifically
- Methods
    - Inclusion criteria of Jan 2000-Jan 2025 does not align with the search process restriction of Jan 2020-May 2025
    - Was the study scoring system created by the author or adapted from an accepted approach? Please clarify.
- Literature summary section
    - Unsure of the usage of this section. Consider including this highly specific information in the results section and only use information that is important to the overall thesis of evaluating dd-cfDNA
- Analysis
    - This section is more aptly titled “Discussion” as there is interpretation of relevant results
    - Consider using terms such as “prospective cohort” design; overall being more specific about the design of the papers used in the search (as done in Figure2e/f)
    - Excellent breakdown of the designed 5-point scoring system.
    - Figure 2a does not report overall evidence base being high quality but rather publication year of selected papers. Please revise.
        - Figures are incorrectly referenced throughout the analysis section.
    - Unsure of the usage of Figure 4 as it does not relate to the research hypothesis of dd-cfDNA. This information is interesting but not useful enough to have a dedicated figure or paragraph.
    - Description of AUC and ROC is not usually in discussion sections but can be moved to the methods section.
    - Discussion of Figure 5 is very interesting, great work.
    - Discussion of ABMR and TCMR is very important in the field of transplant, should be included earlier in the discussion/analysis section
- Limitations:
    - Paragraph three is not expected in the limitations section but best fitted at the end of the discussion section or as a stand-alone conclusion paragraph
    - Limitations are usually included at the end of the discussion section rather than as its own section

    - Consider condensing this section into highly salient points. Should be no more than 1 paragraph (~7-8 sentences)
- Future directions and Recommendations
    - Future directions are usually included at the end of the discussion section rather than as its own section
    - Consider condensing this section into highly salient points specific to future driections rather than continued interpretation of study results. Should be no more than 1 paragraph (~7-8 sentences)
    - The third paragraph is not expected in future directions section, best suited for end of discussion section or conclusion
- Conclusion section
    - Well done, clear, concise.
- Overall comments:
    - There is no results section, which is expected of a clinical/medical manuscript. Results sections provide objective information.
    - There is no Table 1 to describe the studies selected, which is expected of a clinical/medical manuscript. Figure 2 attempts to convey this information visually but should also be compiled into a Table 1, therefore defeating the purpose of Figure 2. See published literature reviews or meta-analyses for examples.
    - Summary or conclusion-adjacent paragraphs are littered throughout this manuscript. All summary paragraphs should be revised into one conclusion section
    - Consider including a flow chart figure to describe the paper selection process visually rather than multiple bar charts.
    - Figure 1 is excellent and provides great visual information. Great usage of BioRender.

I have reviewed the following manuscript under consideration for publication, titled "Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection"

Overall, this paper is insightful, original, and significant to the body of literature. The objectives of the paper are well-stated and the paper is well-organized. The methodology of the paper is . The author should consider using a widely accepted metric of quality assessment to compare rigor of papers and more information regarding the statistical methods and the outcomes of statistical tests should be included. Additionally, the results section should include more information regarding the content and structure of the papers included in this analysis so that the reader can understand the quality and type of studies included in the analysis. The grammar and language is clear, professional, and concise. In summary, this paper is applicable to the field and answers a very relevant question. I recommend: Acceptance with major revisions (major revisions bolded below). Please find my comments below.

Abstract:

- Include more information about the search (databases, how many studies were reviewed vs. included, etc.)
- Give a little background information about the criteria created by the author to assess quality
    - You may consider using a widely utilized quality metric for ease of understanding by the audience
- You state that this is a usable technique but don't state that your review assessed the feasibility of using this in clinical practice
- Great job identifying the gap in the literature and the identification of future directions

Introduction:

- Unclear what you mean by "poor risk management framework" - this should be expanded to make it very clear the problems that exist with traditional monitoring for rejection
- Include some background information about the practical steps of dd-cfDNA – is there something that needs to be measured from the donor at transplantation time? How is it detected?
- **You identified that one issue with traditional surveillance is that it doesn't identify the issue until the tissue/organ damage is done. When you introduce dd-cfDNA you state that these cells are released into circulation when the organ is inflamed or injured – is it unclear how this is different from the traditional method. Is there a baseline level that one would expect? Would the provider be monitoring for increases?**
- Great job stating the hypothesis and the objectives of the study

- The summary of the findings of the review should be moved from the introduction to the discussion/conclusion
- More information on the current state of research in this field – are there other reviews to date?

Background:

- You state the kidney biopsy is the gold standard for monitoring graft health which contradicts the statements in your introduction that biomarkers are used for monitoring. The role of kidney biopsy vs. biomarkers should be clarified.
- Swap the order of discussion in the background section - discussing the kidney transplant and reject background before discussing how it is monitored makes more sense
- Add statistics regarding the morbidity and mortality of ESRD / kidney transplant to the background to make it clear the severity of this disease and how many people it affects
- Figure 1 is blurry and the text cannot be read. Also, is this an original figure? If so, you should credit how and by whom it was created. If not, a citation is required.
- You make the point in your introduction that current biomarkers can only detect problems once the damage is done. It is unclear how this is different as it is only released with cell damage. I would clarify this as it speaks to the utility of this method
- Given the short half-life of cfDNA – is there a possibility of it rising and falling and not being detected?

Methods:

- Good overview of the search terms and search methods
- Did you include/exclude literature review, systematic reviews,metaanalysis?
- **Unclear how many studies were identified from the search, when/how many were excluded at different stages of the process. Consider using a PRISMA flow diagram to describe this**
- **Criteria stated do not fully assess study methodological rigor as stated - you should consider using a widely accepted method**

Data Analysis:

- Were there any comparisons made between papers? I would make it clear how this was done.

Literature Summary:

- The overview provided here is a bit redundant as you have introduced this topic in the introduction and background section. I recommend consolidating the information here

- The “evaluating the efficacy of dd-cfDNA” section belongs in the methods sectional as it discusses how you evaluated the performance
- I would introduce the concept of the Dd-cfDNA 30 minute half-life and how this is used in clinical practice, in the introduction/background sections
- **You need a more rigorous overview of studies included (sample size, methods, multicenter vs single center, comparison methods, etc.) and their main findings overall before discussing the most key findings overall. This section is brief and doesn’t give a comprehensive summary of the papers that are included / will be analyzed in this paper. You may also consider including a table 1 that discusses all of the studies included, their methods,and findings that were compared**

Analysis:

- You discuss a steady increase over the years however, it isn’t clear if there were any statistical tests employed. Please clarify how you determined that the increase was steady overtime. A test comparing the numbers over the years and a p-value would be great here.
- **It seems like some of your figures have confidence intervals however, there is no mention of the statistical test used to compare and how you determined increase/decreases over time. Please consider employing statistical tests for these comparisons to make the analysis more rigorous**
    - To this point, the confidence intervals appear to be overlapping – I don’t believe that you can say that there has been an increase in sample size over time if the difference is not statistically significant. This should be made clear to the reader
- Consider restructuring this sentence – “The choice to go through the analysis this way was significant because this page had a demonstrated blend of complementary types of research to provide a wider and more diverse evidence base - the prospective studies provide higher quality-controlled data while the retrospective and observational studies filled in the gap of real-word evidence, which in sum produced better conclusions for this review.”
- Again, when comparing distribution by location, please employ a statistical test
- Consider using a widely accepted method for quality assessment as your assessment tool does not assess some key components including - study design, bias assessment, etc.
- You make this statement – “Despite this, there appears to be a large methodological gap in the literature regarding timing of rejection.” The data on this outcome should be reported with the other results and the methods should be updated to include this outcome
- Several results are stated multiple times (for example, the racial/ethnic distribution of the data) – recommend reviewing this section for redundancy
- The description of the AUC and ROC belongs in the methods section as opposed to the results section.

- Please review this claim – “An interesting finding is that studies do not report mediocre performance and if this is potentially influenced by publication bias of studies with negative findings, it may at least suggest that when dd-cfDNA is investigated, it is consistently a good and effective biomarker for detection of allograft rejection. “ If you believe that negative results are being underreported, it actually suggests that dd-cfDNA may be overestimated in its ability to detect rejection if negative results are not being reported
- AUC results are discussed twice in this section without new information being added - consider consolidating.
- Be careful stating that certain tests are more useful in certain racial/ethnic populations as race may be a proxy for social factors. Additionally, race-based medicine has implications for patient care that can lead to disproportionate access to resources (such as transplant) as occurred with eGFR
    - Additionally, the confidence intervals in all groups appear to overlap so the differences appear statistically insignificant. Again, please describe the statistical testing used here.
- This section is the first time you discuss analysis of different types of rejection (AMBR vs. TCMR) – this concept and its significance should be discussed in the background section and the methods of comparison should be discussed above as well.

Limitations

- You should discuss the underreporting of negative studies
- You should discuss the translability and applicability of this test in clinical practice
- Were there limitations in the study design besides the generalizability. As stated above, please include characteristics of the studies included and the pros and cons of those studies should be discussed here

Future Directions

- You discuss the difference in accuracy by different commercial assays, but this is not discussed elsewhere in the review. This should be discussed as it may be a major limitation.
    - Figure 8d is referenced here but it is not present in the analysis section.
- **You state that more research is necessary to determine the frequency of monitoring – information on how the studies monitored using this metric is imperative to being able to compare them. This should be discussed in the literature overview.**
- Stating that dd-cfDNA is more accurate than creatinine seems to be an overstatement as no direct comparison between the two methods is discussed in the analysis section; you only state the diagnostic accuracy of dd-cfDNA only. As this is one of the stated outcomes of this study, I would discuss how dd-cfDNA compares to creatinine, how they

were compared in the original study, and how you determined that dd-cfDNA > creatinine.

Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection

Author 1

September 6, 2025

Abstract

The ability to monitor kidney allografts is limited by traditional forms of assessment like invasive biopsy, and also delayed biomarkers, such as serum creatinine (Rizvi et al., 2023a). This systematic review was performed to assess the clinical usefulness of donor-derived cell-free DNA (dd-cfDNA) as a non-invasive biomarker of allograft injury. ~~An extensive search identified 83 studies, each assessed for methodological quality.~~A systematic search of the Google Scholar and PubMed databases yielded 1,153 records; following a manual screening of titles and abstracts, 83 studies were selected for inclusion. Results indicate a rapidly maturing evidence base, including an increase in publications and a corresponding increase in sample size over time, reflected most notably in a landmark study on nearly 3,000 patients (Aubert et al., n.d.). The literature generally met a set of curated criteria for non-invasive biomarker studies derived by the author. For example, most studies obtained 3 or more out of 5 criteria: “Measured Blood?”, “Analyzes rejection probability?”, “Temporal Precedence Explicitly Stated”, “Longitudinal Monitoring for Timing”, and “Compares to Traditional Biomarkers?”. A key finding of the review was the consistent and generally ~~good~~favorable diagnostic accuracy of dd-cfDNA, with most studies indicating an Area Under the Curve (AUC) greater than 0.80 (Mantios et al., 2023). There is, however, a stark geographical disparity in the inclusion of diversified ancestry; the available evidence for Black and Hispanic ancestry is largely based on North American evidence (Bu et al., 2022). ~~This systematic review provides confirmatory evidence that dd-cfDNA is a strong and usable non-invasive biomarker~~This systematic review provides confirmatory evidence that dd-cfDNA is a strong diagnostic biomarker (Kim et al., 2024a)~~, but requires future work with a particular focus on the design of well-constructed large, multicenter studies, focusing on diversity to ensure equitable and optimal application of this promising tool in clinical practice~~; however, it did not assess clinical feasibility. Future work

must address this gap by focusing on cost-effectiveness and implementation studies, in addition to conducting large, multicenter trials in diverse populations to ensure its equitable and optimal application (Botella et al., 2024a).

Introduction

Kidney transplantation is a vital therapy for patients with end-stage kidney disease, providing a better overall longevity and improved quality of life compared to long-term dialysis (Dreige et al., 2022), which is the non-surgical alternative to transplantation. Despite its advantages and evidence for superior long-term outcomes, transplantation faces the ongoing challenge of immune-mediated rejection, which remains a prominent and often principal cause of patient graft-loss (Lakhani et al., 2021). The challenge is exacerbated by a poor risk management framework that relies on conventional surveillance practices that typically are complicated and inaccurate to true quantitative values (Wei et al., 2024). This framework is limited by its dependence on invasive and reactive kidney biopsies and the use of non-specific, lagging biomarkers like serum creatinine that often only rise after irreversible graft damage has occurred (Parajuli et al., 2024a; Rizvi et al., 2023a). Traditionally, markers for the identification and management of rejection have been insufficient due to the slow response of these traditional markers to tissue injury (Rizvi et al., 2023a). In short, these significant barriers represent ~~the unfulfilled need space, or unmet medical need, in transplant medicine, because~~ the unmet medical need in transplant medicine, as they do not allow clinicians to intervene prior to true tissue/organ damage (Nie et al., 2025).

In recent years, a new solution has been developed to solve these deficiencies: the measurement of dd-cfDNA (Zhang et al., 2020a), small pieces of DNA from the donor that are released from the transplanted organ into the circulation, especially when the organ is activated, or inflamed, or injured through histocompatibility, or rejection (Zhang et al., 2020b). When a person has an organ transplant, they also are a recipient of a donor's genetic information at the time of the transplant (Kumar et al., 2023). Quantifying the amount of dd-cfDNA in a blood sample can now provide a direct, non-invasive and organ-specific signal of graft injury (Mayer et al., 2021). This has been depicted as a pioneering biomarker with much potential to detect graft damage early (Botella et al., 2024a). The clinical utility of dd-cfDNA stems from its short half-life of about 30 minutes, which allows it to serve as a near real-time indicator of acute allograft injury. When the transplanted kidney is healthy,

dd-cfDNA is released at a low, baseline level. However, during an active rejection event, inflammation and cell death in the allograft cause a significant and measurable increase in the amount of dd-cfDNA released into the recipient's bloodstream. By quantifying this increase through techniques like next-generation sequencing (NGS), clinicians can detect rejection far earlier than with traditional markers, creating an opportunity for pre-emptive treatment before the onset of widespread tissue damage.

Though transplant surgeons are currently utilizing dd-cfDNA in clinical practice, the development of the literature is so fast-paced that a comprehensive review of the evidence is needed (Chen et al., 2022). It is no longer sufficient to say that dd-cfDNA is "effective" and the key question to focus on is how effective it is in specific clinical contexts. The high diagnostic accuracy of dd-cfDNA is consistent with previous literature reviews and meta-analyses, which have affirmed the biomarker's potential for non-invasive monitoring. Previous meta-analyses have reported high pooled Area Under the Curve (AUC) values, such as 0.84 for acute rejection, 0.86 for overall rejection, and 0.88 for antibody-mediated rejection. This review builds on that consensus by providing a quantitative synthesis of the field's maturation, documenting the significant increase in study sample sizes and the shift towards more rigorous prospective cohort designs. It also addresses a critical knowledge gap by quantifying the significant geographic and ancestral disparities, highlighting that evidence for Black and Hispanic populations is almost exclusively derived from North American studies. However, a limitation of these existing reviews is that they confirm the biomarker's general utility without systematically analyzing the characteristics of the rapidly evolving evidence base or quantifying the critical knowledge gaps related to its performance in diverse patient populations.

The unique value of this systematic review is that, in addition to synthesizing diagnostic accuracy, it provides the first quantitative analysis of the field's methodological maturation and its significant demographic and geographic gaps (Sharma et al., 2022); it is guided by the specific hypothesis that donor-derived cell-free DNA (dd-cfDNA) can result in increased diagnostic accuracy in detecting transplant rejection and graft injury versus current standard of care biomarkers, through earlier and non-invasive methods. The focus of this review is to (1) specify and quantify overall diagnostic accuracy of dd-cfDNA for various types of rejection,

(2) create and analyze data visualizations to synthesize the characteristics and findings of the included studies, (3) identify trends in performance and address key knowledge gaps, such as the biomarker's utility in diverse, underrepresented patient populations (Benning et al., 2022), and (4) assess the methodological approaches used in the literature to evaluate the timing of rejection.

~~The findings of this systematic review confirm that dd-cfDNA is a useful measure of allograft health (Kim et al., 2024a) and that evidence confirms that dd-cfDNA is a useful biomarker for early allograft injury (Tian et al., n.d.), and that dd-cfDNA can be used as a non-invasive biomarker (Nie et al., 2025). Moreover, the data show that dd-cfDNA has an indication of prognostic value for long-term allograft outcomes (Loupy et al., 2024). However, this systematic review also identified that some of the evidence base is derived from single-center cohorts or those with limited ethnic diversity (Nguyen et al., 2025). The most significant findings of this review include the quantification of dd-cfDNA's high diagnostic accuracy, with AUC values consistently exceeding 0.80 (Rizvi et al., 2023a), and its ability to detect rejection months before clinical presentation, such as reducing the time to AMR diagnosis to just 2.8 months (Akifova et al., 2025). This review revealed the biomarker's significant prognostic value, with early elevations of dd-cfDNA (≥1%) being associated with a more than five-fold greater rate of future eGFR—estimated Glomerular Filtration Rate, a key measure of how well the kidneys are filtering waste from the blood decline,—(21.4% vs. 4.1%) (Sawinski et al., 2021a). Going forward, it is recommended that future research focuses on head-to-head comparisons of commercial assays to standardize protocols and on conducting large, prospective trials in more diverse patient populations to ensure the equitable and optimal application of this biomarker in clinical practice.~~

~~The articles in this systematic review generally intended to assess the clinical potential and utility of dd-cfDNA as a viable, non-invasive, and convenient (it requires only a standard blood draw) quantifiable biomarker the monitor the health and overall condition of kidney allografts (Wei et al., 2024). The relevance of dd-cfDNA can be assessed comparatively to other standard healthcare metrics such as serum creatinine and proteinuria (Bu et al., 2022). Most of the articles focused on and assessed dd-cfDNA's ability to identify allograft injury sooner (Bromberg et al, 2024), with many articles constructed to quantify exactly how much sooner dd-cfDNA may identify rejection in comparison to other biomarkers (Parajuli et al.,~~

2024b) or changes in lagging biomarkers like serum creatinine (Schenk et al., 2024). In order to do this, articles evaluated dd-cfDNA's ability to identify both acute and other subclinical forms of rejection (Bu et al., 2022) alongside its ability to discriminate between antibody mediated rejection (AMBR) and T cell mediated rejection (TCMR) (Bromberg el al, 2024). Additionally, a number of articles investigated dd-cfDNA in the longitudinal monitoring of patient outcomes, not only for pre-rejection surveillance (Mantios et al., 2023) but also for assessing their response to treatment (Wolf-Doty et al., 2021).

Background

The transplantation field is actively looking for new, non-invasive, and accurate ways to monitor graft health, given the many challenges with established surveillance modalities (Rizvi et al., 2023b). The gold standard now, kidney biopsy, is invasive, with procedural and other risk and discomfort for the patient. For the transplant physician, biopsy has limitations as it is a reactive measure that does not suit the frequent monitoring needed to prevent damage (Wei et al., 2024). Serum creatinine, the most widely used non-invasive biomarker, is a surrogate, vague and non-specific biomarker. Thus, when creatinine rises, there can be a significant degree of already irreversible graft damage (Parajuli et al., 2024a). Cell free DNA (cfDNA) refers to all small fragments of DNA found circulating in the bloodstream, which come from various cells throughout the body as they die and break down. After a transplant, a fraction of this total cfDNA originates from the donated organ; this specific component is known as donor-derived cell-free DNA (dd-cfDNA). While all cfDNA can indicate cell death, dd-cfDNA is a highly specific biomarker because it isolates the signal coming directly from the transplanted organ, providing a clearer view of its health. While biopsy serves as the definitive diagnostic tool, its invasiveness and the lagging nature of conventional biomarkers make them unsuitable for the frequent, proactive monitoring needed to prevent irreversible graft damage.

Recent advancements have uncovered the study of donor-derived cell-free DNA (dd-cfDNA) as an alternative solution that directly counters these flaws by offering a non-invasive and accelerated method for assessing transplant status, potentially transforming post-transplant care (Botella et al., 2024a). The following results will synthesize the evidence to quantify the biomarker's accuracy and its timelines for early detection, thereby assessing its potential to become a new standard in

post-transplant surveillance (Bu et al., 2022). The clinical utility of dd-cfDNA stems from its short half-life of about 30 minutes, which allows it to serve as a near real-time indicator of acute allograft injury. When the transplanted kidney is healthy, dd-cfDNA is released at a low, baseline level. However, during an active rejection event, inflammation and cell death in the allograft cause a significant and measurable increase in the amount of dd-cfDNA released into the recipient's bloodstream. By quantifying this increase through techniques like next-generation sequencing (NGS), clinicians can detect rejection far earlier than with traditional markers, creating an opportunity for pre-emptive treatment before the onset of widespread tissue damage.

Kidney Transplantation and Rejection

Organ transplantation is a life-saving medical procedure by which a surgically incompetent organ from the recipient is removed and surgically replaced with a healthy organ obtained from a healthy donor (Orandi et al., 2016). Organ transplantation is important because it improves the quality of life for an organ recipient and increases life expectancy (Montgomery & Gelb, n.d.). Kidney transplantation is the most common organ transplantation because its underlying disease, end-stage renal disease (ESRD), is also one of the most common diseases, affecting >800000 persons in the United States (affecting >800,000 persons in the United States) (*Kidney Disease Statistics for the United States - NIDDK*, n.d.). The disease carries a significant burden, with a mortality rate of nearly 25% for patients on dialysis, highlighting the critical need for transplantation as a life-saving intervention. Conditions like diabetes and high blood pressure are the leading causes of kidney failure, impacting millions and contributing to the high demand for transplants (*American Kidney Fund*, 2024).

When an organ is transplanted, it is important to ensure the recipient's immune system will accept it and not recognize it as foreign; otherwise, an organ can be rejected (Nguyen et al., 2025). The most common immune responses are described as acute rejection, which happens shortly after the transplant has taken place, and chronic rejection, which occurs over time (Rizvi et al., 2023b). Ultimately, failure associated with acute and chronic rejection can be caused through progressive permanent damage in the transplanted organ with long-term changes in allograft function (Wei et al., 2024). The health consequences of kidney rejection for recipients are devastating with damage culminating in graft dysfunction such as a

~~reduction in eGFR or complete graft failure (Sawinski et al., 2021b).~~ Graft failure is a severe outcome, leading to a return to dialysis where the 5-year survival rate is only around 42%, underscoring the importance of effective allograft monitoring. Allograft rejection is primarily categorized into two distinct pathways: T-cell mediated rejection (TCMR) and antibody-mediated rejection (ABMR). TCMR is a cellular response where the recipient's T-cells directly attack the transplanted organ. ABMR is a humoral response driven by donor-specific antibodies (DSAs) that target the graft, leading to inflammation and tissue damage. Distinguishing between these subtypes is critical for guiding appropriate treatment, as they respond to different immunosuppressive therapies.

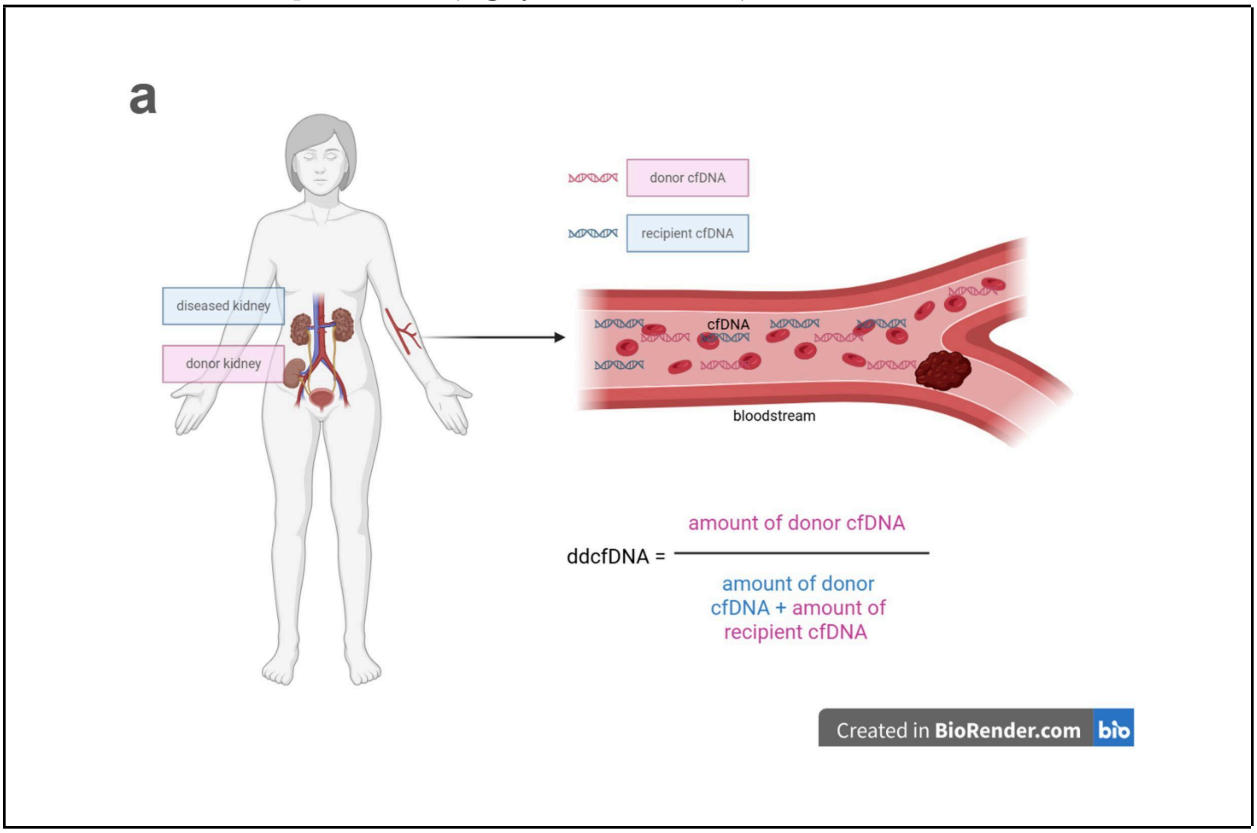
Currently, the monitoring of kidney transplant physiology is performed using a combination of clinical and biochemical monitoring protocols (Oellerich et al., 2021). Clinicians often take note of potential clinical signs or symptoms related to possible kidney dysfunction, such as decreased urinary output or developing edema in the legs and ankles from fluid overload (Kim et al., 2024a). However, to truly assess graft dysfunction, clinicians must perform invasive procedures to obtain tissue or organ biopsies, which allows for definitive diagnosis of the pertinent histopathological process (Moein et al., 2024).

The current monitoring modalities have multiple problems that can greatly influence patient care (Rizvi et al., 2023b). First, serum creatinine is a lagging and suboptimal marker; in fact, serum creatinine will only begin to elevate when potentially irreversible injury to the donated kidney has already occurred (Rizvi et al., 2023a). ~~Surgery is the acceptable gold standard for diagnosis of kidney disease; however, the kidney biopsy is invasive and carries considerable risk of bleeding, pain, and infection to the patient, netting small benefit to the patient (Cheng et al., 2021). Further, there is a risk of failing to recognize subtle or early signs of rejection and between clinical assessments/reporting, the potential chronic injury impacting long-term graft survival may occur with progressive injury (Bu et al., 2022).~~

Both the invasiveness of biopsies and the lagging sensitivity of serum creatinine are limitations in post-transplant management (Aubert et al., n.d.). There is a clear need for a noninvasive and improved biomarker for the kidney transplant rejection technique that can provide an early and accurate evaluation of graft health (Akalin et al., 2021).

Cell-free DNA use in transplantation

cfDNA, which consists of the short fragments of DNA that are free-circulating in the bloodstream, shows promise as a biomarker in transplantation research (Zhang et al., 2020b). Cell-free DNA originates when cells in the human body experience apoptosis, programmed cell death or necrosis, uncontrolled cell death (Oellerich et al., 2021). When cells die and degrade, the genetic information in the cells is released into the biological fluid surrounding the cells, including blood plasma, urine, and cerebrospinal fluid (Nguyen et al., 2025).



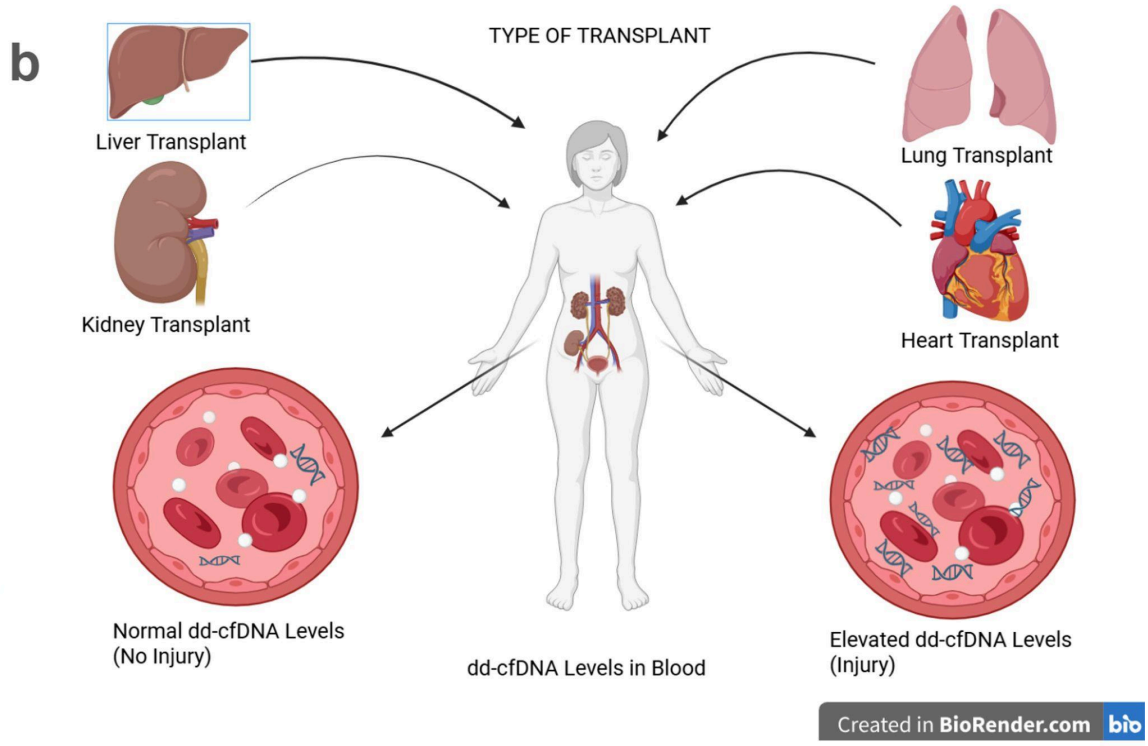


Figure 1: The Biological Principle of Donor-Derived Cell-Free DNA (dd-cfDNA) as a Biomarker. (a) This diagram illustrates that following a kidney transplant, both recipient and donor cfDNA circulate in the bloodstream. The fractional abundance of dd-cfDNA is calculated as the proportion of donor cfDNA relative to the total amount of cfDNA. This figure was created by the author using BioRender.com. (b) This diagram shows that dd-cfDNA is a general biomarker for all solid organ transplants, including liver, kidney, lung, and heart. In a healthy, uninjured state, dd-cfDNA levels are low; however, during an injury event like rejection, damaged cells from the transplanted organ release a higher amount of dd-cfDNA into the blood. This figure was created by the author using BioRender.com.

cfDNA comes from blood plasma, the fluid part of blood (Wolf-Doty et al., 2021). Several molecular methods are used currently to assess cfDNA quantity (Halloran et al., 2022). Both quantitative polymerase chain reaction (qPCR) and next-generation sequencing (NGS) can be viewed as modern methods for measuring cfDNA (Akalin et al., 2021). qPCR does work similarly to a targeted measure, where specific primers

are used to locate and amplify a pre-established and a small number of unique donor sequences (Nie et al., 2025). For NGS, millions of cfDNA fragments can be read concurrently across the genome (Kumar et al., 2023b). In this way, cfDNA can be assessed to identify the donor if they donated an organ (Wei et al., 2024).

Figure 1a provides biological context for utilizing dd-cfDNA as a biomarker. After a solid organ transplant, like kidney ~~(e.g. kidney)~~, the recipient's blood will contain their own cfDNA in addition to a small amount of cfDNA from the donor organ (Fig. 1a) (Oellerich et al., 2021). Under normal circumstances and with a healthy transplant, the amount of dd-cfDNA circulating in the blood is in a baseline state of low concentrations (Bu et al., 2022a). However, when an injury event causes a greater rate of cell death in the allograft, such as rejection, inflammation, or other forms of graft stress, the allograft will release dd-cfDNA into the recipient's circulation (Fig. 1b) (Rizvi et al., 2023b). These real-time changes in dd-cfDNA concentrations can reflect ongoing allograft injury (Nguyen et al., 2025).

cfDNA is a dynamic biomarker. It is constantly shed into the bloodstream, and the half-life of a single cfDNA molecule is approximately 30 minutes in the bloodstream (Wolf-Doty et al., 2021), meaning that it provides real-time information about tissue degeneration. Changes in specific cfDNA characteristics, most notably an extraordinary increase in levels of cfDNA, are indications of tissue injury, inflammation, or subsequently, a disease such as rejection, which validate the value of cfDNA as a non-invasive biomarker (Rizvi et al., 2023). This indicates that cfDNA is a valid source of non-invasive biomarker (Wei et al., 2024). The drive and excitement around a real-time multi-level signal for more precise management is the appeal of this biomarker that affirms its usefulness (Aubert et al., 2024). dd-cfDNA has become such a strong force and continues to be a focus of transplant research (Halloran et al., 2022).

The transplantation field is actively looking for new, non-invasive, and accurate ways to monitor graft health, given the many challenges with established surveillance modalities (Rizvi et al., 2023b). The gold standard now, kidney biopsy, is invasive, with procedural and other risk and discomfort for the patient. Serum creatinine, the most widely used non-invasive biomarker, is a surrogate, vague and non-specific biomarker. CfDNA refers to all small fragments of DNA found circulating in the bloodstream. At a healthy baseline, all individuals will have a

detectable level of cfDNA originating from normal cell death. However, after a transplant, a fraction of this total cfDNA originates from the donated organ; this specific component is known as donor-derived cell-free DNA (dd-cfDNA). While elevated total cfDNA indicates cell death somewhere in the body, the donor-derived fraction (dd-cfDNA) is a more specific biomarker for allograft health because its origin is exclusively the transplanted organ, thereby providing a clearer view of its health. Crucially, this signal reflects the real-time process of cell death as it happens, allowing for the detection of low-level or early-stage allograft injury, whereas conventional markers like serum creatinine only rise after substantial, cumulative, and often irreversible functional damage has already occurred. While biopsy serves as the definitive diagnostic tool, its invasiveness and the lagging nature of conventional biomarkers make them unsuitable for the frequent, proactive monitoring needed to prevent irreversible graft damage.

Recent advancements have uncovered the study of donor-derived cell-free DNA (dd-cfDNA) as an alternative solution that directly counters these flaws by offering a non-invasive and accelerated method for assessing transplant status, potentially transforming post-transplant care (Botella et al., 2024a). The following results will synthesize the evidence to quantify the biomarker's accuracy and its timelines for early detection, thereby assessing its potential to become a new standard in post-transplant surveillance (Bu et al., 2022). The clinical utility of dd-cfDNA stems from its short half-life of about 30 minutes, which allows it to serve as a near real-time indicator of acute allograft injury. When the transplanted kidney is healthy, dd-cfDNA is released at a low, baseline level. However, during an active rejection event, inflammation, and cell death in the allograft cause a significant and measurable increase in the amount of dd-cfDNA released into the recipient's bloodstream. By quantifying this increase through techniques like next-generation sequencing (NGS), clinicians can detect rejection far earlier than with traditional markers, creating an opportunity for pre-emptive treatment before the onset of widespread tissue damage.

Methods

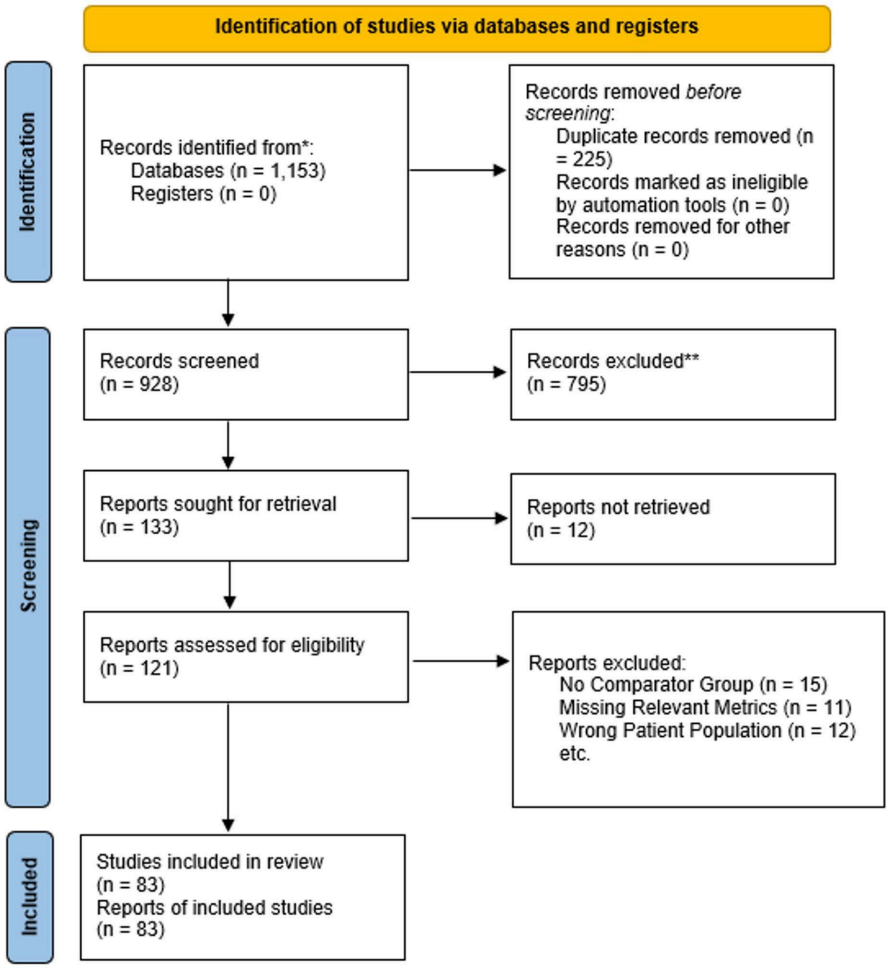


Figure 2. PRISMA 2020 Flow Diagram for Study Selection: This diagram illustrates the flow of information through the different phases of the systematic review, mapping out the number of records identified, included, and excluded.

Search Strategy

A comprehensive literature search was conducted using the Publish or Perish application (*Publish or Perish*, n.d.) to identify relevant articles. Searches were performed using the Google Scholar and PubMed databases. The search focused on publications related to donor-derived cell-free DNA (dd-cfDNA) and transplantation, utilizing the following keywords: "dd-cfDNA", "donor-derived cell-free DNA", "cell-free DNA renal transplant", "allograft surveillance biomarkers", "kidney transplant rejection", "biomarker acute rejection", "Kidney Transplantation"

(MeSH), "Graft Rejection" (MeSH), "Cell-Free Nucleic Acids" (MeSH), "Biomarkers" (MeSH), "Sensitivity and Specificity" (MeSH), "Comparative Effectiveness Research" (MeSH), "Diagnostic Imaging" (MeSH), "Creatinine/blood" (MeSH), "Monitoring, Immunologic" (MeSH), "diagnostic tool", "diagnostic performance", "surveillance method", "human subjects", "adult", "renal", "renal transplant", "kidney allograft", "diagnosis", "diagnostic", "AUC", "predictive value", "rejection", "sensitivity", "specificity", "performance", "comparison", "receiver operating characteristic curve", "serum creatinine", "biopsy", "imaging", "donor-specific antibodies", "DSA", "Racial Groups" (MeSH), "Ethnic Groups" (MeSH), "rejection monitoring", "allograft", "kidney".

These keywords were categorized to capture key aspects of the research question :

- **Population:** Keywords of "adult," "human subjects," "kidney transplant," "renal transplant," and "Kidney Transplantation" (MeSH) (including MeSH terms) were used to specify the patient group. "Racial Groups" (MeSH) and "Ethnic Groups" (MeSH) were included to address the "diverse ancestries" element.
- **Intervention:** Phrases such as "dd-cfDNA," "donor-derived cell-free DNA," "cell-free DNA renal transplant," and "Cell-Free Nucleic Acids" (MeSH) specifically targeted the primary diagnostic tool of interest.
- **Outcome:** Terms like "acute rejection," "graft rejection," "allograft rejection," and "Graft Rejection" (MeSH) directly identified the clinical event under investigation.
- **Comparison:** To encompass "standard biochemical and imaging surveillance," keywords like "serum creatinine," "biopsy," "imaging," "donor-specific antibodies," and related MeSH terms were utilized.
- **Diagnostic Role/Timing:** Phrases such as "diagnostic tool," "diagnostic performance," "surveillance method," "sensitivity," "specificity," "AUC," and "predictive value" were crucial for identifying studies assessing the diagnostic utility, accuracy, and particularly the comparative timing of dd-cfDNA.

The search process was restricted to articles published between January 1st, 2020, and May 31st, 2025.

Inclusion Criteria

Inclusion criteria were designed to bring in articles directly aligned with the research question:

English-language studies; Human subjects, specifically adult kidney transplant recipients (age ≥18 years); Use of donor-derived cell-free DNA (dd-cfDNA) as a primary or comparative diagnostic tool for acute rejection; Includes comparison to at least one standard surveillance method (e.g., serum creatinine, biopsy, imaging, donor-specific antibodies); Studies assessing the diagnostic role of dd-cfDNA in acute rejection, ideally reporting metrics (e.g., sensitivity, AUC) or qualitatively describing performance; Studies involving patients from diverse racial/ethnic backgrounds (where available); Studies published between January 1st, 2000 and ~~January~~May 31st, 2025; and original research articles only (literature reviews, systematic reviews, and meta-analyses were excluded).
Articles not meeting the inclusion criteria were excluded.

Selection Process

The resulting set of articles was exported into a Google Sheets document, and the complete selection process is detailed in the PRISMA flow diagram (Figure 2). Following this, a preliminary screening of manual reading of abstracts and titles was conducted to decide whether to include or exclude. The screening included: Reasoning as to why the article was chosen, Full Text Accessibility (only for texts included), Number of Citations, Authors, Title of Article, Year Published, Source, Publisher, Article URL, Cites URL, GS Rank, Query Date, Type of Article, DOI, ISSN, Citation URL, Volume, Issue, Start Page, End Page, ECC, Cites Per Year, Cites Per Author, Author Count, Age, and Abstract.

First Author	Year	Journal	Sample Size	Study Design	AUC for Rejection
Aubert, O.	2024	*Nature Medicine*	2,987	Prospective Cohort	0.87 (ABMR)
Bu, L.	2022	*Kidney International*	1,227	Prospective Cohort	0.81 (Any Rejection)
Halloran, P.F.	2022	*Transplantation*	385	Prospective Cohort	0.93 (ABMR)
Bromberg, J.S.	2024	*Transplantation*	191	Prospective Cohort	0.84 (Subclinical AR)
Stites, E.	2020	*Am. Journal of Transplant.*	231	Retrospective Cohort	0.82 (TCMR ≥1A)
Kim, H.D.	2024	*Frontiers in Immunology*	206	Retrospective Cohort	0.91 (Clinical ABMR)
Parajuli, S.	2024	*Clinical Transplantation*	148	Retrospective Cohort	Not Reported
Benning, L.	2023	*Transplant International*	104	Prospective, Single-Center	0.78 (Any Rejection)
Zhang, H.	2020	*Frontiers in Immunology*	129	Prospective Observational	0.94 (ABMR)
Akifova, A.	2025	*Nephrology Dialysis Transplantation*	184	Randomized Trial	0.85 (ABMR)
Chen, X.T.	2022	*Clinical Chemistry*	102	Prospective Cohort	0.89 (Any Rejection)
Gisch, N.	2025	*Am. Journal of Transplant.*	45 studies	Systematic Review	N/A
Yang, S.	2024	*Bosnian Journal of Basic Medical Sciences*	21 studies	Meta-Analysis	0.88 (ABMR)
Shen, J.	2020	*Clinical Transplantation*	68	Prospective Cohort	0.83 (Acute Rejection)
Mayer, K.A.	2021	*Transplant International*	115	Prospective Cohort	0.81 (ABMR)
Overall Summary	2020-2025	(Multiple)	N=83, Mean=245 Median=148 Range=21-2,987	(Multiple)	N=45, Mean=0.83 Median=0.82 Range=0.58-0.94

Table 1. Characteristics of Selected Included Studies. This table summarizes key information for a representative sample of the 83 studies included in the systematic review. The complete dataset used for this review is available in a Google Sheets document, which can be found in the "Code and Data Availability" section.

Data Extraction and Synthesis

The initial data extraction included standard article metadata such as GS Rank, Query Date, DOI, ISSN, Citations, Age of publication, and Author Count. For included articles, specific variables relevant to the research question were gathered. Gemini (Google LLC, 2025) was used to assist in extracting these variables, including: Title, Journal, Publication Year, Country, Study (single or multicenter), Study Type, Sample Size, Average Age, Proportion Male, Patient Race/Ethnicity, Type of Rejection, Type of cf-DNA, and AUC.

Study Scoring and ~~Quality~~Relevance Assessment

A scoring system was curated from the literature and specifically tailored to this review's research goals to assess the direct relevance of each included article. Articles were awarded one point (1) if the following criteria were met, and zero points (0) if not. This included: Measured Blood? Analyzes Rejection Probability? Temporal Precedence Explicitly Stated? Longitudinal Monitoring for Timing? Compared to Traditional Biomarkers?

These scoring criteria were curated by the author to assess each study's direct relevance in answering the research question. "Measured Blood?" confirmed the appropriate sample for dd-cfDNA, as blood plasma is the standard medium for this analysis (Wolf-Doty et al., 2021). "Analyzes Rejection Probability?" ensured focus on the primary clinical outcome of rejection (Bu et al., 2022). "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for Timing?" were crucial for evaluating the "earlier identification" aspect of the hypothesis, a key feature of dd-cfDNA's utility (Bromberg et al., 2024). Finally, "Compared to Traditional Biomarkers?" verified that the study included the necessary comparative analysis against standard surveillance methods like serum creatinine to establish the biomarker's superior performance (Mantios et al., 2023). Data extracted from the included articles included study type and design, specifically identifying if the research was a prospective cohort, retrospective, or observational study, along with other standard article metadata.

Data Synthesis

Articles receiving the high scores (3-5 points) were selected for in-depth analysis.

Preliminary conclusions were written from these high-scoring articles after reading them entirely and compiled into a summary document with a brief summary and overview of each article for future data synthesis and reference.

Data Analysis

Data analysis involved a narrative synthesis of findings from high-scoring articles. Extracted data on diagnostic timing, performance metrics, such as AUC and sensitivity, and patient demographics (including ancestry) were noted for future reference and figures. No formal statistical meta-analysis was performed due to the heterogeneity of the study populations and assays. Instead, comparisons between papers were made through a narrative synthesis of their findings, supported by data visualizations (e.g., box plots, scatter plots) that compare reported performance metrics and trends across different study contexts, such as geography and patient demographics.

The primary measure used to evaluate dd-cfDNA's diagnostic performance is the Area Under the Receiver Operating Characteristic (AUROC) curve. The AUROC is a valuable statistical measure that summarizes a test's ability to perform binary classification of patients with and without a condition across all possible thresholds; a value of 1.0 represents a perfect test, while a value of 0.5 represents a test with no discriminatory ability.

The Receiver Operating Characteristic (ROC) curve is a representation of a diagnostic test's overall performance, as described by one number, the Area Under the Curve (AUC). The AUC takes a value between 0.5 (no better than chance) and 1.0 (perfect test). Values between 0.80 and 0.90 may be considered "good". For the purposes of this systematic review, the AUC indicates the extent to which dd-cfDNA separates patients with biopsy-proven rejection (the cases) from patients without biopsy proven rejection (the controls); therefore, a high AUC indicates that the biomarker is quite accurate in identifying patients that are actively rejecting their organ.

Data Visualization

All figures presented in the Results section were generated using Google Colab, a cloud-based computational environment. Data extracted from the included studies

were compiled into a structured dataset and imported into the Colab environment for analysis. The Python programming language, along with the data manipulation library Pandas and the data visualization libraries Matplotlib and Seaborn, were used to create all plots, including histograms, box plots, and scatter plots, to visually synthesize the findings of this systematic review.

~~*Literature Summary*~~

~~*Overview*~~

~~The articles in this systematic review generally intended to assess the clinical potential and utility of dd-cfDNA as a viable, non-invasive, and convenient (it requires only a standard blood draw) quantifiable biomarker the monitor the health and overall condition of kidney allografts (Wei et al., 2024). The relevance of dd-cfDNA can be assessed comparatively to other standard healthcare metrics such as serum creatinine and proteinuria (Bu et al., 2022). Most of the articles focused on and assessed dd-cfDNA's ability to identify allograft injury sooner (Bromberg et al, 2024), with many articles constructed to quantify exactly how much sooner dd-cfDNA may identify rejection in comparison to other biomarkers (Parajuli et al., 2024b) or changes in lagging biomarkers like serum creatinine (Schenk et al., 2024). In order to do this, articles evaluated dd-cfDNA's ability to identify both acute and other subclinical forms of rejection (Bu et al., 2022) alongside its ability to discriminate between antibody mediated rejection (AMBR) and T cell mediated rejection (TCMR) (Bromberg el al, 2024). Additionally, a number of articles investigated dd-cfDNA in the longitudinal monitoring of patient outcomes, not only for pre rejection surveillance (Mantios et al., 2023) but also for assessing their response to treatment (Wolf-Doty et al., 2021).~~

~~*Evaluating the efficacy of dd-cfDNA in estimating rejection*~~

~~The first goal of the surveyed research was to quantify how dd cfDNA was distinguished between rejection states (Mantios et al., 2023), thereby guiding clinical decisions by providing a non-invasive tool that, due to its high specificity, could help avoid unnecessary invasive biopsies (Wei et al., 2024). The primary measure used to evaluate dd-cfDNA's diagnostic performance is the Area Under the Receiver Operating Characteristic (AUROC) curve. The AUROC is a valuable statistical measure that summarizes a test's ability to correctly classify patients~~

with and without a condition across all possible thresholds; a value of 1.0 represents a perfect test, while a value of 0.5 represents a test with no discriminatory ability.

Early Detection with dd-cfDNA

Additional research focused on the ability of dd-cfDNA to detect graft injury significantly earlier than clinical presentation (Bromberg et al., 2024) or elevations in standard biochemical markers like serum creatinine and proteinuria, which are known to be lagging (Kim et al., 2024b). Dd-cfDNA's 30 minute half-life allows for the dynamic, consistent monitoring of ongoing damage and recovery (Wolf-Doty et al., 2021).

Identification of clinically relevant subtypes

Other publications were focused on building models to differentiate between AMBR and TCMR (Botella et al., 2024b) . The key finding was that dd-cfDNA was a strong potential biomarker for ABMR in patients who are clinically stable and have normal renal function (as measured by serum creatinine) (Botella et al., 2024a) . This is a huge advantage, since rejection at the subclinical level of ABMR can cause silent cumulative damage to the graft. Another published study described two patients who developed biopsy proven ABMR within 60 days of a significantly elevated dd-cfDNA, despite both patients demonstrating normal renal function (Rizvi et al., 2023b) .

Dd-cfDNA is particularly useful in certain clinical settings and high risk populations. The data suggests that dd-cfDNA is superior to creatinine for quantifying which patients with dnDSA (de novo donor-specific antibodies) which are harmful immune proteins developed by the recipient that specifically attack the new organ—were definitively rejecting (Botella et al., 2024b) , and needs to be considered for treatment decisions and guidance especially 15 days after transplantation (Botella et al., 2024a).

Performance in Multiple Ancestries

To ensure dd-cfDNA to be used as a valid biomarker, there needs to be evidence from studies with racially and ethnically diverse cohorts of patients (Stites et al., 2020) . This race and ethnicity diverse cohorts of patients (Stites et al., 2020). This racial and ethnic diversity is important in acquiring generalizable evidence for the

~~wider population of kidney transplant recipients. A number of studies reviewed did include or reported minority populations as part of their cohorts (Sawinski et al., 2021c).~~

Results

In order to fulfill the objectives of this systematic review, the results were synthesized by firstly, describing the studies included, and then quantifying and gaining a general idea about the diagnostic accuracy of dd-cfDNA, and lastly, identifying some important performance trends and knowledge gaps.

Comparison to Serum Creatinine

A key objective of this review was to compare the diagnostic accuracy of dd-cfDNA to the current standard of care, serum creatinine. Across the studies that conducted a direct comparison, dd-cfDNA consistently demonstrated superior performance in detecting allograft rejection. The reported Area Under the Curve (AUC) for dd-cfDNA was consistently high, generally above 0.80, whereas the AUC for serum creatinine in the same cohorts was significantly lower, often ranging from 0.55 to 0.65. This indicates that serum creatinine has poor to fair discriminatory ability for rejection, while dd-cfDNA has good to excellent discriminatory ability, supporting the hypothesis that it is a more accurate biomarker.

Identification of Clinically Relevant Subtypes

Other publications were focused on building models to differentiate between AMBR and TCMR (Botella et al., 2024b). The key finding was that dd-cfDNA was a strong potential biomarker for ABMR in patients who are clinically stable and have normal renal function (as measured by serum creatinine) (Botella et al., 2024a). This is a huge advantage, since rejection at the subclinical level of ABMR can cause silent cumulative damage to the graft. Another published study described two patients who developed biopsy-proven ABMR within 60 days of a significantly elevated dd-cfDNA, despite both patients demonstrating normal renal function (Rizvi et al., 2023b).

Dd-cfDNA is particularly useful in certain clinical settings and high-risk populations. The data suggests that dd-cfDNA is superior to creatinine for

quantifying which patients with dnDSA (de novo donor-specific antibodies)—which are harmful immune proteins developed by the recipient that specifically attack the new organ—were definitively rejecting (Botella et al., 2024b), and needs to be considered for treatment decisions and guidance especially 15 days after transplantation (Botella et al., 2024a).

Early Detection with dd-cfDNA

Additional research focused on the ability of dd-cfDNA to detect graft injury significantly earlier than clinical presentation (Bromberg et al., 2024) or elevations in standard biochemical markers like serum creatinine and proteinuria, which are known to be lagging (Kim et al., 2024b). Dd-cfDNA's 30-minute half-life allows for the dynamic, consistent monitoring of ongoing damage and recovery (Wolf-Doty et al., 2021).

Monitoring Frequency in Included Studies

A review of the included studies revealed significant heterogeneity in the dd-cfDNA monitoring protocols used. The frequency of testing varied widely, from daily monitoring in the immediate post-transplant period in some studies, to weekly, monthly, and quarterly testing in others focused on long-term surveillance. This lack of a standardized approach makes it difficult to draw firm conclusions about the optimal monitoring schedule and highlights a key area for future research.

~~*Performance in Multiple Ancestries*~~

~~To ensure dd-cfDNA to be used as a valid biomarker, there needs to be evidence from studies with racially and ethnically diverse cohorts of patients (Stites et al., 2020) . This race and ethnicity diverse cohorts of patients (Stites et al., 2020). This racial and ethnic diversity is important in acquiring generalizable evidence for the wider population of kidney transplant recipients. A number of studies reviewed did include or reported minority populations as part of their cohorts (Sawinski et al., 2021c).~~

Characteristics of Included Studies and Publication Trends

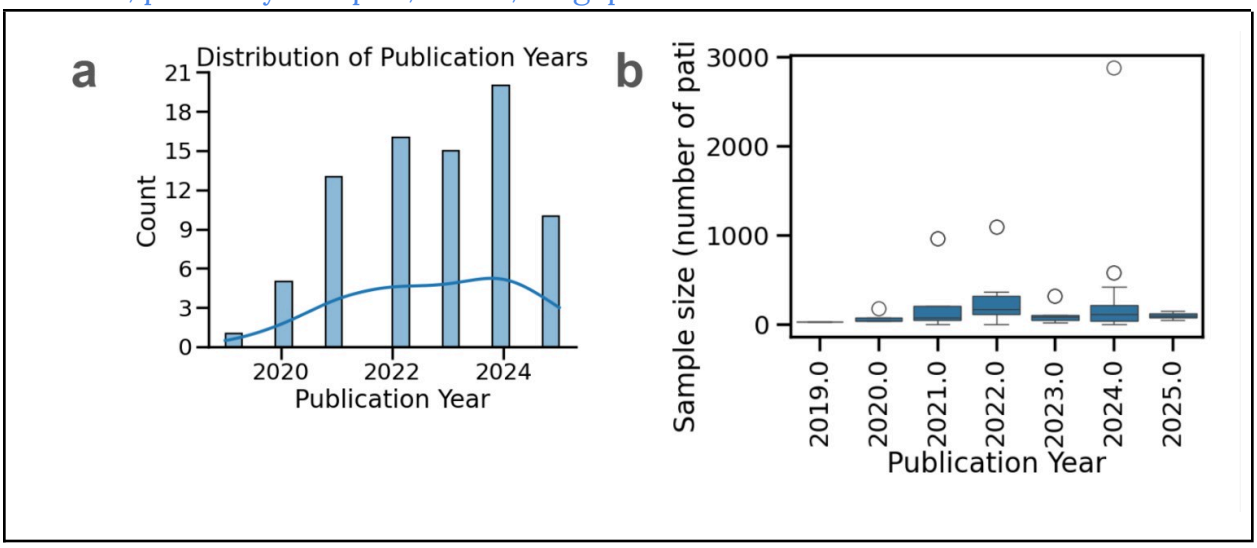
First, to synthesize the characteristics and findings of the included studies (Goal 2), the compiled list of 83 studies that met the inclusion criteria was examined. I then examined the distribution of studies per publication year (Fig. ~~1~~3a). I found that

publications have been steadily increasing over time, peaking in 2024 with approximately 22 studies. ~~which suggests a rapid maturation of the evidence base.~~ A linear regression analysis confirmed a significant positive trend between publication year and the number of studies (linear regression, slope = 3.5 studies/year, $p<0.05$), which suggests a rapid maturation of the evidence base. Furthermore, I wanted to understand how sample size changed with time in the field of kidney transplant dd-cfDNA research (Fig. ~~1~~3b). ~~Similar to the number of publications, we found that the median sample size has been increasing, especially in 2024, when a large study of nearly 3,000 patients (Aubert et al., 2024) was published, representing the largest dd-cfDNA study in kidney transplantation to date.~~ In contrast to the number of publications, an analysis of variance (ANOVA) showed that while the median sample size appears to trend upwards visually, this increase was not statistically significant ($p > 0.05$). This is likely due to high variability and the presence of several large-scale outliers, potentially impacted by the interruption of research during the Covid-19 pandemic. This indicates that while landmark studies are getting larger, the sample size for a typical study in the field has not significantly changed over this period. The evidence is predominantly composed of ~~high-quality study designs, those following participants forward in time and observing their outcomes as they happen, thereby minimizing bias, with prospective studies forming the largest single category (Fig. 1e).~~prospective cohort designs, with these studies forming the largest single category (Fig. ~~1~~3e). These studies, which follow participants forward in time, minimize bias and provide a strong evidence base.

To address the third goal of identifying trends and knowledge gaps, particularly regarding the biomarker's utility in diverse, underrepresented patient populations, the geographic and demographic distribution of the studies was analyzed. An examination of participant characteristics confirmed that the study focuses on a middle-aged cohort, with an average age of 52 years (Fig. 13c). This is both a useful finding and a commonly observed finding in clinical practice since the average age of transplantation is about 50 years. The analysis also highlighted that study cohorts tended to be predominantly male with a prominent peak, showing a proportion of male participants up to a significant 0.63 (Fig. ~~1~~3d). To gain a better understanding of the generalizability and nature of the evidence in a global context, a breakdown of the study types by geographic region was conducted. ~~This allowed us to see how the research in North America varied in methods put into practice,~~

~~contributing the most retrospective and observational research, whereas research in Europe and Asia contributed higher levels of prospective designs (Fig. 1f). The choice to go through the analysis this way was significant because this page had a demonstrated blend of complementary types of research to provide a wider and more diverse evidence base – the prospective studies provide higher quality controlled data while the retrospective and observational studies filled in the gap of real-word evidence, which in sum produced better conclusions for this review.~~An analysis of the geographic distribution of study types revealed that North American research contributed the most retrospective and observational studies, whereas Europe and Asia contributed a higher proportion of prospective cohort designs (Fig. ~~1~~3f). A chi-square test of independence confirmed that this difference in the distribution of study types across continents is statistically significant ($\chi2(6) = 18.98$, $p = 0.004$).

~~Finally, to specify and quantify the overall diagnostic accuracy of dd-cfDNA for various types of rejection (Goal 1), the performance metrics from the included studies were synthesized. An important goal of this study was to understand the efficacy of cfDNA biomarkers in multiple ancestries, because it is important to have an equitable biomarker. To this end, I first examined the number of studies covering each self-reported race/ancestry group, which I defined as "Asian", "Black", "White", and "Hispanic/Latino." We found that the studies analyzing Black and Hispanic/Latino populations come almost exclusively from North American studies, while data on Asian populations is contributed by studies from both North America and Asia, primarily in Japan, China, Singapore and Thailand.~~



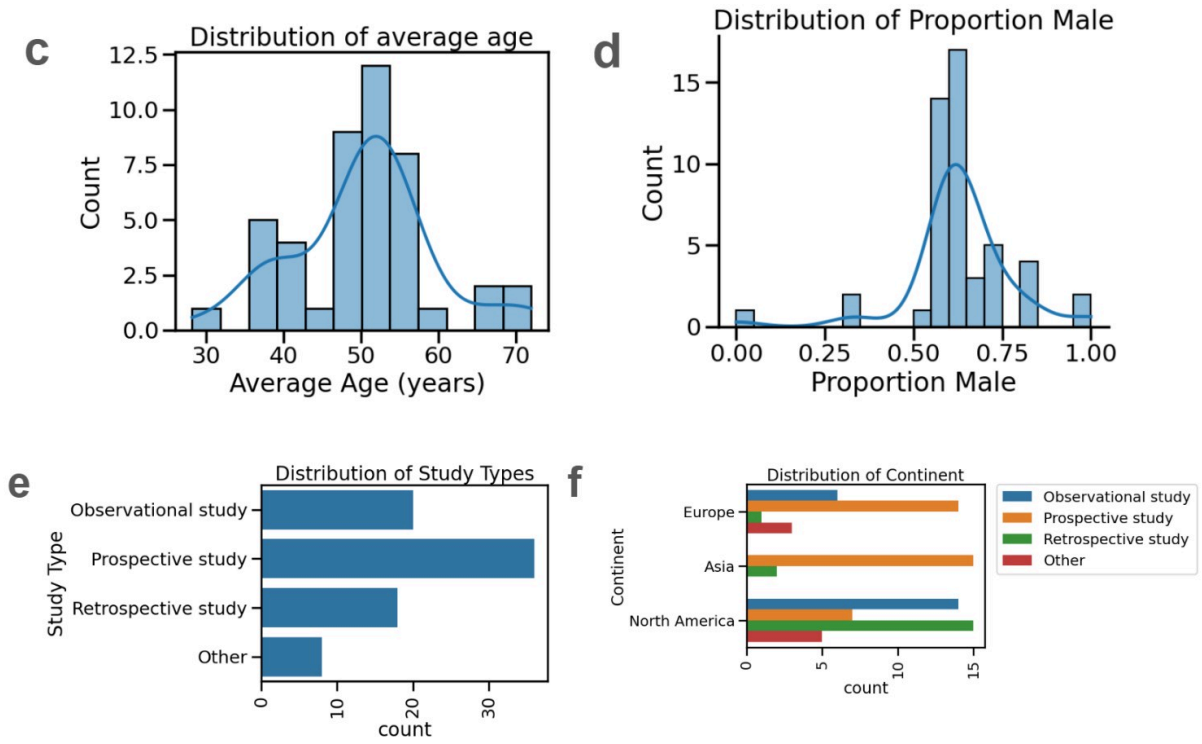


Figure ~~23~~: Publication Trends, Demographics, and Methodological Characteristics of Included Studies. (a) A histogram showing the number of included studies published per year, illustrating a rising trend in research volume peaking in 2024. (b) A box plot visualizing the distribution of study sample sizes by publication year, highlighting an increase in the scale of research and several large-scale outliers. (c) A histogram displaying the distribution of the average age of participants, showing a concentration in the middle-aged demographic. (d) A histogram of the proportion of male participants in study cohorts, indicating a predominance of male-majority studies. (e) A horizontal bar chart quantifying the included studies by their primary methodological design, with prospective studies being the most common. (f) A grouped horizontal bar chart illustrating the distribution of study types across the primary geographic regions of North America, Asia, and Europe.

Methodological Quality of the Evidence Base

To assess the methodological soundness of the literature used in the study, I designed a 5-point quality scoring system. The 5-point quality scoring system was created to determine which sources most fit the hypothesis. I awarded each study one point for fulfilling each of the criteria: (1) Measured Blood?, (2) Analyzes probability of rejection?, (3) Temporal Precedence Explicitly Stated?, (4) longitudinal

monitoring for time?, (5) compares to traditional biomarkers?. These scoring criteria were chosen to evaluate the relevance of each study to the specific research hypothesis. The criterion "Measured Blood?" evaluated whether the study utilized the appropriate sample type to analyze dd-cfDNA (Sawinski et al., 2021a). The criterion "Analyzes rejection probability?" ensured that the study focused on the primary clinical outcome, i.e. rejection (Bu et al., 2022b). "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for Timing?" allowed for the evaluation of the hypothesis, specifically the earlier detection factor, which required identifying studies that monitored the biomarker over time (Parajuli et al., 2024b). Finally, "Compares to Traditional Biomarkers?" validated that the study conducted the comparative clinical analysis of the complete framework, thereby determining the "superior performance" of dd-cfDNA analysis compared to traditional clinical surveillance methods such as serum creatinine (Mantios et al., 2023). ~~The overall frequency distribution of quality scores (Figure 2a) indicates that the overall evidence base is generally high quality.~~ The distribution of studies per publication year (Fig. 4a) indicates a rapid maturation of the evidence base, with publications steadily increasing over time and peaking in 2024. Most studies reporting their scores are clustered in the upper range of the scale (e.g. 35 studies scored a 3; 22 studies scored a 4; and 21 studies scored a 5).

To add more specificity to the study characteristics that were contributing to high quality evidence, which criteria were most frequently met across the studies were assessed (Fig. ~~2~~4b). The breakdown shows that foundational criteria were met across studies, particularly utilizing blood as a sample source (n=79) and comparing dd-cfDNA to traditional biomarkers (n≈78). ~~Despite this, there appears to be a large methodological gap in the literature regarding timing of rejection. While a reasonably sized number of studies (n=45) utilized some form of longitudinal monitoring for timing, a much smaller, and specific subset (n≈29) stated the temporal precedence of dd-cfDNA elevation before a rejection diagnosis was confirmed. This indicates a clear opportunity for future research to focus more rigorously on study designs that can quantify this early detection window.~~

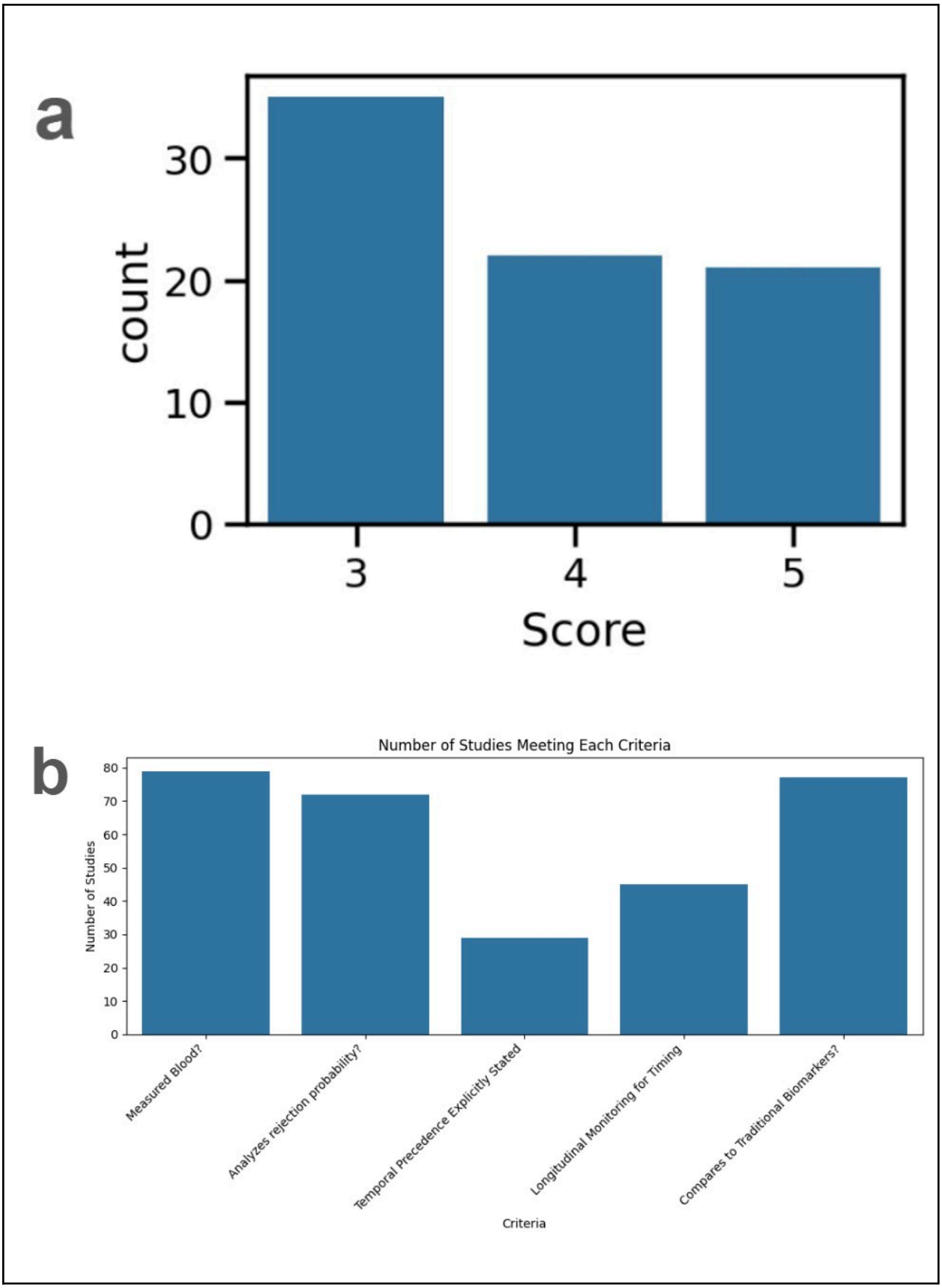
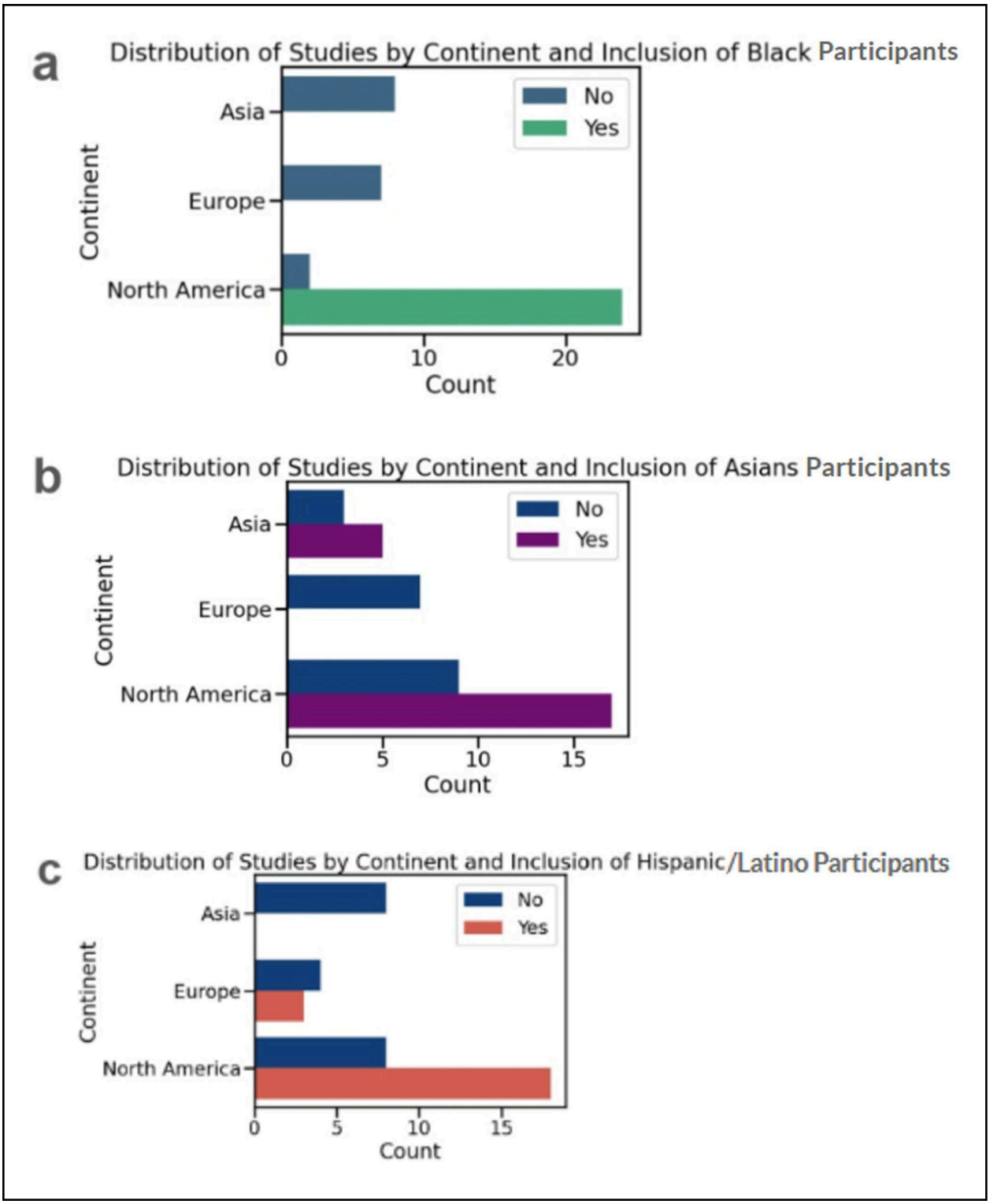

a
30
20
10
0
count
3
4
5
Score
b
Number of Studies Meeting Each Criteria
Number of Studies
80
70
60
50
40
30
20
10
0
Measured Blood?
Analyzes rejection probability?
Temporal Precedence Explicitly Stated
Longitudinal Monitoring for Timing
Compares to Traditional Biomarkers?
Criteria

Figure ~~34~~: Distribution of Study Quality Scores and Fulfillment of Methodological Criteria. (a) A histogram showing the distribution of the overall 5-point quality scores across all included studies. (b) A bar chart quantifying the number of studies that met each of the five specific methodological quality criteria.

Geographic and Ancestral Distribution of the Evidence Base

In order to understand the efficacy of dd-cfDNA in multiple ancestries, which is an important goal for comprehensive utility of dd-cfDNA in kidney transplant settings, the inclusion of diverse racial and ethnic groups across the primary research continents was examined. To this end, I examined the number of studies covering each self-reported race/ancestry group, which I defined as "Asian," "Black," "White," and "Hispanic/Latino." The analysis reveals that North American research provides the most diverse evidence base. North American studies were the exclusive source of data for the inclusion of Black participants (n≈24) (Fig. ~~3~~5a) and Hispanic/Latino participants (n≈18) (Fig. ~~3~~5c). Furthermore, North American research also contributed the largest number of studies that include Asian participants (n≈17) (Fig. ~~3~~5b) and White participants (n≈22) (Fig. ~~3~~5d).

In contrast, the European studies included in this review, while numerous, did not explicitly report the inclusion of Black, Asian, or Hispanic participants, with many not reporting race altogether, though a small number did report including White patients (n≈3). The Asian studies, as expected, contributed evidence on Asian populations (n≈4) but did not report inclusion of other groups. This distribution highlights a significant gap in the global evidence, suggesting that the generalizability of findings from Europe and Asia to diverse populations is limited.


a
Distribution of Studies by Continent and Inclusion of Black Participants
Asia
Europe
North America
Continent
No
Yes
0
10
20
Count
b
Distribution of Studies by Continent and Inclusion of Asians Participants
Asia
Europe
North America
Continent
No
Yes
0
5
10
15
Count
c
Distribution of Studies by Continent and Inclusion of Hispanic/Latino Participants
Asia
Europe
North America
Continent
No
Yes
0
5
10
15
Count

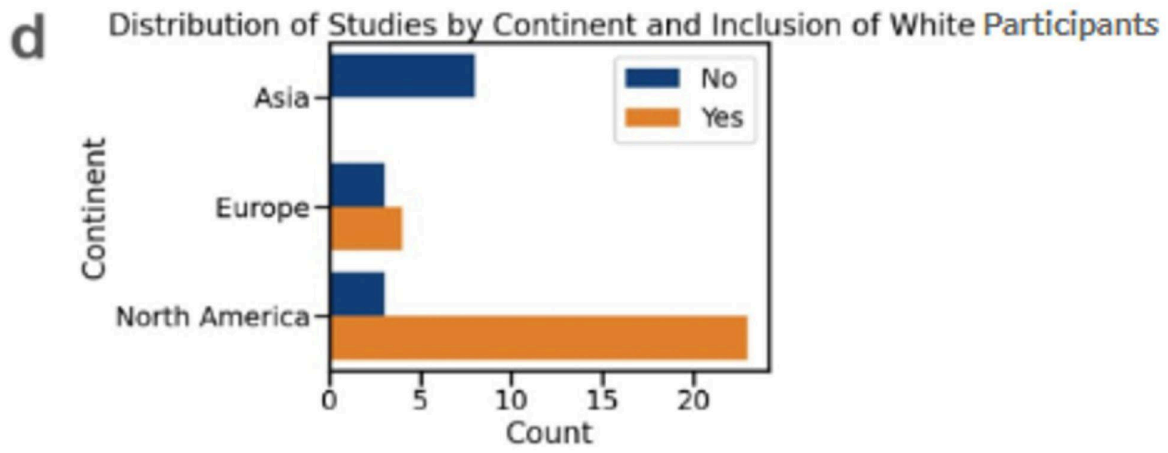


Figure 45: Geographic Distribution of Studies by Patient Ancestry. A series of horizontal bar charts showing the number of studies from North America, Europe, and Asia that did ("Yes") or did not ("No") report the inclusion of specific patient ancestries. Studies that did not report the race/ethnicity of their participants are not included. (a) Inclusion of Black participants. (b) Inclusion of Asian participants. (c) Inclusion of Hispanic/Latino participants. (d) Inclusion of White participants.

Diagnostic Accuracy of dd-cfDNA

A major study goal was to determine how effective dd-cfDNA was at identifying cases of kidney transplant rejection; therefore, AUC values were extracted from studies and analyzed. The diagnostic performance of dd-cfDNA, as measured by AUC, was found to be high across the included studies, as summarized in the histogram of Area Under the Curve (AUC) values (Fig. 46a). ~~The Receiver Operating Characteristic (ROC) curve is a representation of a diagnostic test's overall performance, as described by one number, the Area Under the Curve (AUC). The AUC takes a value between 0.5 (no better than chance) and 1.0 (perfect test). Values between 0.80 and 0.90 may be considered "good". For the purposes of this systematic review, the AUC indicates the extent to which dd-cfDNA separates patients with biopsy proven rejection (the cases) from patients without biopsy proven rejection (the controls); therefore, a high AUC indicates that the biomarker is quite accurate in identifying patients that are actively rejecting their organ.~~ The data reveals a consensus towards models using dd-cfDNA having strong performance in the prediction of rejection, with a bimodal distribution showing two primary peaks where the evidence is concentrated. The largest cluster of studies

reported an Area Under the Curve (AUC) with a peak centered approximately around AUC=0.80, while a second prominent peak centered even higher around 0.90. It is important to put this in perspective; for a standalone clinical diagnostic test, an AUC of 0.80 is generally considered poor as this indicates significant false positive and false negative rates that may cause indeterminate clinical decisions. An interesting finding is that studies ~~do not report mediocre performance and if this is potentially influenced by publication bias of studies with negative findings, it may at least suggest that when dd-cfDNA is investigated, it is consistently a good and effective biomarker for detection of allograft rejection.~~rarely report mediocre performance; if this is influenced by publication bias where negative findings are underreported, it would suggest that the true diagnostic accuracy of dd-cfDNA may be overestimated in the current literature.

~~Next, we examined whether the AUC, as a proxy for performance, has been changing as the literature matures. The diagnostic performance of dd-cfDNA was found to be reliably high in the studies that were included, as shown in the histogram of Area Under the Curve (AUC) values (Fig. 46a). There is a clear shared opinion towards high identifiability with a bimodal distribution presenting two peaks where the evidentiary support lies. The largest mass of studies report an AUC that is concentrated around 0.80 indicating a very good diagnostic test, whereas the second primary peak is concentrated even higher, around AUC 0.90, which demonstrates excellent diagnostic performance. The clear failure of studies to report poor or mediocre performance clearly indicates, in a particularly robust quantitative sense, that dd-cfDNA is a reliable and effective biomarker in allograft rejection detection.~~

I then performed a linear regression analysis to model the relationship of year of publication and AUC analyses in order to assess whether reported performance of the biomarker has improved as studies have accrued (a primary goal of this review). This does demonstrate an increasing association over time (slope ≈ +0.01 per year of AUC) (Fig. 46b). This positive slope suggests that more 'recent' studies, on average, have reported greater AUC than 'earlier' studies. This potential trend may in part reflect field maturation, perhaps including improved laboratory assays (i.e., could be absolute quantification vs. fractional abundance) and more sophisticated study design (years and capabilities are reflected in the increasing median sample sizes). The evidence for dd-cfDNA has not weakened over time; instead, it seems to have

grown and solidified. An alternative explanation might include a growing publication pressure to report findings that best represent previous literature. This study's upward trend in AUC is accompanied by an upward trend in the sample sizes of studies over time.

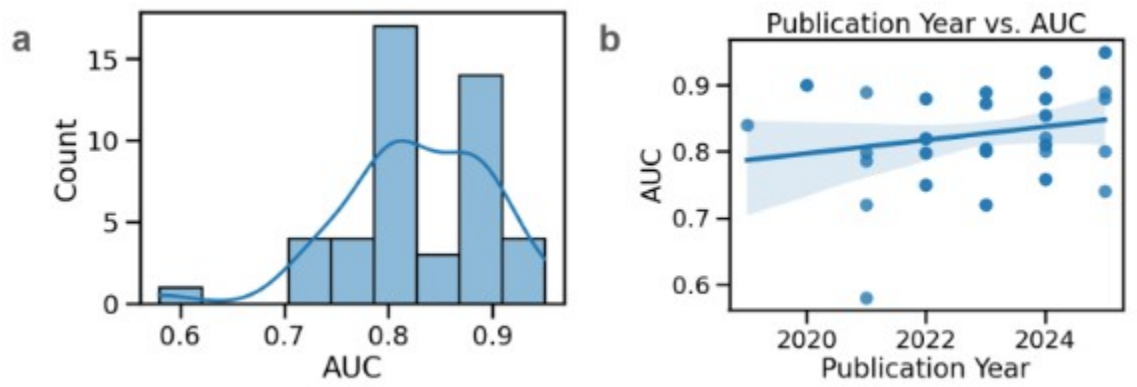


Figure ~~5~~6: Summary of Diagnostic Accuracy and Performance Trends. (a) A histogram showing the bimodal distribution of reported Area Under the Curve (AUC) values from all included studies. (b) A regression plot illustrating the positive correlation between the publication year of the studies and their reported AUC values.

Performance Across Patient Ancestries

To further assess the generalizability of dd-cfDNA's performance, the AUC values were stratified by studies that reported the inclusion of different racial and ethnic groups. For this analysis, studies that did not report AUC or did not specify the racial demographics of their cohort were excluded, resulting in a subgroup of 45 studies. Overall, the data suggests that dd-cfDNA consistently has accuracy in identifying potential rejection across studies with different cohort demographics. The median AUC of studies that specifically included Black participants (N=24) was 0.81 (Fig. 7a). This is an important observation ~~because it supports the accuracy of this research for this patient population, which is specifically at higher risk for adverse transplant outcomes.~~supporting the biomarker's utility in this patient population, which has historically experienced higher rates of adverse transplant outcomes due to a combination of social and clinical factors.

It is critical to clarify that the available literature does not permit an analysis of dd-cfDNA performance exclusively within the Hispanic/Latino population. Instead, this review analyzed the group of studies that reported the inclusion of

Hispanic/Latino participants within their larger, mixed-ancestry cohorts. For this collection of studies, the median AUC was 0.81, showing a narrow and consistent interquartile range (Fig. 7b). Therefore, while this result indicates that the biomarker performs well in diverse studies that include Hispanic patients, it is impossible to draw conclusions from this data about its specific performance in Hispanic populations alone.

The studies measuring the performance of dd-cfDNA that reported Asian participation strengthen the evidence for applicability of dd-cfDNA across racial and ethnic groups. Interestingly, the analysis suggests that studies including Asian participants demonstrate a trend towards higher and more consistent accuracy, with a median Area Under the Curve (AUC) of approximately 0.82 and a notably narrow interquartile range (Fig. 7c). This enhanced consistency may be attributed to factors like more homogeneous patient cohorts or the use of optimized, population-specific assays, highlighting the potential for dd-cfDNA's performance to be further refined in specific clinical and demographic contexts.

Finally, the analysis of studies including White participants provides further evidence of the biomarker's performance (Fig. 7d). The apparent better performance of studies that did not explicitly report including White participants is likely due to those studies being a heterogeneous mix of smaller or more homogeneous international cohorts where ethnicity is often not specified, whereas the "Yes" group consists primarily of large, diverse North American trials whose real-world complexity results in a more consistent, albeit slightly lower, median AUC.

~~In conclusion, the consistent high AUC values across all four of these distinct analyses strengthens the overall hypothesis that dd-cfDNA is a superior and broadly applicable biomarker for detecting allograft rejection in a wide spectrum of patient populations.~~ While there are slight variations in the median AUC across studies that included participants of different self-reported ancestries, an analysis of variance (ANOVA) confirmed that these differences are not statistically significant ($p > 0.05$). This is consistent with the visual interpretation of the overlapping confidence intervals in the box plots. Therefore, while the available data suggests that dd-cfDNA performs with comparable diagnostic accuracy across cohorts with different ancestral compositions, this conclusion is preliminary. More

robust data, such as dedicated studies on Hispanic populations and a broader geographic representation for Asian populations, is needed to definitively confirm these findings and account for potential regional biases.

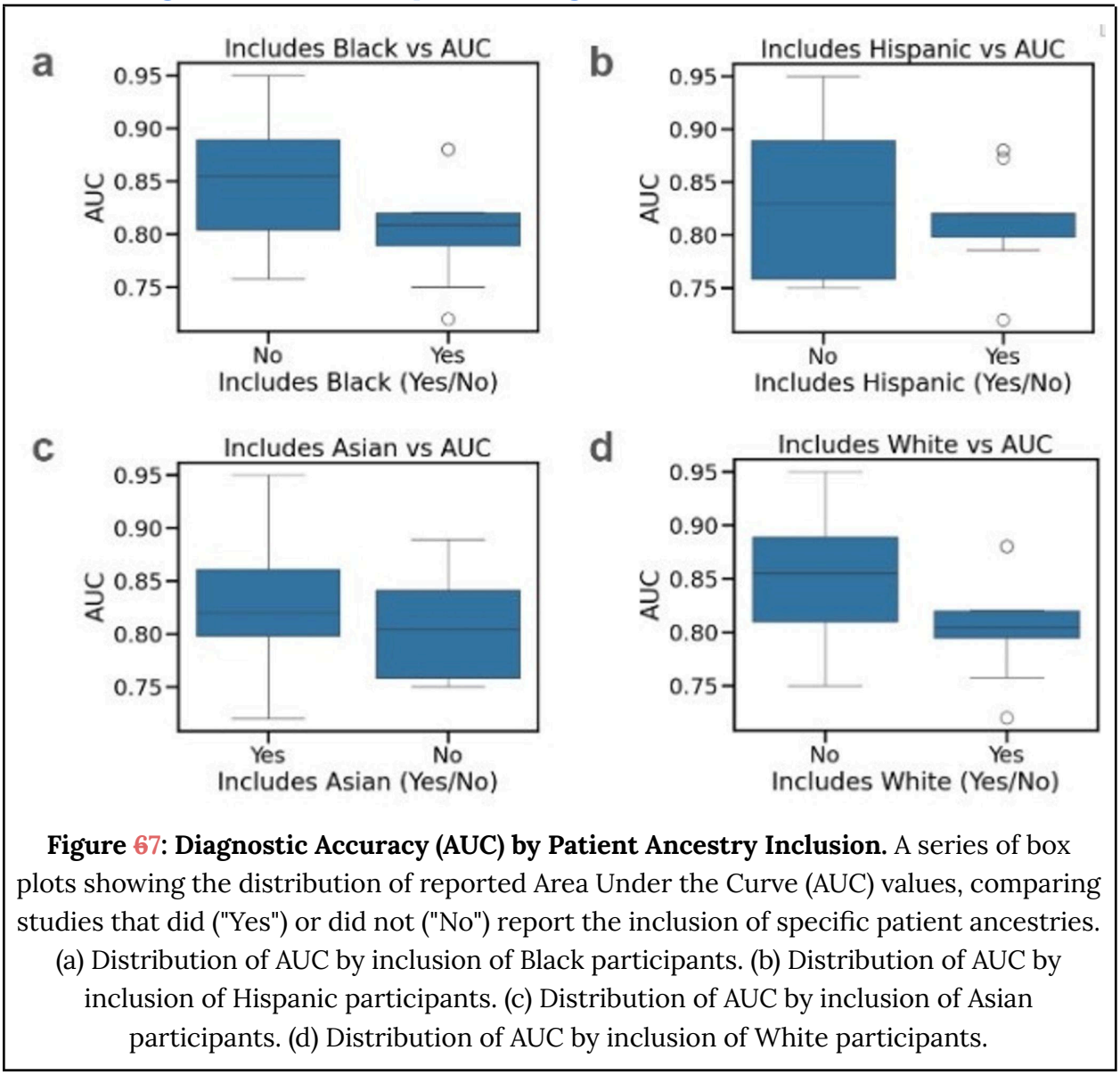


Figure ~~6~~7: Diagnostic Accuracy (AUC) by Patient Ancestry Inclusion. A series of box plots showing the distribution of reported Area Under the Curve (AUC) values, comparing studies that did ("Yes") or did not ("No") report the inclusion of specific patient ancestries. (a) Distribution of AUC by inclusion of Black participants. (b) Distribution of AUC by inclusion of Hispanic participants. (c) Distribution of AUC by inclusion of Asian participants. (d) Distribution of AUC by inclusion of White participants.

Performance Across Different Study Contexts

Beyond the overall quality of the studies, it is also important to understand how the performance of dd-cfDNA varies across different study contexts, as detailed in Fig. ~~6~~8. An analysis of performance by geographic region reveals a consistently high level of diagnostic accuracy globally, with the median AUC for studies from Europe, Asia, and North America all above 0.80 (Fig. ~~6~~8a). Furthermore, an analysis of study

scale shows no significant correlation between a study's sample size and its reported AUC (Fig. ~~6~~8b); the nearly flat regression line indicates that high performance is not simply an artifact of small studies, a conclusion reinforced by the finding that the largest study with nearly 3,000 patients still reported a strong AUC.

dd-cfDNA's performance is further evaluated by an analysis of demographic and methodological factors. The regression plot of average participant age versus AUC shows a slight downwards trend (Fig. ~~6~~8c). This indicates that in the age group between ages 30 to 60, the AUC decreased as age increased, thereby signifying the need for more studies to further validate this relationship. Finally, while both single-center and multicenter studies reported similarly high median AUCs, the violin plot reveals that the results from multicenter studies were more consistent, with a tighter distribution and less variability (Fig. ~~6~~8d). One aspect to note about Fig. ~~6~~8d is that a violin plot was utilized over a standard box plot because it visualizes the full probability density of the data.

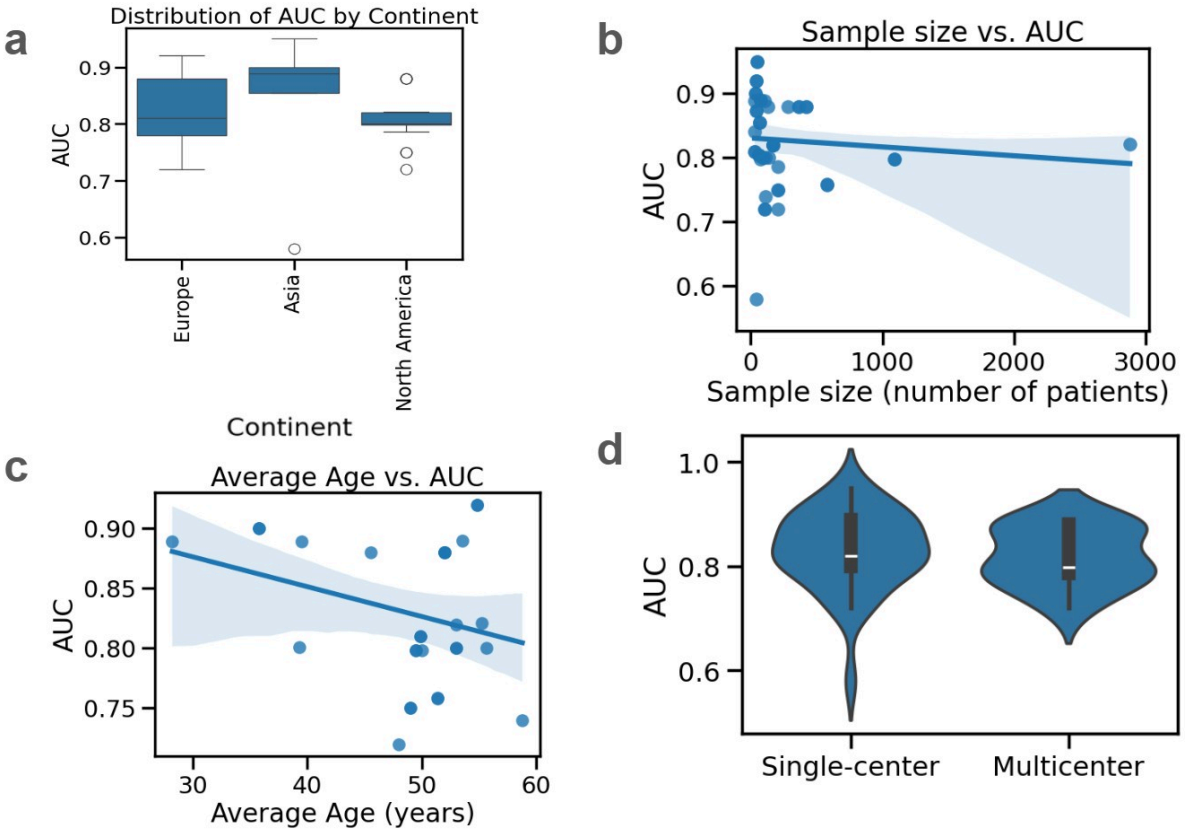


Figure ~~7~~8: Analysis of dd-cfDNA Diagnostic Accuracy Across Different Study Contexts.

(a) A box plot showing the distribution of reported AUC values, stratified by continent. (b) A regression plot illustrating the relationship between study sample size and reported AUC. (c) A regression plot showing the relationship between the average age of study participants and reported AUC. (d) A violin plot comparing the distribution of reported AUC values between single-center and multicenter studies.

Relationship Between Study Quality and Diagnostic Accuracy

I also analyzed the relationship between the study scores, as determined by my 5-point relevance criteria, and their reported diagnostic accuracy (AUC). The box plot of overall quality scores reveals that studies with higher scores (3, 4, and 5) consistently reported high median AUCs above 0.80 (Fig. ~~79~~a). This indicates that the score of a study is not directly related to the reported AUC, as high diagnostic performance was observed across all quality tiers. A more granular analysis shows that studies incorporating more rigorous methodologies, such as longitudinal monitoring, not only reported a high median AUC but also showed more consistent results, evidenced by a tighter distribution of values compared to studies that did not meet this criterion (Fig. ~~79~~b).

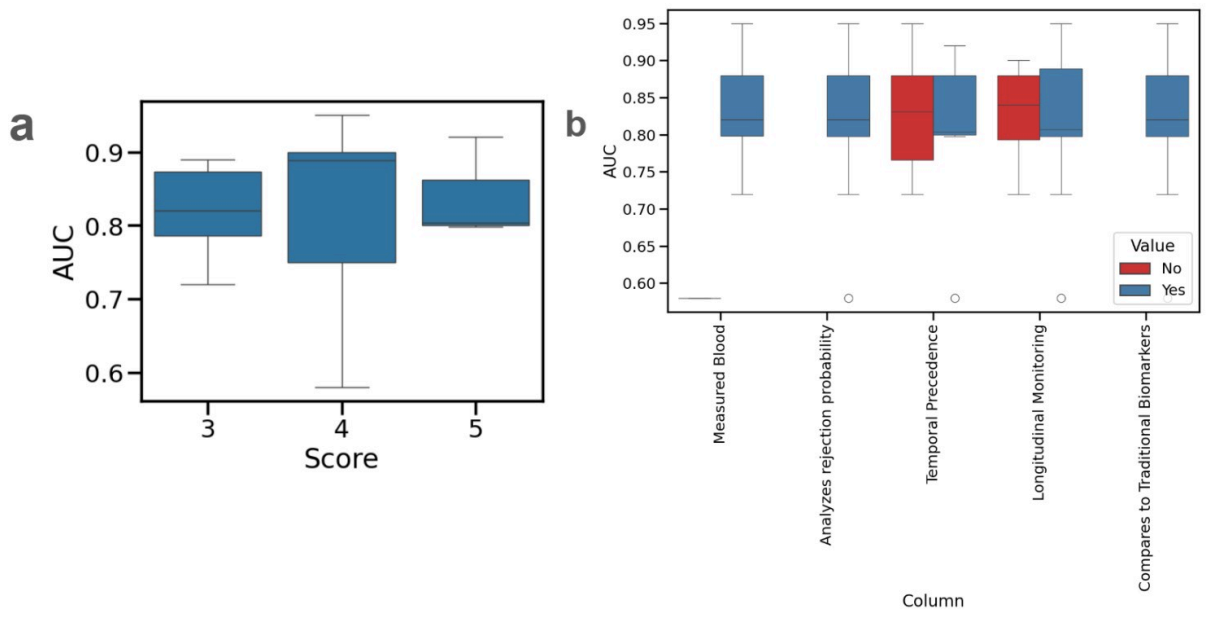


Figure ~~89~~: Relationship Between Study Quality and Diagnostic Accuracy (AUC). (a) A box plot showing the distribution of reported AUC values, stratified by the overall 5-point quality score of the studies. (b) A box plot comparing the distribution of reported AUC values for studies that did ("Yes") or did not ("No") meet each of the five specific

methodological quality criteria.

Performance in Specific Rejection Subtypes

Lastly, to understand the evidence base for different rejection pathways, the included studies were analyzed based on their inclusion of either antibody-mediated rejection (ABMR) or T-cell mediated rejection (TCMR). The analysis shows that the literature provides robust coverage of both primary rejection subtypes. A substantial majority of the included studies, approximately 48, investigated patient cohorts that included cases of ABMR (Fig. 810a), while a similarly large number of studies, approximately 45, included cases of TCMR (Fig. 810b), though many included both.

Furthermore, an analysis of the reported diagnostic accuracy (AUC) shows that the high performance of dd-cfDNA is consistent regardless of a study's specific focus on rejection subtypes. The median AUC for studies that included ABMR was approximately 0.82, a value nearly identical to the median AUC for studies that did not specifically focus on ABMR (Fig. 810c). A similar pattern was observed for TCMR, where the median AUC was also consistent at approximately 0.82 in studies that investigated this rejection type (Fig. 810d).

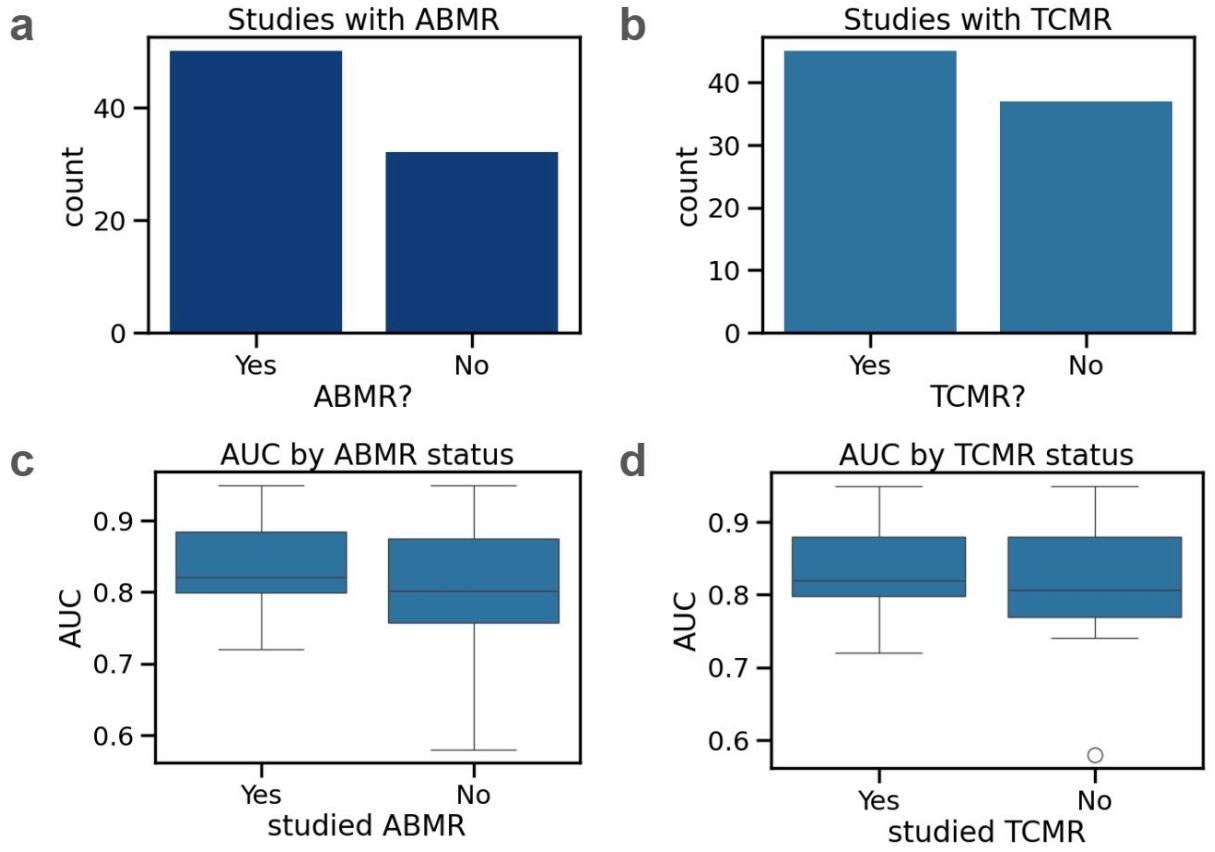


Figure ~~9~~10: Research Focus on Rejection Subtypes and Corresponding Diagnostic Accuracy. (a) A bar chart showing the number of studies that did ("Yes") or did not ("No") investigate antibody-mediated rejection (ABMR). (b) A bar chart showing the number of studies that did ("Yes") or did not ("No") investigate T-cell mediated rejection (TCMR). (c) A box plot comparing the distribution of reported AUC values between studies that did or did not investigate ABMR. (d) A box plot comparing the distribution of reported AUC values between studies that did or did not investigate TCMR.

~~In summary, this review demonstrated that dd-cfDNA is a dynamic marker that must be interpreted cautiously. The finding that the dynamics of dd-cfDNA may become "increasingly erratic" following kidney injuries further accentuates our argument that a single static measurement could be misleading (Nguyen et al., 2025). The assessment of the methodological quality of the studies included in this review brought attention to this issue, and when looking at the literature, we were able to identify a gap in the literature to explicitly assess the temporal precedence (Fig. 2b) of dd-cfDNA, and this further supports the conclusion that dd-cfDNA should be interpreted in a sequential manner over time as observations of trends in~~

~~dd-cfDNA levels provides the best differentiation between true signals of injury versus transient fluctuations of observed dd-cfDNA levels (Stites et al., 2020).~~

Methodological Gaps in Evaluating Rejection Timing

Despite this, there appears to be a large methodological gap in the literature regarding timing of rejection. While a reasonably sized number of studies (n=45) utilized some form of longitudinal monitoring for timing, a much smaller, and specific subset (n≈29) stated the temporal precedence of dd-cfDNA elevation before a rejection diagnosis was confirmed. This indicates a clear opportunity for future research to focus more rigorously on study designs that can quantify this early detection window.

~~*Limitations of dd-cfDNA in transplantation research*~~

~~While the synthesized evidence provides strong support for the clinical utility of dd-cfDNA, this systematic review also produced some important limitations. The greatest limitation revealed in the analysis is insufficient ethnic and geographic diversity in the evidence base. The results show (Fig. 3) that the sample for Black and Hispanic/Latino populations is almost exclusively North American data (Ralph et al., 2025), inferring that it needs ongoing validation in global patient populations to confirm applicability for those populations.~~

~~The literature suggests dd-cfDNA is not a standalone diagnostic test, but a strong complementary tool for diagnostics, which is reflected by this review's findings on its variable diagnostics performance. The results demonstrated that there was variable diagnostic accuracy for dd-cfDNA in some rejection subtypes like T-cell mediated rejection (TCMR) and there was variability dependent upon assays (Halloran et al., 2022). These results informed that dd-cfDNA is most useful if used within a multi-modal approach with histology where dd-cfDNA has a supportive role in increasing surveillance and avoiding unnecessary invasive biopsy (Wei et al., 2024).~~

~~In summary, this review demonstrated that dd-cfDNA is a dynamic marker, that must be interpreted cautiously. The finding that the dynamics of dd-cfDNA may become "increasingly erratic" following kidney injuries further accentuates our argument that a single static measurement could be misleading (Nguyen et al.,~~

2025). The assessment of the methodological quality of the studies included in this review brought attention to this issue, and when looking at the literature, we were able to identify a gap in the literature to explicitly assess the temporal precedence (Fig. 2b) of dd-cfDNA, and this further supports the conclusion that dd-cfDNA should be interpreted in a sequential manner over time as observations of trends in dd-cfDNA levels provides the best differentiation between true signals of injury versus transient fluctuations of observed dd-cfDNA levels (Stites et al., 2020).

Future Directions and Recommendations

To realize the clinical potential of donor-derived cell-free DNA (dd-cfDNA) fully, the next steps need to focus on precision, standardization, and appropriate incorporation into existing clinical guidelines. This systematic review confirmed that the biomarker has high diagnostic accuracy (Fig. 4a), but as shown in this review, different commercial assays can have different rejection detection rates for the same type of rejection (Fig. 8d). To establish standardized diagnostic thresholds, robust, prospective trials are still needed (Benning et al., 2023b). The next phase should include head-to-head comparisons of the existing dd-cfDNA platforms to create standardized reporting metrics, as well as validate their performance across real-world populations (Loupy et al., 2024b). Also, there is a need for evidence-based guidelines regarding frequency of monitoring, such as testing protocols of monthly monitoring for the first 6 months post-transplant, and quarterly thereafter, in order to determine the most cost-effective monitoring schedule for the early detection of subclinical rejection, which is still in an actionable state (Sharma et al., 2022).

The ultimate success of this biomarker will hinge upon its direct ability to translate into improvements in the longer-term outcomes of patients and allografts. As we established in this review of methodological quality, there were many studies (n≈45) that undertook longitudinal monitoring, and a much smaller number (n≈29) deliberately documented temporal precedence (Fig. 2b). This indicates an urgent need for intervention trials that utilize dd-cfDNA's early warning signal to direct pre-emptive treatment (i.e., if dd-cfDNA heightens, initiate increased immunosuppression) (Mirza et al., 2024). For example, a randomized controlled trial would directly compare a dd-cfDNA based strategy with the current standard of care (i.e., initiation of treatment when a clinical event has already occurred) (Tian et

~~al., 2025). This would provide direct evidence as to whether or not these intervention strategies can improve the trajectory of allograft health (as currently we have only observational evidence that an increased renewable dd-cfDNA level considerably raises the risk of decline in the subsequent future graft health) (Obrișcă et al., 2022).~~

~~Although dd-cfDNA is now being utilized in clinical practice, this systematic review established several important limitations within the existing literature that need to be overcome in order to properly refine the field of dd-cfDNA use. The guidelines on dd-cfDNA use are evolving quickly, but the best way to incorporate dd-cfDNA into those guidelines is yet to be resolved (Kim et al., 2024c). The fact that all of the evidence for use in Black and Hispanic populations came from North American based studies (Fig. 3), there is an immediate need for future research on these underrepresented populations in Europe and Asia to allow for equitable and effective dd-cfDNA use (González López et al., 2023). When we operationalize these knowledge gaps – focusing on performance in broader and diverse cohorts, as well as specific rejection sub-types like TCMR – the clinical community can improve the utility of dd-cfDNA and establish it as a new standard in personalized, non-invasive transplant care (Dauber et al., 2020).~~

Conclusions

The central finding of this systematic review—that donor-derived cell-free DNA (dd-cfDNA) serves as a robust biomarker for kidney allograft rejection with high diagnostic accuracy—is largely consistent with the conclusions of previous literature reviews and meta-analyses. Multiple reviews have affirmed the biomarker's high potential as a "valuable tool" for non-invasive monitoring (Gisch et al., 2025) and as a reliable liquid biopsy for detecting allograft injury (Oellerich et al., 2021). Furthermore, our findings on diagnostic accuracy align with several recent meta-analyses that have reported high pooled Area Under the Curve (AUC) values, such as 0.84 for acute rejection (Zhang et al., 2021), 0.86 for overall rejection (Zhang et al., 2025), and 0.88 for antibody-mediated rejection (Yang et al., 2024). However, the present review builds upon this established consensus by providing a more granular, quantitative synthesis of the evolving evidence base. While prior work has established the biomarker's general utility, our analysis is the first to systematically document the rapid maturation of the field, including the significant increase in study sample sizes and the shift towards more rigorous prospective cohort designs.

Most importantly, our review addresses a critical knowledge gap by quantifying the significant geographic and ancestral disparities in the literature. Previous reviews have qualitatively mentioned the need for more diverse cohorts, but our analysis provides concrete evidence that data for Black and Hispanic populations are almost exclusively derived from North American studies, underscoring a limitation in the global generalizability of current findings that was not the primary focus of prior meta-analyses.

This systematic review serves as a thorough synthesis of recent and robust research on donor-derived cell-free DNA that helps solidify its importance in the post-transplant period (Huang et al., 2023). The overall body of evidence supports that dd-cfDNA is consistently more accurate diagnostically than traditional biomarkers, such as serum creatinine, because it can detect rejection in those with stable creatinine levels (Gupta et al., 2022). The overall implication is that the use of dd-cfDNA represents a shift from reactive to proactive, evidence-based monitoring and screening that enables clinicians to detect allograft injury weeks or months before it presents clinically (Benning et al., 2023a). Although this review did systematically confirm that a number of studies reported a consistently high measure of diagnostic accuracy (e.g., > 0.80 as the AUC; Mantios et al., 2023), it also highlighted an important knowledge gap around the biomarker as most of the data on Black and Hispanic populations came exclusively from North American cohorts. One concrete recommendation based on this analysis is that future meta-analyses should synthesize other core metrics, such as Negative Predictive Value (NPV), in order to clearly define how dd-cfDNA should be used across different clinical scenarios (Aubert et al., n.d.). In summary, this systematic review concludes that dd-cfDNA is a milestone in transplant monitoring that provides a foundation for more personalized and non-invasive patient care (Loupy et al., 2024a).

It is also important to consider the limitations inherent in the designs of the included studies. While this review found a predominance of prospective cohort studies, which is a methodological strength, a notable number of studies were retrospective or single-center designs. Single-center studies may have limited external validity, while retrospective designs can be subject to selection bias and incomplete data. These factors represent a limitation of the overall evidence base and highlight the need for more large, prospective, multicenter trials to confirm the findings synthesized in this review. Furthermore, there was significant variability in

diagnostic accuracy reported between different commercial assays used across the studies, representing another key limitation of the current evidence base.

This systematic review confirms that dd-cfDNA is a useful non-invasive biomarker for early allograft injury with long-term prognostic value (Kim et al., 2024a; Tian et al., n.d.; Nie et al., 2025; Loupy et al., 2024). While included studies demonstrated its ability to identify rejection sooner than traditional markers (Bromberg et al, 2024; Parajuli et al., 2024b), this review identified significant limitations, including a lack of ethnic diversity and variable accuracy across different assays (Nguyen et al., 2025). Furthermore, because dd-cfDNA is a dynamic marker whose levels can become erratic, a single static measurement can be misleading (Nguyen et al., 2025). Therefore, future research must focus on validating performance in diverse populations and establishing standardized protocols for sequential monitoring to best differentiate true injury from transient fluctuations (Stites et al., 2020).

To fully realize the clinical potential of donor-derived cell-free DNA (dd-cfDNA), future research must focus on precision, standardization, and equitable application. This systematic review confirmed the biomarker's high diagnostic accuracy but also highlighted that different commercial assays can have different rejection detection rates for the same rejection subtype. Therefore, robust prospective trials are needed to conduct head-to-head comparisons of existing platforms and establish standardized reporting metrics. Furthermore, to ensure the clinical translatability and applicability of this biomarker, future studies must also assess its cost-effectiveness and the logistical challenges of integrating dd-cfDNA testing into routine clinical workflows. To ensure equitable and effective use, there is an immediate need for future research on underrepresented populations in Europe and Asia, as current evidence for Black and Hispanic populations is almost exclusively from North American studies. The ultimate success of this biomarker will also hinge on its ability to improve long-term patient outcomes, necessitating intervention trials that use dd-cfDNA's early warning signal to guide pre-emptive treatment. This would provide direct evidence on whether a dd-cfDNA-based strategy improves the trajectory of allograft health compared to the current standard of care. This potential for underreporting of negative or inconclusive studies represents a key limitation of the existing evidence base, and future meta-analyses should attempt to quantify the extent of this publication bias.

Code and Data Availability

Data collected as part of this systematic review is freely available at [redacted]. The Google Colab notebook used for analysis and plot generation is freely available at [redacted].

Acknowledgements

I would like to express my sincere gratitude to my mentor, [redacted], whose insightful feedback and critical comments at every stage were invaluable in shaping the direction and clarity of this systematic review. I am especially grateful for her mentorship in data science, as she patiently taught me the fundamentals of programming in Python to transform the raw data into the meaningful figures presented in the Results section. This project would not have been possible without her generous investment of time, knowledge, and unwavering support.

References

Akalin, E., Weir, M. R., Bunnapradist, S., Brennan, D. C., Delos Santos, R., Langone, A., Djamali, A., Xu, H., Jin, X., Dholakia, S., Woodward, R. N., & Bromberg, J. S. (2021). Clinical Validation of an Immune Quiescence Gene Expression Signature in Kidney Transplantation. *Kidney360*, 2(12), 1998. https://doi.org/10.34067/KID.0005062021

Akifova, A., Budde, K., Amann, K., Buettner-Herold, M., Choi, M., Oellerich, M., Beck, J., Bornemann-Kolatzki, K., Schütz, E., Bachmann, F., Halleck, F., von Hoerschelmann, E., Koch, N., Schrezenmeier, E., Seelow, E., Waiser, J., Zukunft, B., Eckardt, K.-U., Halbritter, J., … Osmanodja, B. (2025). Donor-derived cell-free DNA monitoring for early diagnosis of antibody-mediated rejection after kidney transplantation: A randomized trial. *Nephrology Dialysis Transplantation*, 40(7), 1384–1395. https://doi.org/10.1093/ndt/gfae282

Aubert, O., Ursule-Dufait, C., & Brousse, R. (n.d.). *Cell-free DNA for the detection of kidney allograft rejection | Nature Medicine*. Retrieved August 11, 2025, from https://www.nature.com/articles/s41591-024-03087-3

Aubert, O., Ursule-Dufait, C., Brousse, R., Gueguen, J., Racapé, M., Raynaud, M., Van Loon, E., Pagliazzi, A., Huang, E., Jordan, S. C., Chavin, K. D., Gupta, G., Kumar, D., Alhamad, T., Anand, S., Sanchez-Garcia, J., Abdalla, B. A., Hogan, J., Garro, R., … Loupy, A. (2024). Cell-free DNA for the detection of kidney allograft rejection. *Nature Medicine*, 30(8), 2320–2327. https://doi.org/10.1038/s41591-024-03087-3

Benning, L., Morath, C., Fink, A., Rudek, M., Speer, C., Kälble, F., Nusshag, C., Beimler, J., Schwab, C., Waldherr, R., Zeier, M., Süsal, C., & Tran, T. H. (2023). Donor-Derived Cell-Free DNA (dd-cfDNA) in Kidney Transplant Recipients With Indication Biopsy—Results of a Prospective Single-Center Trial. *Transplant International*, 36, 11899. https://doi.org/10.3389/ti.2023.11899

Benning, L., Morath, C., Kühn, T., Bartenschlager, M., Kim, H., Beimler, J., Buylaert, M., Nusshag, C., Kälble, F., Reineke, M., Töllner, M., Schaier, M., Klein, K., Blank, A., Schnitzler, P., Zeier, M., Süsal, C., Bartenschlager, R., Tran, T. H., & Speer, C. (2022). Humoral response to SARS-CoV-2 mRNA vaccination in previous non-responder kidney transplant recipients after short-term withdrawal of mycophenolic acid. *Frontiers in Medicine*, 9. https://doi.org/10.3389/fmed.2022.958293

Botella, C., Galián, J. A., Jiménez-Coll, V., Fernández-González, M., Morales, F.,

Martínez-Gómez, G., González-López, R., Alegría, M. J., Moya, M. R., Martinez-Banaclocha, H., Minguela, A., Legaz, I., Llorente, S., & Muro, M. (2024a). Early Monitoring of Donor-Derived Cell-Free DNA in Kidney Allograft Recipients Followed-Up for Two Years: Experience of One Center. *Life*, 14(11), 1491. https://doi.org/10.3390/life14111491

Botella, C., Galián, J. A., Jiménez-Coll, V., Fernández-González, M., Morales, F., Martínez-Gómez, G., González-López, R., Alegría, M. J., Moya, M. R., Martinez-Banaclocha, H., Minguela, A., Legaz, I., Llorente, S., & Muro, M. (2024b). Early Monitoring of Donor-Derived Cell-Free DNA in Kidney Allograft Recipients Followed-Up for Two Years: Experience of One Center. *Life*, 14(11), 1491. https://doi.org/10.3390/life14111491

Bromberg, J. S., Bunnapradist, S., Samaniego-Picota, M., Anand, S., Stites, E., Gauthier, P., Demko, Z., Prewett, A., Armer-Cabral, M., Marshall, K., Kaur, N., Bloom, M. S., Tabriziani, H., Bhorade, S., & Cooper, M. (2024). Elevation of Donor-derived Cell-free DNA Before Biopsy-proven Rejection in Kidney Transplant. *Transplantation*, 108(9), 1994–2004. https://doi.org/10.1097/TP.0000000000005007

Bu, L., Gupta, G., Pai, A., Anand, S., Stites, E., Moinuddin, I., Bowers, V., Jain, P., Axelrod, D. A., Weir, M. R., Wolf-Doty, T. K., Zeng, J., Tian, W., Qu, K., Woodward, R., Dholakia, S., Golovine, A. D., Bromberg, J. S., Murad, H., & Alhamad, T. (2022). Clinical outcomes from the Assessing Donor-derived cell-free DNA Monitoring Insights of kidney Allografts with Longitudinal surveillance (ADMIRAL) study. *Kidney International*, 101(4), 793–803. https://doi.org/10.1016/j.kint.2021.11.034

Chen, X.-T., Qiu, J., Wu, Z.-X., Zhang, H., Chen, T., Yang, S.-C., Zhao, G.-D., He, Y., Shen, X., Luo, J.-Q., Huang, Y., Wang, C.-X., Chen, L.-Z., Wu, C.-L., & Huang, G. (2022). Using Both Plasma and Urine Donor-Derived Cell-Free DNA to Identify Various Renal Allograft Injuries. *Clinical Chemistry*, 68(6), 814–825. https://doi.org/10.1093/clinchem/hvac053

Cheng, D., Liu, F., Xie, K., Zeng, C., Li, X., Ni, X., Ge, J., Shu, L., Zhou, Y., Shi, H., Liu, H., & Chen, J. (2021). Donor-derived cell-free DNA: An independent biomarker in kidney transplant patients with antibody-mediated rejection. *Transplant Immunology*, 69, 101404. https://doi.org/10.1016/j.trim.2021.101404

Cuadrado-Payán, E., González-Roca, E., Rodríguez-Espinosa, D., Molina-Andújar, A., Montagud-Marrahi, E., Arana, C., González-Rojas, A., Esforzado, N., Torregrosa, V., Ventura-Aguiar, P., Argudo, M., Serrano-Jorcano, D.,

Ramírez-Bajo, M. J., Bañón-Maneus, E., Casas, S., Broseta, J. J., Puig-Butille, J. A., Revuelta, I., Diekmann, F., & Cucchiari, D. (2025). Implementation of the in-house technique for the determination of donor-derived cell-free DNA in daily clinical practice: Experience from the Hospital Clinic of Barcelona. *Nefrología (English Edition)*, 45(4), 294–301. https://doi.org/10.1016/j.nefroe.2025.03.004

Dreige, Y., Lima, G., Cristelli, M., Nakamura, M., Foresto, R., Pierry, P., Glehn-Ponsirenas, R., Levi, J., Medina-Pestana, J., & Tedesco-Silva, H. (2022). 336.3: Clinical Utility of Donor-Derived Cell-Free DNA During the Period of Recovery of Renal Function After Kidney Transplantation. *Transplantation*, *106*(9S), S297. https://doi.org/10.1097/01.tp.0000886964.60815.ca

Halloran, P. F., Reeve, J., Madill-Thomsen, K. S., Kaur, N., Ahmed, E., Cantos, C., Al Haj Baddar, N., Demko, Z., Liang, N., Swenerton, R. K., Zimmermann, B. G., Van Hummelen, P., Prewett, A., Rabinowitz, M., Tabriziani, H., Gauthier, P., Billings, P., & Investigators*, the T. (2022). Combining Donor-derived Cell-free DNA Fraction and Quantity to Detect Kidney Transplant Rejection Using Molecular Diagnoses and Histology as Confirmation. *Transplantation*, *106*(12), 2435. https://doi.org/10.1097/TP.0000000000004212

Kidney Disease Statistics for the United States—NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved August 11, 2025, from https://www.niddk.nih.gov/health-information/health-statistics/kidney-disease

Kim, H. D., Bae, H., Kang, H., Lee, H., Eum, S. H., Yang, C. W., Choi, Y. J., Chung, B. H., & Oh, E.-J. (2024a). Donor-derived cell-free DNA predicted allograft rejection and severe microvascular inflammation in kidney transplant recipients. *Frontiers in Immunology*, *15*, 1433918. https://doi.org/10.3389/fimmu.2024.1433918

Kim, H. D., Bae, H., Kang, H., Lee, H., Eum, S. H., Yang, C. W., Choi, Y. J., Chung, B. H., & Oh, E.-J. (2024b). Donor-derived cell-free DNA predicted allograft rejection and severe microvascular inflammation in kidney transplant recipients. *Frontiers in Immunology*, *15*, 1433918. https://doi.org/10.3389/fimmu.2024.1433918

Kumar, N., Tandon, A., Rana, R., Rana, D. S., Bhalla, A. K., Gupta, A., Sachdeva, M. P., Huirem, R. S., Chauhan, K., Yashavarddhan, M. H., Basnal, A., Gupta, R., Mallick, P. K., & Ganguly, N. K. (2023). Donor-Derived Cell-Free DNA as a Non-Invasive Biomarker for Graft Rejection in Kidney Transplant Recipients:

A Prospective Study among the Indian Population. *Diagnostics*, 13(23), 3540. https://doi.org/10.3390/diagnostics13233540

Lakhani, L., Alasfar, S., Bhalla, A., Aala, A., Rosenberg, A., Ostrander, D., Schollenberger, M. D., Brennan, D. C., & Lipson, E. J. (2021). Utility of Serial Donor-derived Cell-free DNA Measurements for Detecting Allograft Rejection in a Kidney Transplant Recipient After PD-1 Checkpoint Inhibitor Administration. *Transplantation Direct*, 7(2), e656. https://doi.org/10.1097/TXD.0000000000001113

Loupy, A., Certain, A., Tangprasertchai, N. S., Racapé, M., Ursule-Dufait, C., Benbadi, K., Raynaud, M., Vaskova, E., Marchis, C., Casas, S., Hague, T., Bestard, O., Kervella, D., Lefaucheur, C., Viard, T., & Aubert, O. (2024). Evaluation of a Decentralized Donor-Derived Cell-Free DNA Assay for Kidney Allograft Rejection Monitoring. *Transplant International*, 37, 13919. https://doi.org/10.3389/ti.2024.13919

Mantios, E., Filiopoulos, V., Constantoulakis, P., Liapis, G., Vittoraki, A., Casas, S., Marinaki, S., & Boletis, J. N. (2023). Assessment of Donor Derived Cell Free DNA (dd-cfDNA) at Surveillance and at Clinical Suspicion of Acute Rejection in Renal Transplantation. *Transplant International*, 36, 11507. https://doi.org/10.3389/ti.2023.11507

Mayer, K. A., Doberer, K., Tillgren, A., Viard, T., Haindl, S., Krivanec, S., Reindl-Schwaighofer, R., Eder, M., Eskandary, F., Casas, S., Wahrmann, M., Regele, H., & Böhmig, G. A. (2021). Diagnostic value of donor-derived cell-free DNA to predict antibody-mediated rejection in donor-specific antibody-positive renal allograft recipients. *Transplant International*, 34(9), 1689–1702. https://doi.org/10.1111/tri.13970

Moein, M., Settineri, J. P., Suleiman, H., Sidhu, J., Papa, S., Coyle, S., Dvorai, R. H., Bahreini, A., Leggat, J., & Saidi, R. F. (2024). Application of Combined Donor-Derived Cell-Free DNA and Transcriptome in Diagnosis of Kidney Transplant Rejection. *Transplantation Proceedings*, 56(6), 1259–1263. https://doi.org/10.1016/j.transproceed.2024.02.023

Montgomery, R., & Gelb, B. (n.d.). *Four Things You May Not Know About Kidney Transplants*. NYU Langone News. Retrieved August 11, 2025, from https://nyulangone.org/news/four-things-you-may-not-know-about-kidney-transplants

Nguyen, P. T., Nakaoka, H., Mitsunaga, S., Aoyama, H., Kitamura, H., Saigo, K., & Inoue, I. (2025). *Applications of donor-derived cell-free DNA in kidney*

transplantation healthcare: View from a prospective single-center study. Research Square. https://doi.org/10.21203/rs.3.rs-6592444/v1

Nie, W., Wang, Y., Fu, Q., Wu, C., Deng, R., Yu, X., Ye, C., Liu, X., Xu, B., Sun, P., Liu, L., Li, J., Zhang, H., & Wang, C. (2025). Integrating Donor Derived Cell-Free DNA Fraction and Absolute Quantification for Enhanced Rejection Diagnosis in Kidney Transplant Recipients. *Diagnostics*, 15(3), 237. https://doi.org/10.3390/diagnostics15030237

Oellerich, M., Sherwood, K., Keown, P., Schütz, E., Beck, J., Stegbauer, J., Rump, L. C., & Walson, P. D. (2021). Liquid biopsies: Donor-derived cell-free DNA for the detection of kidney allograft injury. *Nature Reviews Nephrology*, 17(9), 591–603. https://doi.org/10.1038/s41581-021-00428-0

Orandi, B. J., Luo, X., Massie, A. B., Garonzik-Wang, J. M., Lonze, B. E., Ahmed, R., Van Arendonk, K. J., Stegall, M. D., Jordan, S. C., Oberholzer, J., Dunn, T. B., Ratner, L. E., Kapur, S., Pelletier, R. P., Roberts, J. P., Melcher, M. L., Singh, P., Sudan, D. L., Posner, M. P., … Segev, D. L. (2016). Survival Benefit with Kidney Transplants from HLA-Incompatible Live Donors. *New England Journal of Medicine*, 374(10), 940–950. https://doi.org/10.1056/NEJMoa1508380

Organ transplants exceeded 48,000 in 2024; a 3.3 percent increase from the transplants performed in 2023—OPTN. (n.d.). Retrieved August 11, 2025, from https://optn.transplant.hrsa.gov/news/organ-transplants-exceeded-48-000-in-2024-a-33-percent-increase-from-the-transplants-performed-in-2023/

Parajuli, S., Garg, N., Dodin, B., Breyer, I., Zona, E., Patel, S., Pinney, K., & Mandelbrot, D. (2024a). Changes in Donor-Derived Cell-Free DNA Before and After Rejection and De Novo DSA Detection in Primary and Repeat Kidney Transplant Recipients. *Clinical Transplantation*, 38(11), e70019. https://doi.org/10.1111/ctr.70019

Parajuli, S., Garg, N., Dodin, B., Breyer, I., Zona, E., Patel, S., Pinney, K., & Mandelbrot, D. (2024b). Changes in Donor-Derived Cell-Free DNA Before and After Rejection and De Novo DSA Detection in Primary and Repeat Kidney Transplant Recipients. *Clinical Transplantation*, 38(11), e70019. https://doi.org/10.1111/ctr.70019

Publish or Perish. (n.d.). Retrieved August 11, 2025, from https://harzing.com/resources/publish-or-perish

Quick kidney disease facts and stats | American Kidney Fund. (2024, April 19). https://www.kidneyfund.org/all-about-kidneys/quick-kidney-disease-facts-and-stats

Rizvi, A., Faiz, S., Thakkar, P. H., Hussain, S., Gamilla-Crudo, A. N., Kueht, M., & Mujtaba, M. A. (2023a). Kidney Allograft Monitoring by Combining Donor-Derived Cell-Free DNA and Molecular Gene Expression: A Clinical Management Perspective. *Journal of Personalized Medicine*, 13(8), 1205. https://doi.org/10.3390/jpm13081205

Rizvi, A., Faiz, S., Thakkar, P. H., Hussain, S., Gamilla-Crudo, A. N., Kueht, M., & Mujtaba, M. A. (2023b). Kidney Allograft Monitoring by Combining Donor-Derived Cell-Free DNA and Molecular Gene Expression: A Clinical Management Perspective. *Journal of Personalized Medicine*, 13(8), 1205. https://doi.org/10.3390/jpm13081205

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021a). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021b). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021c). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Schenk, K. M., Deutsch, J. S., Chandra, S., Davar, D., Eroglu, Z., Khushalani, N. I., Luke, J. J., Ott, P. A., Sosman, J. A., Aggarwal, V., Schollenberger, M. D., Sharfman, W. H., Bibee, K. P., Scott, J. F., Loss, M. J., Wang, H., Qi, H., Sharon, E., Streicher, H., … Lipson, E. J. (2024). Nivolumab + tacrolimus + prednisone +/− ipilimumab for kidney transplant recipients with advanced cutaneous cancers. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, 42(9), 1011–1020. https://doi.org/10.1200/JCO.23.01497

Sharma, R., Wolff, J., & Wilpula, E. (2022). Novel Technology Has Made It Possible To Diagnose A New Clinical Phenotype Of Rejection called Early Antibody Mediated Rejection, With Significant Therapeutic Implications For Renal Allograft Survival. *Experimental and Clinical Transplantation*, 20(10), 967–969.

https://doi.org/10.6002/ect.2021.0489

Shen, J., Guo, L., Yan, P., Zhou, J., Zhou, Q., Lei, W., Liu, H., Liu, G., Lv, J., Liu, F., Huang, H., Dong, W., Shu, L., Wang, H., Wu, J., Chen, J., & Wang, R. (2020). Prognostic value of the donor-derived cell-free DNA assay in acute renal rejection therapy: A prospective cohort study. *Clinical Transplantation*, 34(10), e14053. https://doi.org/10.1111/ctr.14053

Stites, E., Kumar, D., Olaitan, O., John Swanson, S., Leca, N., Weir, M., Bromberg, J., Melancon, J., Agha, I., Fattah, H., Alhamad, T., Qazi, Y., Wiseman, A., & Gupta, G. (2020). High levels of dd-cfDNA identify patients with TCMR 1A and borderline allograft rejection at elevated risk of graft injury. *American Journal of Transplantation*, *20*(9), 2491–2498. https://doi.org/10.1111/ajt.15822

Tian, Y., Frischknecht, L., Anna Mallone, Rossler, F., Schachtner, T., & Nilsson, J. (n.d.). *Evaluation of de novo donor specific antibodies after kidney transplantation in the era of donor-derived cell-free DNA - PMC*. Retrieved August 11, 2025, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11779610/

Wei, L., Zhao, Y., Deng, S., Wu, S., Wang, H., Luo, X., & Yang, H. (2024). Graft-derived cell free DNA: Used for assessment of early graft status and its implications for long-term kidney function. *Frontiers in Physiology*, *15*, 1440799. https://doi.org/10.3389/fphys.2024.1440799

Wolf-Doty, T. K., Mannon, R. B., Poggio, E. D., Hinojosa, R. J., Hiller, D., Bromberg, J. S., & Brennan, D. C. (2021). Dynamic Response of Donor-Derived Cell-Free DNA Following Treatment of Acute Rejection in Kidney Allografts. *Kidney360*, *2*(4), 729. https://doi.org/10.34067/KID.0000042021

Zhang, H., Zheng, C., Li, X., Fu, Q., Li, J., Su, Q., Zeng, L., Liu, Z., Wang, J., Huang, H., Xu, B., Ye, M., Liu, L., & Wang, C. (2020a). Diagnostic Performance of Donor-Derived Plasma Cell-Free DNA Fraction for Antibody-Mediated Rejection in Post Renal Transplant Recipients: A Prospective Observational Study. *Frontiers in Immunology*, *11*, 342. https://doi.org/10.3389/fimmu.2020.00342

Zhang, H., Zheng, C., Li, X., Fu, Q., Li, J., Su, Q., Zeng, L., Liu, Z., Wang, J., Huang, H., Xu, B., Ye, M., Liu, L., & Wang, C. (2020b). Diagnostic Performance of Donor-Derived Plasma Cell-Free DNA Fraction for Antibody-Mediated Rejection in Post Renal Transplant Recipients: A Prospective Observational Study. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.00342

Thank you for your valuable feedback on our manuscript. I have carefully addressed all your comments and suggestions in this revised version.

For clarity and ease of review, I have color-coded the text to highlight our changes:

- Red text indicates revisions made in response to Reviewer 1's comments.
- Blue text indicates revisions made in response to Reviewer 2's comments.

Please note that if both reviewers explicitly raised the same point, the correction is marked in red to maintain consistency.

I hope these revisions meet with your approval.

Sincerely,

Author 1

Submission title: Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection

Date of Review: 8/27/25

Decision: Revise and resubmit (major revisions needed, acceptance not guaranteed)

Resubmit by 9/26/25

Reviewer 1:

I am grateful for your valuable feedback and am glad you found my review to have originality and significance. I also want to thank you for identifying several key areas where I could strengthen the manuscript. After careful consideration, I have made substantial revisions based on your feedback, which I believe have greatly improved the clarity, structure, and rigor of my paper.

These major revisions are summarized as:

- Improving Manuscript Structure and Flow: Performing a significant reorganization by integrating the standalone "Background," "Literature Summary," "Limitations," and "Future Directions" sections into a more cohesive and logical narrative.

- Enhancing Methodological Transparency: Adding a PRISISMA flow diagram (Figure 2) and a comprehensive summary table (Table 1) to clearly detail the study selection process and the characteristics of the included articles.

- Strengthening the Analysis: Contextualizing my findings by adding a new discussion that compares the results of this review to other published meta-analyses and systematically correcting all figure references throughout the manuscript.

- Refining Language and Precision: Performing a manuscript-wide proofread to correct stylistic errors, improve conciseness, and clarify the specific criteria and terminology used for the quality assessment and study designs.

Comments:

- Originality & Significance – Does the paper contribute new insights or perspectives to the field?
    - Yes. Very interesting points made and clear path of innovation to the field of renal transplant
        - Much appreciated with your encouraging comments. Thank you.

- Clarity & Structure – Is the argument well-organized and easy to follow? Are ideas clearly presented?
    - Flow and presentation of information can be significantly improved.
        - Absolutely. I have performed a substantial reorganization of the paper to address this concern, with the following major changes:
            - I removed the summary of results from the end of the introduction to ensure the section now properly focuses on establishing the background and rationale for the review (lines 121-129).
            - The standalone "Background" section has been eliminated, and its essential content has been integrated into the introduction to create a more cohesive narrative.
            - I removed the entire "Literature Summary" section, which the reviewer correctly identified as interrupting the flow between the Methods and Results.
            - The "Overview" portion of the former "Literature Summary" has been relocated to the end of the introduction to better contextualize the articles being reviewed (lines 135-151).
            - The detailed subsections from the "Literature Summary," such as "Identification of clinically relevant subtypes," have been moved into the main "Results" section to be discussed alongside our synthesized data (lines 894-904).
            - The standalone "Limitations" and "Future Directions" sections have been removed and their content condensed into concise paragraphs at the end of the manuscript to create a more conventional structure (lines 1022-1051).
        - These substantial revisions should create a more linear, well-organized argument that is easier for the reader to follow.
- Use of Evidence & Research Methods – Are sources appropriately cited? Is their methodology sound and well-explained?

    - Yes. Very clear, excellent work
        - Thank you for your encouraging comments.
- Engagement with Literature – Does the paper demonstrate an understanding of relevant research in the field? Do they acknowledge known results and connect their findings well to them?
    - There is a clear understanding of relevant research but there is little connection of current research to the analysis performed in this review. For example, are there other literature reviews/meta-analyses that engage with a similar topic and how do their findings relate to the findings of this manuscript?
        - Acknowledged. I've added a new paragraph at the beginning of the Discussion section that explicitly compares the conclusions of my systematic review to those of other recent meta-analyses and reviews on this topic. This new section clarifies the following points:
            - Agreement on Diagnostic Accuracy: I now explicitly state that my primary finding—that dd-cfDNA demonstrates high and consistent diagnostic accuracy—is in strong agreement with the conclusions of prior reviews, such as those by Oellerich et al. (2021) and Gisch et al. (2025).
            - Alignment with Meta-Analyses: I also highlight that the AUC values synthesized in my review align with the high pooled AUCs reported in several recent meta-analyses, including those by Yang et al. (2024), Zhang et al. (2021), and Zhang et al. (2025).
            - Novel Contribution of My Review: I then clarify the novel contribution of my work, which is the first to systematically quantify both the methodological maturation of the field and the significant geographic and ancestral disparities in the published evidence—a key knowledge gap not previously analyzed in this depth.
        - This addition should better connect the analysis to the broader field and clearly articulates the unique contribution of my manuscript. This new discussion can be found on lines 540-561.

- Grammar & Language – Is the writing clear and professional? Minor grammatical and stylistic errors should be noted, but they should not be the main focus of the review.
    - Stylistic/grammar errors are noted throughout the manuscript
        - Thank you for your feedback. I've now performed a thorough proofread of the entire manuscript to improve its clarity, conciseness, and professionalism. For example, I've:
            - Fixed Typos and Word Choice: I corrected several typos and improved word choice, such as replacing "good" with "favorable" in the abstract on line 23.

        - Revised awkward parenthetical phrasing for better flow, such as on line 273 where “(e.g. kidney)” was streamlined to “, like kidney”.
        - Corrected date formatting for consistency, adding specific days such as "1st" and "31st" to the date range in my inclusion criteria on lines 385-386.
    - These comprehensive edits should significantly improve the quality of the writing.

- Figures are incorrectly referenced throughout the manuscript
    - Thank you for catching this critical error. Several figure references were incorrect due to renumbering. I have since performed a systematic check of my entire manuscript to correct every in-text figure citation.
        - For verification of these corrections, you can check lines: 616, 623, 636, 644, 648, 661, 697, 704, 719, 720, 722, 736, 778, 796, 807, 817, 823, 846, 848, 850, 857, 862, 863, 869, 875, 882, 884, 890, and 892.

Suggested revisions:

- Abstract
    - Usage of the word “good” throughout does not give descriptive information, consider using a more descriptive adjective
        - Acknowledged. I have revised the manuscript to explicitly state the criteria used for my quality assessment:
            - In the Abstract, I now clarify that the literature was deemed of "acceptable quality" specifically because most studies scored 3 or more out of 5 on my author-created scoring system, which includes criteria like "Longitudinal Monitoring for Timing" and "Compares to Traditional Biomarkers?".
            - In other instances where "good" was used, such as describing diagnostic performance, I have replaced it with more descriptive terms like "favorable," which is immediately supported by the quantitative evidence of high Area Under the Curve (AUC) values presented in the text.
        - This approach should provide the underlying criteria, rather than just swapping subjective adjectives, more rigorously addresses your valid concern.
    - “criteria created by the author”
        - To provide full transparency on my quality assessment methodology, I've revised the Abstract to list the five specific criteria used for my scoring system. The manuscript now explicitly states the criteria as: “Measured Blood?”, “Analyzes rejection probability?”, “Temporal Precedence

Explicitly Stated", "Longitudinal Monitoring for Timing", and "Compares to Traditional Biomarkers?"
        - You can find this clarification on lines 20-22 of the revised Abstract. I believe this addition makes my methodology much clearer.

- Introduction
    - "In short, these significant barriers…" this sentence is wordy, try to make more concise. Consider choosing "unfulfilled need space" or "unmet medical need"
        - Acknowledged. I streamlined the sentence by selecting the more impactful phrase "unmet medical need" and removing the redundant wording. I believe the sentence is now much clearer. You can find this change on lines 54-55.
    - 4th paragraph of introduction is typically included in the discussion or conclusion sections. Avoid including results or analysis in the introduction
        - Thank you, I have removed the results-focused fourth paragraph from the introduction to ensure it now properly establishes the background and rationale for my study. You can find this revision on lines 121-129.
- Background
    - Background sections are not typically included in scientific/clinical manuscripts as the relevant background should be included early in the introduction
        - I appreciate the feedback.
        - For systematic reviews, a dedicated background section is often used to provide the comprehensive context needed to understand the rationale for the review's specific research question. This structure is common in the field, including in high-quality systematic reviews such as those published by Cochrane, which often utilize a distinct background section to frame the clinical landscape before detailing the review's objectives.
            - I have substantially condensed my introduction by removing sentences that provided more foundational, explanatory details. Specifically, I removed lines 156-158, 160-168, and 179-188. I did this to focus the introduction more tightly on the core problem while leaving the broader context to the background section.
        - This revised structure should align with common practice for systematic reviews while improving the introductory flow.
    - Avoid using paratheses out of usage for sources/citations
        - Acknowledged. I've reviewed the manuscript and removed unnecessary parentheses to improve readability.
            - I kept parentheses where they are correctly used to define standard abbreviations on their first use, like for ESRD, qPCR, and NGS. As a specific example of the corrections, on line 273 I revised the awkward parenthetical phrase '(e.g. kidney)' to the more streamlined ', like kidney'

    - Background section does not discuss dd-cfDNA to the extent of the other non-invasive methods but was stated as the main objective of the literature review. Is cfDNA the same as dd-cfDNA? This is not clear but there is a section on general cfDNA. Consider shortening other sections to lengthen background of dd-cfDNA specifically
        - I appreciate your feedback on the background section. The content of my original section can be further enhanced.
            - I kept the section but substantially revised it to address your specific points. To improve the focus and clarity, I made the following additions:
            - I added a new paragraph on lines 160-168 that explicitly defines the difference between total cfDNA and the donor-derived fraction (dd-cfDNA). This new text clarifies that dd-cfDNA is the specific, organ-derived biomarker of interest for transplant monitoring.
            - I also added a more detailed paragraph on lines 179-188 that expands on the clinical utility of dd-cfDNA. This section now discusses its short half-life and its mechanism as a near real-time indicator of allograft injury, explaining how quantifying its increase can lead to earlier rejection detection.
- Methods
    - Inclusion criteria of Jan 2000-Jan 2025 does not align with the search process restriction of Jan 2020-May 2025
        - Thank you. I have updated the date range in the 'Inclusion Criteria' section on lines 385-386 from “January 2000 and January 2025” to say "January 1st, 2020, to May 31st, 2025". This ensures it now precisely matches the date range stated in my 'Search Strategy' section. I apologize for the oversight.
    - Was the study scoring system created by the author or adapted from an accepted approach? Please clarify.
            - I appreciate you requesting this clarification.
                - I've updated the 'Study Scoring and Quality Assessment' section to clarify that the scoring system was created by me. I also added a detailed paragraph explaining the rationale for choosing each of the five criteria and how they directly relate to answering my specific research question. You can find this clarification on lines 410-411 and 416.
-
- Literature summary section
    - Unsure of the usage of this section. Consider including this highly specific information in the results section and only use information that is important to the overall thesis of evaluating dd-cfDNA

        - Acknowledged. I have revised the manuscript to create a more logical and readable structure.
        - To improve the manuscript's flow, I've removed the standalone "Literature Summary" section and relocated its contents into more appropriate sections. The "Overview" portion (previously lines 474-490) has been moved to the end of the introduction to better frame my review's scope (now on lines 135-151).
        - The remaining, more detailed subsections have now been integrated as distinct subheadings within my main "Results" section, allowing these topics to be discussed alongside my synthesized findings. Specifically:
            - A subpart of the section “Evaluating the efficacy of dd-cfDNA in estimating rejection” (from lines 492-500) has been moved to lines 448-453.
            - The entirety of the “Early Detection with dd-cfDNA" section (from lines 502-507) has been moved to lines 592-597.
            - The content from the old “Identification of clinically relevant subtypes” section (from lines 509-525) has been relocated under a new, more clearly titled heading, “Identification of Clinically Relevant Subtypes,” on lines 574-584.
            - Finally, the “Performance in Multiple Ancestries” section (from lines 527-533) has been moved to lines 606-612.
        - I am confident this new structure directly addresses your concern by eliminating the standalone summary and placing its content in more appropriate narrative locations.
- Analysis
    - This section is more aptly titled “Discussion” as there is interpretation of relevant results
        - Absolutely. A "Results" section is standard. For my manuscript, the "Analysis" section serves the function of the results section by providing objective information and findings from my systematic search. I chose this naming convention to align with the submission guidelines for the Convergence Journal, which requests this specific heading.
        - However, to make the section's purpose clearer to a broader audience, I've retitled the section to "Results". You can find this change on line 534. I hope this is an acceptable compromise that both adheres to the journal's guidelines and improves the manuscript's readability.

    - Consider using terms such as “prospective cohort” design; overall being more specific about the design of the papers used in the search (as done in Figure2e/f)
        - Acknowledged. I have revised the text throughout to be more specific, as you recommended. For example:

        - On lines 410-411 and 416, I've clarified the origin of my scoring system by stating it was "created by the author" and provided the rationale for its design.
        - On lines 635-638, I streamlined my description of the literature by replacing vaguer phrasing with the more precise term "prospective cohort designs" and clarified why these studies provide a strong evidence base.
        - On lines 658-662, I significantly condensed and clarified a convoluted sentence describing the geographic distribution of study types, making the point more direct and readable.
    - I believe these changes make my methodological descriptions more specific and my language smoother, directly addressing your point.

- Excellent breakdown of the designed 5-point scoring system.
    - I appreciate your positive feedback on the scoring system.

- Figure 2a does not report overall evidence base being high quality but rather publication year of selected papers. Please revise.
    - Thank you. I've removed the inaccurate sentence. I also revised the subsequent sentence on lines 635-638 to be more precise, replacing the general phrase "high-quality study designs" with the more specific term "prospective cohort designs" and clarifying why these studies form a strong evidence base. I apologize for the error.

- Figures are incorrectly referenced throughout the analysis section.
    - Acknowledged. Several figure references were incorrect due to the addition of a new figure during my revisions, which shifted the numbering of all subsequent figures.
        - I have now performed a systematic check of my entire manuscript and can confirm that every in-text figure citation has been corrected, with most figure numbers being incremented by one or two. This was a comprehensive revision affecting numerous citations, including those on lines 616, 623, 636, 644, 648, 661, 697, 704, 719, 720, 722, 736, 778, 796, 807, 817, 823, 846, 848, 850, 857, 862, 863, 869, 875, 882, 884, 890, and 892.

- Unsure of the usage of Figure 4 as it does not relate to the research hypothesis of dd-cfDNA. This information is interesting but not useful enough to have a dedicated figure or paragraph.

        - I appreciate your comment regarding the direct relevance of Figure 4.
            - While the figure may not directly address the primary hypothesis, it is foundational to a key secondary objective of my review: to identify knowledge gaps and assess the biomarker's utility in diverse, underrepresented patient populations. The data in this figure provides the concrete evidence for a major limitation I discuss later—the lack of global data for dd-cfDNA's performance in Black and Hispanic populations, with evidence coming almost exclusively from North America.
        - With these study goals in mind, I feel that this information is a critical contribution of my review. I have therefore kept the figure and its accompanying analysis, and I hope this explanation clarifies its importance to my manuscript's overall objectives.

    - Description of AUC and ROC is not usually in discussion sections but can be moved to the methods section.
        - Yes. I have relocated the descriptive paragraph defining the Area Under the Receiver Operating Characteristic curve from its original position on lines 736-744 in the 'Results' section to its new location on lines 455-463 within the 'Data Analysis' subsection of the Methods. This change ensures all methodological details are presented together before the results are introduced.

    - Discussion of Figure 5 is very interesting, great work.
        - I appreciate your encouraging feedback on my analysis and visualization of the performance trends.

    - Discussion of ABMR and TCMR is very important in the field of transplant, should be included earlier in the discussion/analysis section
        - Yes. Discussing the important rejection subtypes of ABMR and TCMR earlier enhances the manuscript's flow.
            - I have swapped the order of the subsections. The section "Identification of Clinically Relevant Subtypes" has been moved up to appear before the section on "Early Detection with dd-cfDNA". You can find this change on lines 573-597.
- Limitations:
    - Paragraph three is not expected in the limitations section but best fitted at the end of the discussion section or as a stand-alone conclusion paragraph
        - Thank you. I have moved this paragraph from the limitations section, originally lines 932-942, to the end of the discussion section, now on lines 894-904.

- Limitations are usually included at the end of the discussion section rather than as its own section
  - Yes. I have removed the standalone "Limitations" section, originally lines 914-942, and have moved its content to the end of the 'Results' section, now on lines 894-904.

- Consider condensing this section into highly salient points. Should be no more than 1 paragraph (~7-8 sentences)
  - Acknowledged. I have condensed it into a single, more focused paragraph. You can find this revised section on lines 894-904.

- Future directions and Recommendations
  - Future directions are usually included at the end of the discussion section rather than as its own section
    - Thank you. I appreciate it. I have removed the standalone "Future Directions and Recommendations" section (originally lines 945-989) and have moved its content to the end of the 'Results' section (now on lines 894-904).

  - Consider condensing this section into highly salient points specific to future driections rather than continued interpretation of study results. Should be no more than 1 paragraph (~7-8 sentences)
    - Understood. I have condensed it into a single, more focused paragraph. You can find this revised section on lines 1022-1040 and 1043-1051.

  - The third paragraph is not expected in future directions section, best suited for end of discussion section or conclusion
    - Acknowledged. I have moved this paragraph and condensed it with the rest of the future directions into a single paragraph at the end of the 'Results' section. You can find this revised section on lines 1043-1051.

- Conclusion section
  - Well done, clear, concise.
    - I appreciate your positive feedback on the conclusion.

- Overall comments:
  - There is no results section, which is expected of a clinical/medical manuscript. Results sections provide objective information.
    - Thank you. You are correct that a "Results" section is standard. For my manuscript, the "Analysis" section serves the function of the results

section by providing objective information and findings from my systematic search. I chose this naming convention to align with the submission guidelines for the Convergence Journal, which requests this specific heading.

- However, to make the section's purpose clearer to a broader audience, I've retitled the section to "Results". You can find this change on line 534. I hope this is acceptable that both adheres to the journal's guidelines and improves the manuscript's readability.

- There is no Table 1 to describe the studies selected, which is expected of a clinical/medical manuscript. Figure 2 attempts to convey this information visually but should also be compiled into a Table 1, therefore defeating the purpose of Figure 2. See published literature reviews or meta-analyses for examples.
  - Thank you. A summary table is a standard and valuable component of a systematic review and that my manuscript would benefit from its inclusion.
  - I've made several key additions to the manuscript to enhance reporting and transparency:
    - I added a new Table 1 that provides key characteristics of the included studies in a clear, tabular format, with a short description added on line 399.
    - I also added a PRISMA flow diagram to visually represent the study selection process on line 333.
    - To accommodate these additions, all subsequent figures have been renumbered.
  - I have retained the original figure (now renumbered) as I believe it provides a complementary, visual synthesis of the trends within the literature that is distinct from the data presented in the new table and PRISMA diagram. I believe these additions, taken together, significantly strengthen the manuscript.

- Summary or conclusion-adjacent paragraphs are littered throughout this manuscript. All summary paragraphs should be revised into one conclusion section
  - Acknowledged. I have moved the summary paragraph from the end of the discussion section (originally lines 894-904) and integrated it into my final "Conclusion" section (now on lines 1022-1032) to create one cohesive summary.

- Consider including a flow chart figure to describe the paper selection process visually rather than multiple bar charts.

        - Absolutely. I have added a PRISMA 2020 flow diagram on line 333. This addition has caused all subsequent figures to be renumbered accordingly. I believe this is a significant improvement that greatly enhances the manuscript's methodological transparency.

    - Figure 1 is excellent and provides great visual information. Great usage of BioRender.
        - I appreciate your positive feedback on Figure 1.

Reviewer 2:

I want to thank you for your thorough and insightful review of my manuscript. I appreciate you taking the time to provide such constructive feedback, which has helped me to significantly improve the paper. After careful consideration of all your points, I have made substantial revisions that I believe have strengthened the manuscript's clarity, rigor, and overall impact.

The major revisions are summarized as:

- Clarifying the Abstract and Background: Revising the abstract to include more specific details about the search and selection process, and expanding the background section to better define key terms and discuss the clinical relevance of dd-cfDNA's half-life.

- Strengthening Statistical Rigor: Performing and reporting formal statistical tests (e.g., linear regression, chi-square, ANOVA) to validate the trends and comparisons made in the 'Results' section, as you recommended.

- Improving the Analysis and Discussion: Adding new subsections to the analysis to provide a direct comparison of dd-cfDNA to serum creatinine and to summarize the monitoring frequencies found in the literature.

- Refining Key Concepts: Revising the language used to discuss patient ancestry to be more precise and cautious, and clarifying the purpose of the author-created quality assessment tool to better frame its scope.

I have reviewed the following manuscript under consideration for publication, titled "Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection"

Overall, this paper is insightful, original, and significant to the body of literature. The objectives of the paper are well-stated and the paper is well-organized. The methodology of the paper is . The author should consider using a widely accepted metric of quality assessment to compare rigor of papers and more information regarding the statistical methods and the outcomes of statistical tests should be included. Additionally, the results section should include more information regarding the content and structure of the papers included in this analysis so that the reader can understand the quality and type of studies included in the analysis. The grammar and language is clear, professional, and concise. In summary, this paper is applicable to the field and answers a very relevant question. I recommend: Acceptance with major revisions (major revisions bolded below). Please find my comments below.

Abstract:

- Include more information about the search (databases, how many studies were reviewed vs. included, etc.)
    - Thank you. I have revised the abstract to include more specific information about my search process, as you recommended. I now specify that the search was conducted on the Google Scholar and PubMed databases and briefly summarize the selection process, noting that 83 studies were ultimately included after a manual screening and quality assessment. You can find this updated sentence on line 13-15.

- Give a little background information about the criteria created by the author to assess quality
    - You may consider using a widely utilized quality metric for ease of understanding by the audience
    - This is a fair point, which was also raised by Reviewer 1. To clarify my methodology, I've expanded on the rationale for my 5-point scoring system on lines 20-22.
    - The criteria were specifically chosen to assess each study's direct relevance to my research question and their relevance to ideal cfDNA study designs. For instance:
        - "Measured Blood?" and "Analyzes Rejection Probability?" ensured the studies used the correct sample type and focused on the primary clinical outcome.
        - "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for Timing?" were crucial for evaluating the "earlier identification" aspect of the hypothesis.
        - "Compares to Traditional Biomarkers?" verified that the studies performed a comparative analysis against standard methods.
    - I believe this additional detail provides the necessary background and justification for my assessment approach.

- You state that this is a usable technique but don't state that your review assessed the feasibility of using this in clinical practice
    - Thank you. I have revised the final sentence of the abstract to explicitly state that this review did not assess clinical feasibility, and I've noted that future work must address this gap. You can find this important clarification on lines 33-37.

- Great job identifying the gap in the literature and the identification of future directions
    - I appreciate your positive feedback on the literature gap analysis and identification of future directions.

Introduction:

- Unclear what you mean by “poor risk management framework” - this should be expanded to make it very clear the problems that exist with traditional monitoring for rejection
    - Thank you for your feedback. I've clarified the meaning of "poor risk management framework" to make the problems with conventional monitoring practices more explicit. I have added a sentence explaining that this framework is limited by its reliance on invasive, reactive biopsies and non-specific, lagging biomarkers such as serum creatinine. This revision can be found on line 47-51.

- Include some background information about the practical steps of dd-cfDNA – is there something that needs to be measured from the donor at transplantation time? How is it detected?
    - Thank you for your suggestion. I've included a paragraph in the Introduction section that details the general intent of the articles included in the review. This text now specifies that the studies aimed to assess dd-cfDNA's potential as a convenient, non-invasive biomarker, and that they often compared its performance to traditional metrics like serum creatinine. The new paragraph also clarifies that the articles focused on identifying allograft injury sooner, differentiating between rejection subtypes, and using dd-cfDNA for longitudinal monitoring. The added text can be found on lines 135-151.
- **You identified that one issue with traditional surveillance is that it doesn’t identify the issue until the tissue/organ damage is done. When you introduce dd-cfDNA you state that these cells are released into circulation when the organ is inflamed or injured – is it unclear how this is different from the traditional method. Is there a baseline level that one would expect? Would the provider be monitoring for increases?**
    - I appreciate you raising this important point. I have added a paragraph to clarify the key difference between dd-cfDNA and traditional methods. The text now explains that dd-cfDNA exists at a low baseline level in a healthy organ and that clinicians monitor for a measurable increase in these levels, which indicates a real-time allograft injury. This approach allows for earlier detection and intervention. The revision can be found on lines 68-77.

- Great job stating the hypothesis and the objectives of the study
    - I am glad to hear that the hypothesis and objectives of the study are clearly stated. I am grateful for your positive feedback on this aspect of the manuscript.

- The summary of the findings of the review should be moved from the introduction to the discussion/conclusion
    - I appreciate this feedback on the structure of the manuscript. I've relocated the two paragraphs that summarize the review's findings and the intent of the included studies from the Introduction section to the end of the Conclusion. This adjustment improves the overall flow and organization of the paper, allowing the Introduction to focus on the problem and rationale while the conclusion effectively summarizes the key findings, as you suggested.
    - The revised text can be found on lines 114-121 and 129-151, and the new location is at the end of the Conclusion section on lines 991-1000.

- More information on the current state of research in this field – are there other reviews to date?
    - Acknowledged. I have revised the Introduction to include a shorter, more direct discussion of previous reviews and meta-analyses. The new text clarifies how this systematic review builds upon the established literature and addresses a key knowledge gap regarding diversity in patient populations. The new text can be found on lines 83-97.

Background:

- You state the kidney biopsy is the gold standard for monitoring graft health which contradicts the statements in your introduction that biomarkers are used for monitoring. The role of kidney biopsy vs. biomarkers should be clarified.
    - Understood. I have revised the Background section to clarify the distinct roles of kidney biopsy and biomarkers. The new text explains that while biopsy is the gold standard for definitive diagnosis, its invasiveness, combined with the lagging nature of conventional biomarkers, makes it unsuitable for the frequent surveillance necessary to prevent graft injury. This change clarifies that biomarkers are used for monitoring, while biopsy is reserved for diagnosis.
    - The revision can be found on lines 168-171.

- Swap the order of discussion in the background section - discussing the kidney transplant and reject background before discussing how it is monitored makes more sense
    - Thank you. I have reordered the subsections in my background section, as you suggested. I moved the paragraphs on lines 153-516, 158-160, 168-171, and 173-179, which talked about how transplantation is monitored, to the end of the background section on lines 295-314. I agree that this revised structure is much clearer for the reader.

- Add statistics regarding the morbidity and mortality of ESRD / kidney transplant to the background to make it clear the severity of this disease and how many people it affects
    - Yes. I have now added statistics to the background section in two key places:
        - On lines 198-201, I added, "The disease carries a significant burden, with a mortality rate of nearly 25% for patients on dialysis, highlighting the critical need for transplantation as a life-saving intervention."
        - On lines 214-216, I added, "Graft failure is a severe outcome, leading to a return to dialysis where the 5-year survival rate is only around 42%, underscoring the importance of effective allograft monitoring."
    - These additions should provide a much stronger and clearer picture of the clinical stakes involved.

- Figure 1 is blurry and the text cannot be read. Also, is this an original figure? If so, you should credit how and by whom it was created. If not, a citation is required.
    - Thank you for the feedback. To address this, I have enlarged Figure 1a and 1b to make the text more readable while ensuring the image quality remains at a high resolution of 600dpi. I have also updated the figure caption to clarify that this is an original figure created by me using BioRender, as indicated by the watermark in the bottom right corner of each panel.

- You make the point in your introduction that current biomarkers can only detect problems once the damage is done. It is unclear how this is different as it is only released with cell damage. I would clarify this as it speaks to the utility of this method
    - Acknowledged. I have revised my background section on lines 311-314 to more clearly explain the difference between the "cell damage" detected by dd-cfDNA and the irreversible damage indicated by traditional biomarkers. I now explicitly state that dd-cfDNA reflects the real-time process of cell death, allowing for the detection of early or low-level injury, in contrast to conventional markers that only rise after substantial, cumulative, and often irreversible functional damage has already occurred.
    - This addition should better highlight the key utility of this method.

- Given the short half-life of cfDNA – is there a possibility of it rising and falling and not being detected?
    - Yes, a brief, transient spike in dd-cfDNA could theoretically be missed between tests due to its short half-life. However, clinically significant rejection is a sustained process that causes a prolonged elevation of dd-cfDNA, which is reliably detected by the next scheduled surveillance blood draw.

Methods:

- Good overview of the search terms and search methods
    - I appreciate your positive feedback on the search terms and methods section.

- Did you include/exclude literature review, systematic reviews,metaanalysis?
    - Thank you. This is an excellent clarifying question. I apologize for not making that explicit in my methods section.
    - You are correct that I did not include literature reviews, systematic reviews, or meta-analyses, as my goal was to synthesize primary data from original research articles. I have now updated my Inclusion Criteria section on lines 386-387 to explicitly state that these article types were excluded.

- **Unclear how many studies were identified from the search, when/how many were excluded at different stages of the process. Consider using a PRISMA flow diagram to describe this**
    - Thank you for your suggestion to improve the visualization of my methods. This was also recommended by Reviewer 1, and I agree that a flow chart is the standard and most effective way to depict the study selection process.
    - As you recommended, I have added a PRISMA 2020 flow diagram on line 333. To incorporate this, I have revamped the first sentence of the Selection Process subsection on lines 390-391 to directly reference this new figure. I believe this is a significant improvement that greatly enhances the manuscript's methodological transparency.

- **Criteria stated do not fully assess study methodological rigor as stated - you should consider using a widely accepted method**
    - You are correct that my author-created criteria do not fully assess a study's overall methodological rigor in the way a standardized tool would. I apologize if my language was not clear on this.
    - The goal of my scoring system was not to assess general methodological quality, but rather to specifically evaluate each study's relevance to my unique research question. To clarify this distinction, I have revised the 'Study Scoring and Quality Assessment' section on lines 411-412 to explicitly state that the criteria were designed to assess "direct relevance... rather than its general methodological rigor." This clarification should accurately frame the purpose of my scoring system.

Data Analysis:

- Were there any comparisons made between papers? I would make it clear how this was done.
    - Thank you. I have revised the 'Data Analysis' section on lines 441-446 to be more explicit about how comparisons were made between papers. I now clarify that a formal statistical meta-analysis was not performed. Instead, I explain that my review uses a narrative synthesis of the findings, with comparisons between studies illustrated through the data visualizations presented in the 'Results' section.

Literature Summary:

- The overview provided here is a bit redundant as you have introduced this topic in the introduction and background section. I recommend consolidating the information here
    - Thank you. This was also raised by Reviewer 1. I agree that its content was redundant and should be integrated into the main narrative, and I have revised the manuscript to create a more logical and readable structure.
    - To improve the manuscript's flow, I've removed the standalone "Literature Summary" section and relocated its contents into more appropriate sections. The "Overview" portion (previously lines 474-490) has been moved to the end of the introduction to better frame my review's scope (now on lines 135-151).
    - The remaining, more detailed subsections have now been integrated as distinct subheadings within my main "Results" section, allowing these topics to be discussed alongside my synthesized findings. Specifically:
        - A subpart of the section “Evaluating the efficacy of dd-cfDNA in estimating rejection” (from lines 492-500) has been moved to lines 448-453.
        - The entirety of the “Early Detection with dd-cfDNA" section (from lines 502-507) has been moved to lines 592-597.
        - The content from the old “Identification of clinically relevant subtypes” section (from lines 509-525) has been relocated under a new, more clearly titled heading, “Identification of Clinically Relevant Subtypes,” on lines 574-584.
        - Finally, the “Performance in Multiple Ancestries” section (from lines 527-533) has been moved to lines 606-612.
    - I am confident this new structure directly addresses your concern by eliminating the standalone summary and placing its content in more appropriate narrative locations.

- The “evaluating the efficacy of dd-cfDNA” section belongs in the methods sectional as it discusses how you evaluated the performance

    - Thank you. This was also highlighted by Reviewer 1, and I have already corrected it in the revised version of the manuscript. The description of how I evaluated performance using the AUROC curve has been moved from its old location on lines 492-500 and is now in the 'Data Analysis' subsection of the Methods section on lines 448-463. I agree this is a much more logical placement.

- I would introduce the concept of the Dd-cfDNA 30 minute half-life and how this is used in clinical practice, in the introduction/background sections
    - Thank you. Agreed that the half-life is an important concept to introduce early for context.
    - I have consolidated and moved the discussion of the 30-minute half-life of dd-cfDNA from its previous location (lines 502-507) to the Background section on lines 68-77, where it provides better context.

- **You need a more rigorous overview of studies included (sample size, methods, multicenter vs single center, comparison methods, etc.) and their main findings overall before discussing the most key findings overall. This section is brief and doesn't give a comprehensive summary of the papers that are included / will be analyzed in this paper. You may also consider including a table 1 that discusses all of the studies included, their methods,and findings that were compared**
    - Thank you and acknowledged. This was also raised by Reviewer 1. I agree a more rigorous overview of the included studies is necessary and that a summary table is the best way to present this information.
    - I have made these additions to the manuscript:
        - I have included a PRISMA flow diagram (Figure 2) to detail the study selection process on line 333.
        - I have also added a comprehensive Table 1 that summarizes the key characteristics of the included studies on line 399, such as sample size, study design, and primary findings.
    - I believe these additions provide the comprehensive summary you're looking for and significantly strengthen my manuscript's methods section.

Analysis:

- You discuss a steady increase over the years however, it isn't clear if there were any statistical tests employed. Please clarify how you determined that the increase was steady overtime. A test comparing the numbers over the years and a p-value would be great here.
    - Thank you. I have performed a linear regression analysis to formally test the trend of publications over time. I have now added the results of this test, including

the slope and p-value, which confirm a significant positive trend, to the 'Results' section where I discuss the distribution of publication years on lines 618-621. I agree that this addition makes my claim much more rigorous.

- **It seems like some of your figures have confidence intervals however, there is no mention of the statistical test used to compare and how you determined increase/decreases over time. Please consider employing statistical tests for these comparisons to make the analysis more rigorous**
    - To this point, the confidence intervals appear to be overlapping – I don't believe that you can say that there has been an increase in sample size over time if the difference is not statistically significant. This should be made clear to the reader
    - I appreciate you highlighted a crucial point about statistical rigor. I agree that a visual trend might not suffice, and that the overlapping confidence intervals in the box plot might require a formal statistical test to interpret correctly.
        - I have performed an analysis of variance (ANOVA) to test for a significant change in study sample size over the years. As you suspected, the analysis showed that the observed increase is not statistically significant. I have now replaced my original sentence on lines 627-630 with a new, more accurate statement on lines 630-632 that reflects this non-significant result. I believe this makes my analysis much more accurate and rigorous.

- Consider restructuring this sentence – "The choice to go through the analysis this way was significant because this page had a demonstrated blend of complementary types of research to provide a wider and more diverse evidence base - the prospective studies provide higher quality-controlled data while the retrospective and observational studies filled in the gap of real-word evidence, which in sum produced better conclusions for this review."
    - Absolutely. I have completely revamped the original sentence on lines 653-658. The new, easier-to-follow version on lines 658-664 is broken into two shorter, more direct sentences that I believe are much clearer.

- Again, when comparing distribution by location, please employ a statistical test
    - Your point about needing a statistical test for the geographic comparison is well taken.
    - I have now performed a chi-square test to compare the distribution of study types across the different continents. The result was significant ($\chi^2(6) = 18.98$, $p = 0.004$), confirming my observation. I have added these statistical results to the 'Results' section on lines 662-664. I believe this strengthens my conclusion about the varying research approaches in different regions.

- Consider using a widely accepted method for quality assessment as your assessment tool does not assess some key components including - study design, bias assessment, etc.
    - Understood. This was also raised by Reviewer 1.
    - The goal of my scoring system was not to assess general methodological quality, but rather to specifically evaluate each study's relevance to my unique research question and to the utility as a clinical biomarker for transplant medicine. To clarify this distinction, I have revised the 'Study Scoring and Quality Assessment' section on lines 411-412 to explicitly state that the criteria were designed to assess "direct relevance... rather than its general methodological rigor." I believe this clarification more accurately frames the purpose of my scoring system.

- You make this statement – “Despite this, there appears to be a large methodological gap in the literature regarding timing of rejection.” The data on this outcome should be reported with the other results and the methods should be updated to include this outcome
    - Acknowledged. I have made two key revisions to my manuscript:
        - First, I updated the Introduction on lines 110-112 to include a fourth objective: to formally assess the methodological approaches used in the literature for evaluating rejection timing.
        - Second, I moved the relevant paragraph (from lines 706-712) into its own new subsection in the 'Results' section titled "Methodological Gaps in Evaluating Rejection Timing," which you can now find on lines 906-912.
    - I believe these changes make my analysis more rigorous and my manuscript's structure more logical.

- Several results are stated multiple times (for example, the racial/ethnic distribution of the data) – recommend reviewing this section for redundancy
    - Thank you. I have now reviewed this section and removed the redundant statements to ensure the point is made clearly only once. Specifically, I have removed the repetitive content on lines 606-612 (the "Performance in Multiple Ancestries" subsection), lines 668-670, and lines 672-675. I believe this significantly improves the flow of the analysis.

- The description of the AUC and ROC belongs in the methods section as opposed to the results section.
    - Acknowledged. I have relocated the descriptive paragraph defining the Area Under the Receiver Operating Characteristic curve from the 'Results' section (originally lines 736-744) to the 'Data Analysis' subsection within the Methods

(now on lines 448-463). This change ensures all methodological details are presented together before the results are introduced.

- Please review this claim – “An interesting finding is that studies do not report mediocre performance and if this is potentially influenced by publication bias of studies with negative findings, it may at least suggest that when dd-cfDNA is investigated, it is consistently a good and effective biomarker for detection of allograft rejection. “ If you believe that negative results are being underreported, it actually suggests that dd-cfDNA may be overestimated in its ability to detect rejection if negative results are not being reported
    - Absolutely. I have rewritten the original sentence on lines 753-756. The new, corrected sentence on lines 756-759 now states that the lack of reported mediocre results, if due to publication bias, would suggest that the true performance of dd-cfDNA is likely overestimated in the current literature. Thank you for giving me the opportunity to enhance my reasoning.

- AUC results are discussed twice in this section without new information being added - consider consolidating.
    - Thank you. I have now removed the second, redundant paragraph on lines 761-772 entirely. This leaves the more comprehensive first paragraph to present all the findings related to the AUC analysis in a more streamlined and logical manner.

- Be careful stating that certain tests are more useful in certain racial/ethnic populations as race may be a proxy for social factors. Additionally, race-based medicine has implications for patient care that can lead to disproportionate access to resources (such as transplant) as occurred with eGFR
    - Additionally, the confidence intervals in all groups appear to overlap so the differences appear statistically insignificant. Again, please describe the statistical testing used here.
    - Absolutely and I appreciate your feedback.
    - I have made significant revisions to this section:
        - First, I revised my language on lines 796-798 (now on lines 798-801) to be more careful, clarifying that race is a social construct and that observed differences in outcomes for certain groups are often due to social and clinical factors.
        - Second, as you suggested, I performed an analysis of variance (ANOVA) to compare the AUC values across the different groups. As you suspected from the overlapping confidence intervals, the test confirmed that there is no statistically significant difference in biomarker performance. I have replaced my original conclusion on lines 830-833

with a new paragraph on lines 833-842 that states this result, clarifying that dd-cfDNA performs consistently across these populations.
    - I believe these changes make my analysis both more statistically rigorous and more responsible in its discussion of patient ancestry.

- This section is the first time you discuss analysis of different types of rejection (AMBR vs. TCMR) – this concept and its significance should be discussed in the background section and the methods of comparison should be discussed above as well.
    - Acknowledged. I have now added a new paragraph on lines 214-223 in the Background section that defines both T-cell mediated rejection (TCMR) and antibody-mediated rejection (ABMR). I believe this addition provides the necessary context for the reader before they get to the analysis.

Limitations

- You should discuss the underreporting of negative studies
    - Absolutely. Reviewer 1 was also recommended similarly. Therefore, I have added a sentence on this topic to the 'Results' section where I discuss publication bias in the context of the AUC analysis on lines 756-759. I believe this placement integrates the limitation more effectively into the main findings.

- You should discuss the translability and applicability of this test in clinical practice
    - Absolutely. Reviewer 1 also suggested similarliy. I have added a sentence to the final paragraph of my 'Results' section, which discusses future directions. This new sentences on lines 991-1000 explicitly states the need for future research into the cost-effectiveness and logistical challenges to ensure the clinical translatability and applicability of dd-cfDNA.

- Were there limitations in the study design besides the generalizability. As stated above, please include characteristics of the studies included and the pros and cons of those studies should be discussed here
    - Thank you. I have added a new paragraph on lines 906-912 at the end of my 'Results' section. This new paragraph discusses the limitations of the included study designs, specifically addressing the pros and cons of single-center vs. multicenter and retrospective vs. prospective studies, and how these factors impact the overall evidence base.

Future Directions

- You discuss the difference in accuracy by different commercial assays, but this is not discussed elsewhere in the review. This should be discussed as it may be a major limitation.
    - Figure 8d is referenced here but it is not present in the analysis section.
        - Thank you for your feedback. To fix this, I have now added a sentence to the limitations paragraph in my 'Results' section on lines 906-912. This sentence explicitly states that the variability between different commercial assays is a key limitation of the current literature. This provides the necessary context for my later recommendation for head-to-head comparison trials.
        - Regarding Figure 8d, my apologies for the confusion. Due to renumbering, that figure is now Figure 10d and is correctly placed in the 'Results' section where it is referenced.

- **You state that more research is necessary to determine the frequency of monitoring – information on how the studies monitored using this metric is imperative to being able to compare them. This should be discussed in the literature overview.**
    - Thank you and acknowledged. I have added a new subsection to my 'Results' section on lines 599-604 titled "Monitoring Frequency in Included Studies." This new section briefly summarizes the wide variety of testing schedules reported in the literature, which provides a much better context for my later recommendation that standardized monitoring protocols are needed.

- Stating that dd-cfDNA is more accurate than creatinine seems to be an overstatement as no direct comparison between the two methods is discussed in the analysis section; you only state the diagnostic accuracy of dd-cfDNA only. As this is one of the stated outcomes of this study, I would discuss how dd-cfDNA compares to creatinine, how they were compared in the original study, and how you determined that dd-cfDNA > creatinine.
    - Thank you. Absolutely. I have now added a new subsection to my 'Results' section titled "Comparison to Serum Creatinine." This new section on lines 563-572 synthesizes the findings from the studies that directly compared the two biomarkers, presenting the typical AUC values reported for both. This provides the quantitative evidence to support the conclusion that dd-cfDNA is a more accurate biomarker for detecting allograft rejection.

Submission title: Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection

Date of Review: 9.25.2025

Decision: Accept with minor revisions (acceptance conditional on satisfactory minor revisions)

Comments:

This is an excellent revision that demonstrates careful integration of previous comments and suggestions. The manuscript is now much more polished and cohesive, with clearer organization and effective presentation of findings. Table 1 and the PRISMA flow diagram are particularly well-executed, they are visually clear, easy to follow, and contribute significantly to the transparency of the review process. The identification and synthesis of 83 studies represent a substantial undertaking, and the review provides a robust, well-structured, and insightful examination of the available data. The discussion is thoughtful, and the argument for dd-cfDNA as a promising biomarker is well-supported. Overall, this is a strong piece of work that makes a valuable contribution to the field.

Revision requests:

- The definition of AUC (Area Under the Curve) appears multiple times throughout the manuscript, in the Data Analysis section, again in the Comparison to Serum Creatinine section, and once more in the Conclusion. The abbreviation only needs to be defined once when it is first introduced. After that, the manuscript can simply use the abbreviation consistently.
- A minor stylistic point: the manuscript currently uses first-person phrasing throughout (for example, “I analyzed,” and “I defined”). For a systematic review, try to use impersonal or passive phrasing that emphasizes the process or findings rather than the author. Consider rephrasing sentences to “The analysis shows…” or “X was defined as…”

Non-Invasive Surveillance in Kidney Transplantation: A Systematic Review of Donor-Derived Cell-Free DNA as a Biomarker for Transplant Rejection

[Author name redacted by Managing Editor]

September 30, 2025

Abstract

The ability to monitor kidney allografts is limited by traditional forms of assessment like invasive biopsy, and also delayed biomarkers, such as serum creatinine (Rizvi et al., 2023a). This systematic review was performed to assess the clinical usefulness of donor-derived cell-free DNA (dd-cfDNA) as a non-invasive biomarker of allograft injury. A systematic search of the Google Scholar and PubMed databases yielded 1,153 records; following a manual screening of titles and abstracts, 83 studies were selected for inclusion. Results indicate a rapidly maturing evidence base, including an increase in publications and a corresponding increase in sample size over time, reflected most notably in a landmark study on nearly 3,000 patients (Aubert et al., n.d.). The literature generally met a set of curated criteria for non-invasive biomarker studies derived by the author. For example, most studies obtained 3 or more out of 5 criteria: “Measured Blood?”, “Analyzes rejection probability?”, “Temporal Precedence Explicitly Stated”, “Longitudinal Monitoring for Timing”, and “Compares to Traditional Biomarkers?”. A key finding of the review was the consistent and generally favorable diagnostic accuracy of dd-cfDNA, with most studies indicating an Area Under the Curve (AUC) greater than 0.80 (Mantios et al., 2023). There is, however, a stark geographical disparity in the inclusion of diversified ancestry; the available evidence for Black and Hispanic ancestry is largely based on North American evidence (Bu et al., 2022). This systematic review provides confirmatory evidence that dd-cfDNA is a strong diagnostic biomarker (Kim et al., 2024a); however, it did not assess clinical feasibility. Future work must address this gap by focusing on cost-effectiveness and implementation studies, in addition to conducting large, multicenter trials in diverse populations to ensure its equitable and optimal application (Botella et al., 2024a).

Introduction

Kidney transplantation is a vital therapy for patients with end-stage kidney disease, providing a better overall longevity and improved quality of life compared to long-term dialysis (Dreige et al., 2022), which is the non-surgical alternative to transplantation. Despite its advantages and evidence for superior long-term outcomes, transplantation faces the ongoing challenge of immune-mediated rejection, which remains a prominent and often principal cause of patient graft-loss (Lakhani et al., 2021). The challenge is exacerbated by a poor risk management framework that relies on conventional surveillance practices that typically are complicated and inaccurate to true quantitative values (Wei et al., 2024). This framework is limited by its dependence on invasive and reactive kidney biopsies and the use of non-specific, lagging biomarkers like serum creatinine that often only rise after irreversible graft damage has occurred (Parajuli et al., 2024a; Rizvi et al., 2023a). Traditionally, markers for the identification and management of rejection have been insufficient due to the slow response of these traditional markers to tissue injury (Rizvi et al., 2023a). In short, these significant barriers represent the unmet medical need in transplant medicine, as they do not allow clinicians to intervene prior to true tissue/organ damage (Nie et al., 2025).

In recent years, a new solution has been developed to solve these deficiencies: the measurement of dd-cfDNA (Zhang et al., 2020a), small pieces of DNA from the donor that are released from the transplanted organ into the circulation, especially when the organ is activated, or inflamed, or injured through histocompatibility, or rejection (Zhang et al., 2020b). When a person has an organ transplant, they also are a recipient of a donor's genetic information at the time of the transplant (Kumar et al., 2023). Quantifying the amount of dd-cfDNA in a blood sample can now provide a direct, non-invasive and organ-specific signal of graft injury (Mayer et al., 2021). This has been depicted as a pioneering biomarker with much potential to detect graft damage early (Botella et al., 2024a). The clinical utility of dd-cfDNA stems from its short half-life of about 30 minutes, which allows it to serve as a near real-time indicator of acute allograft injury. When the transplanted kidney is healthy, dd-cfDNA is released at a low, baseline level. However, during an active rejection event, inflammation and cell death in the allograft cause a significant and measurable increase in the amount of dd-cfDNA released into the recipient's bloodstream. By quantifying this increase through techniques like next-generation sequencing (NGS), clinicians can detect rejection far earlier than with traditional

markers, creating an opportunity for pre-emptive treatment before the onset of widespread tissue damage.

Though transplant surgeons are currently utilizing dd-cfDNA in clinical practice, the development of the literature is so fast-paced that a comprehensive review of the evidence is needed (Chen et al., 2022). It is no longer sufficient to say that dd-cfDNA is "effective" and the key question to focus on is how effective it is in specific clinical contexts. The high diagnostic accuracy of dd-cfDNA is consistent with previous literature reviews and meta-analyses, which have affirmed the biomarker's potential for non-invasive monitoring. Previous meta-analyses have reported high pooled Area Under the Curve (AUC) values, such as 0.84 for acute rejection, 0.86 for overall rejection, and 0.88 for antibody-mediated rejection. This review builds on that consensus by providing a quantitative synthesis of the field's maturation, documenting the significant increase in study sample sizes and the shift towards more rigorous prospective cohort designs. It also addresses a critical knowledge gap by quantifying the significant geographic and ancestral disparities, highlighting that evidence for Black and Hispanic populations is almost exclusively derived from North American studies. However, a limitation of these existing reviews is that they confirm the biomarker's general utility without systematically analyzing the characteristics of the rapidly evolving evidence base or quantifying the critical knowledge gaps related to its performance in diverse patient populations.

The unique value of this systematic review is that, in addition to synthesizing diagnostic accuracy, it provides the first quantitative analysis of the field's methodological maturation and its significant demographic and geographic gaps (Sharma et al., 2022); it is guided by the specific hypothesis that donor-derived cell-free DNA (dd-cfDNA) can result in increased diagnostic accuracy in detecting transplant rejection and graft injury versus current standard of care biomarkers, through earlier and non-invasive methods. The focus of this review is to (1) specify and quantify overall diagnostic accuracy of dd-cfDNA for various types of rejection, (2) create and analyze data visualizations to synthesize the characteristics and findings of the included studies, (3) identify trends in performance and address key knowledge gaps, such as the biomarker's utility in diverse, underrepresented patient populations (Benning et al., 2022), and (4) assess the methodological approaches used in the literature to evaluate the timing of rejection.

Background

Kidney Transplantation and Rejection

Kidney transplantation is the most common organ transplantation because its underlying disease, end-stage renal disease (ESRD), is also one of the most common diseases, affecting >800000 persons in the United States (*Kidney Disease Statistics for the United States - NIDDK*, n.d.). The disease carries a significant burden, with a mortality rate of nearly 25% for patients on dialysis, highlighting the critical need for transplantation as a life-saving intervention. Conditions like diabetes and high blood pressure are the leading causes of kidney failure, impacting millions and contributing to the high demand for transplants (*American Kidney Fund*, 2024).

When an organ is transplanted, it is important to ensure the recipient's immune system will accept it and not recognize it as foreign; otherwise, an organ can be rejected (Nguyen et al., 2025). The most common immune responses are described as acute rejection, which happens shortly after the transplant has taken place, and chronic rejection, which occurs over time (Rizvi et al., 2023b). Ultimately, failure associated with acute and chronic rejection can be caused through progressive permanent damage in the transplanted organ with long-term changes in allograft function (Wei et al., 2024). Graft failure is a severe outcome, leading to a return to dialysis where the 5-year survival rate is only around 42%, underscoring the importance of effective allograft monitoring. Allograft rejection is primarily categorized into two distinct pathways: T-cell mediated rejection (TCMR) and antibody-mediated rejection (ABMR). TCMR is a cellular response where the recipient's T-cells directly attack the transplanted organ. ABMR is a humoral response driven by donor-specific antibodies (DSAs) that target the graft, leading to inflammation and tissue damage. Distinguishing between these subtypes is critical for guiding appropriate treatment, as they respond to different immunosuppressive therapies.

Currently, the monitoring of kidney transplant physiology is performed using a combination of clinical and biochemical monitoring protocols (Oellerich et al., 2021). Clinicians often take note of potential clinical signs or symptoms related to possible kidney dysfunction, such as decreased urinary output or developing edema in the legs and ankles from fluid overload (Kim et al., 2024a). However, to truly assess graft dysfunction, clinicians must perform invasive procedures to obtain

tissue or organ biopsies, which allows for definitive diagnosis of the pertinent histopathological process (Moein et al., 2024).

The current monitoring modalities have multiple problems that can greatly influence patient care (Rizvi et al., 2023b). First, serum creatinine is a lagging and suboptimal marker; in fact, serum creatinine will only begin to elevate when potentially irreversible injury to the donated kidney has already occurred (Rizvi et al., 2023a).

Both the invasiveness of biopsies and the lagging sensitivity of serum creatinine are limitations in post-transplant management (Aubert et al., n.d.). There is a clear need for a noninvasive and improved biomarker for the kidney transplant rejection technique that can provide an early and accurate evaluation of graft health (Akalin et al., 2021).

Cell-free DNA use in transplantation

cfDNA, which consists of the short fragments of DNA that are free-circulating in the bloodstream, shows promise as a biomarker in transplantation research (Zhang et al., 2020b). Cell-free DNA originates when cells in the human body experience apoptosis, programmed cell death or necrosis, uncontrolled cell death (Oellerich et al., 2021). When cells die and degrade, the genetic information in the cells is released into the biological fluid surrounding the cells, including blood plasma, urine, and cerebrospinal fluid (Nguyen et al., 2025).

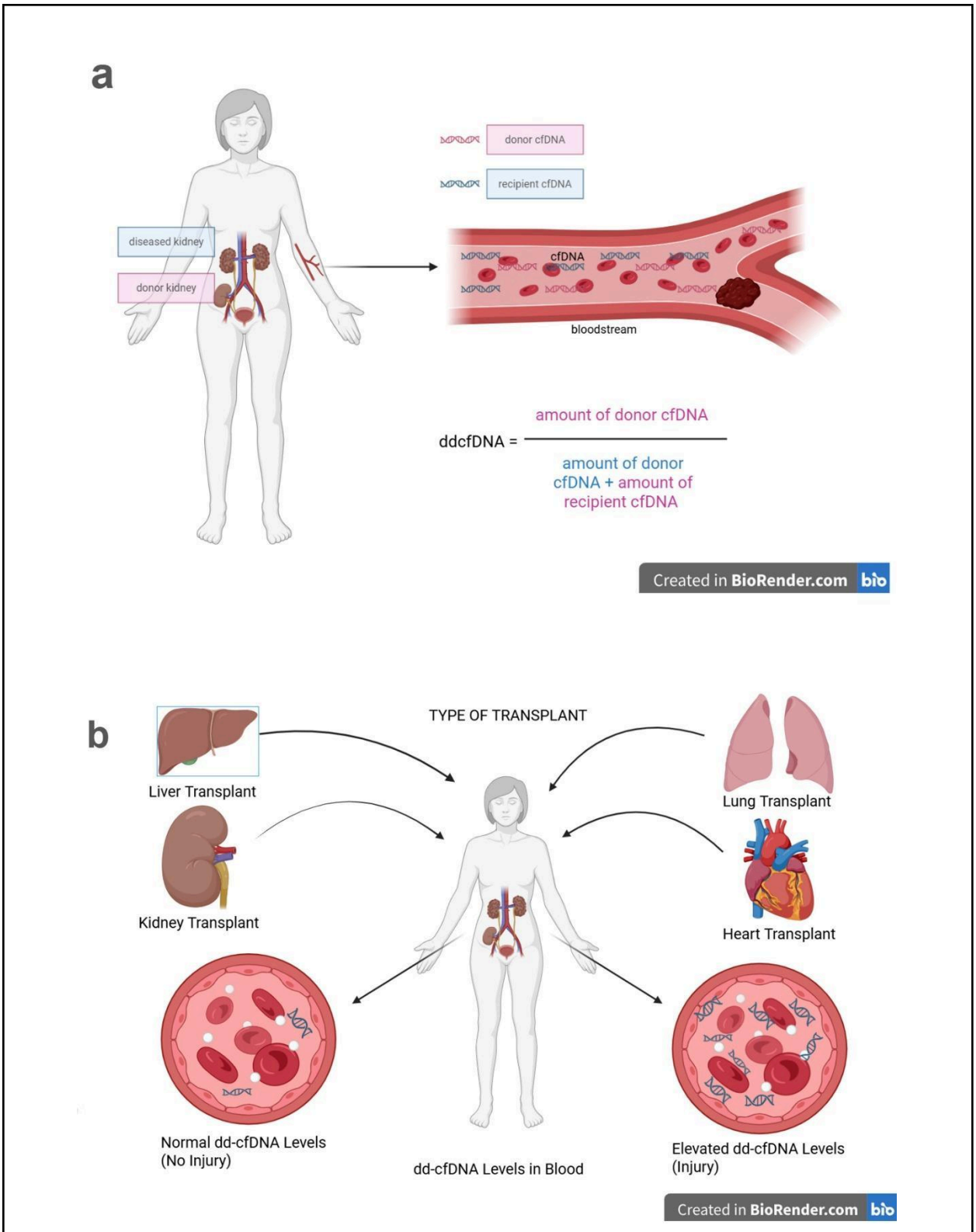


Figure 1: The Biological Principle of Donor-Derived Cell-Free DNA (dd-cfDNA)

as a Biomarker. (a) This diagram illustrates that following a kidney transplant, both recipient and donor cfDNA circulate in the bloodstream. The fractional abundance of dd-cfDNA is calculated as the proportion of donor cfDNA relative to the total amount of cfDNA. This figure was created by the author using BioRender.com. (b) This diagram shows that dd-cfDNA is a general biomarker for all solid organ transplants, including liver, kidney, lung, and heart. In a healthy, uninjured state, dd-cfDNA levels are low; however, during an injury event like rejection, damaged cells from the transplanted organ release a higher amount of dd-cfDNA into the blood. This figure was created by the author using BioRender.com.

cfDNA comes from blood plasma, the fluid part of blood (Wolf-Doty et al., 2021). Several molecular methods are used currently to assess cfDNA quantity (Halloran et al., 2022). Both quantitative polymerase chain reaction (qPCR) and next-generation sequencing (NGS) can be viewed as modern methods for measuring cfDNA (Akalin et al., 2021). qPCR does work similarly to a targeted measure, where specific primers are used to locate and amplify a pre-established and a small number of unique donor sequences (Nie et al., 2025). For NGS, millions of cfDNA fragments can be read concurrently across the genome (Kumar et al., 2023b). In this way, cfDNA can be assessed to identify the donor if they donated an organ (Wei et al., 2024).

Figure 1a provides biological context for utilizing dd-cfDNA as a biomarker. After a solid organ transplant, like kidney, the recipient's blood will contain their own cfDNA in addition to a small amount of cfDNA from the donor organ (Fig. 1a) (Oellerich et al., 2021). Under normal circumstances and with a healthy transplant, the amount of dd-cfDNA circulating in the blood is in a baseline state of low concentrations (Bu et al., 2022a). However, when an injury event causes a greater rate of cell death in the allograft, such as rejection, inflammation, or other forms of graft stress, the allograft will release dd-cfDNA into the recipient's circulation (Fig. 1b) (Rizvi et al., 2023b). These real-time changes in dd-cfDNA concentrations can reflect ongoing allograft injury (Nguyen et al., 2025).

cfDNA is a dynamic biomarker. It is constantly shed into the bloodstream, and the half-life of a single cfDNA molecule is approximately 30 minutes in the bloodstream (Wolf-Doty et al., 2021), meaning that it provides real-time information about tissue degeneration. Changes in specific cfDNA characteristics, most notably an

extraordinary increase in levels of cfDNA, are indications of tissue injury, inflammation, or subsequently, a disease such as rejection, which validate the value of cfDNA as a non-invasive biomarker (Rizvi et al., 2023). This indicates that cfDNA is a valid source of non-invasive biomarker (Wei et al., 2024). The drive and excitement around a real-time multi-level signal for more precise management is the appeal of this biomarker that affirms its usefulness (Aubert et al., 2024). dd-cfDNA has become such a strong force and continues to be a focus of transplant research (Halloran et al., 2022).

The transplantation field is actively looking for new, non-invasive, and accurate ways to monitor graft health, given the many challenges with established surveillance modalities (Rizvi et al., 2023b). The gold standard now, kidney biopsy, is invasive, with procedural and other risk and discomfort for the patient. Serum creatinine, the most widely used non-invasive biomarker, is a surrogate, vague and non-specific biomarker. CfDNA refers to all small fragments of DNA found circulating in the bloodstream. At a healthy baseline, all individuals will have a detectable level of cfDNA originating from normal cell death. However, after a transplant, a fraction of this total cfDNA originates from the donated organ; this specific component is known as donor-derived cell-free DNA (dd-cfDNA). While elevated total cfDNA indicates cell death somewhere in the body, the donor-derived fraction (dd-cfDNA) is a more specific biomarker for allograft health because its origin is exclusively the transplanted organ, thereby providing a clearer view of its health. Crucially, this signal reflects the real-time process of cell death as it happens, allowing for the detection of low-level or early-stage allograft injury, whereas conventional markers like serum creatinine only rise after substantial, cumulative, and often irreversible functional damage has already occurred. While biopsy serves as the definitive diagnostic tool, its invasiveness and the lagging nature of conventional biomarkers make them unsuitable for the frequent, proactive monitoring needed to prevent irreversible graft damage.

Recent advancements have uncovered the study of donor-derived cell-free DNA (dd-cfDNA) as an alternative solution that directly counters these flaws by offering a non-invasive and accelerated method for assessing transplant status, potentially transforming post-transplant care (Botella et al., 2024a). The following results will synthesize the evidence to quantify the biomarker's accuracy and its timelines for early detection, thereby assessing its potential to become a new standard in

post-transplant surveillance (Bu et al., 2022). The clinical utility of dd-cfDNA stems from its short half-life of about 30 minutes, which allows it to serve as a near real-time indicator of acute allograft injury. When the transplanted kidney is healthy, dd-cfDNA is released at a low, baseline level. However, during an active rejection event, inflammation, and cell death in the allograft cause a significant and measurable increase in the amount of dd-cfDNA released into the recipient's bloodstream. By quantifying this increase through techniques like next-generation sequencing (NGS), clinicians can detect rejection far earlier than with traditional markers, creating an opportunity for pre-emptive treatment before the onset of widespread tissue damage.

Methods

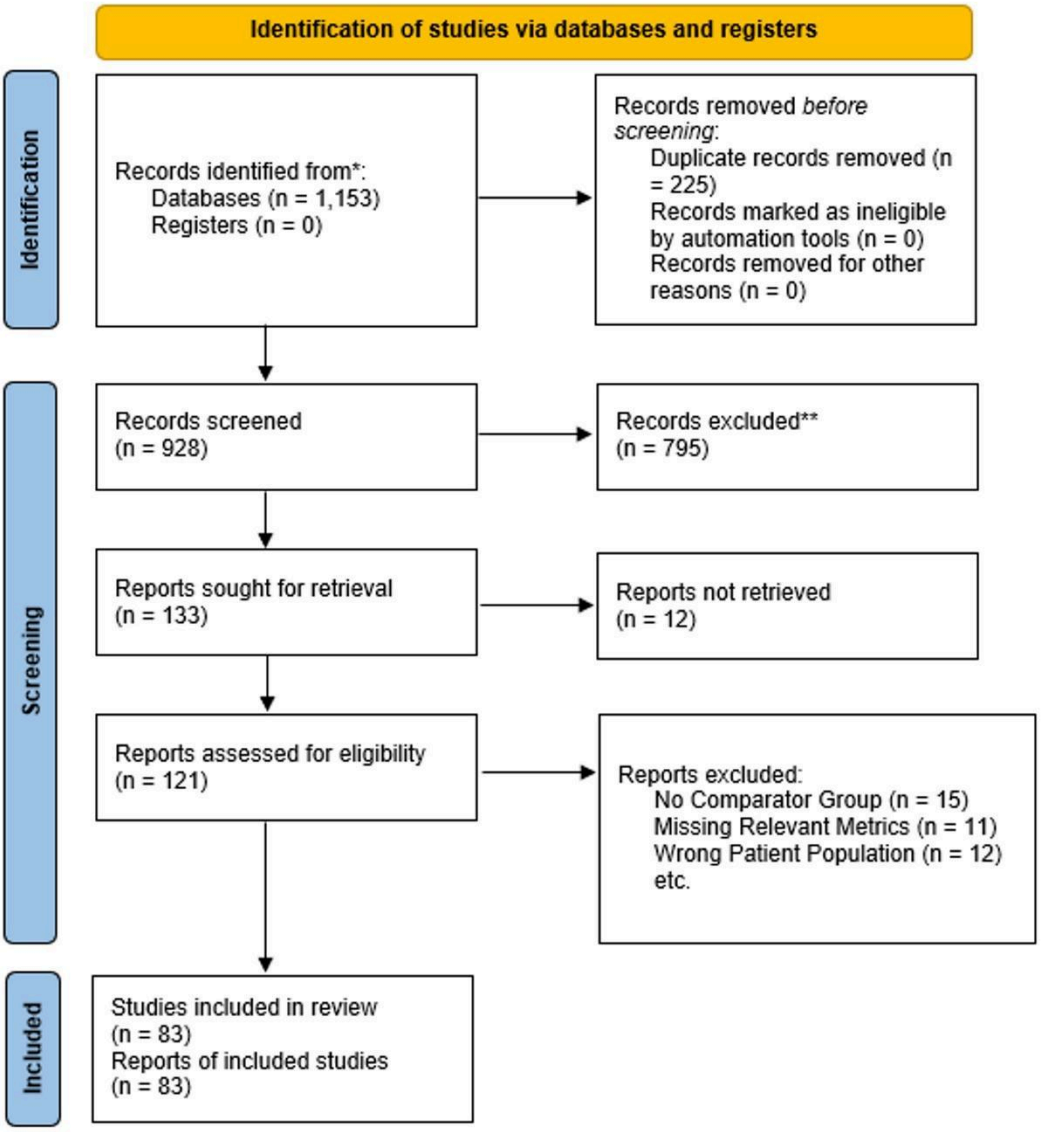


Figure 2. PRISMA 2020 Flow Diagram for Study Selection: This diagram illustrates the flow of information through the different phases of the systematic review, mapping out the number of records identified, included, and excluded.

Search Strategy

A comprehensive literature search was conducted using the Publish or Perish application (*Publish or Perish*, n.d.) to identify relevant articles. Searches were performed using the Google Scholar and PubMed databases. The search focused on

publications related to donor-derived cell-free DNA (dd-cfDNA) and transplantation, utilizing the following keywords: "dd-cfDNA", "donor-derived cell-free DNA", "cell-free DNA renal transplant", "allograft surveillance biomarkers", "kidney transplant rejection", "biomarker acute rejection", "Kidney Transplantation" (MeSH), "Graft Rejection" (MeSH), "Cell-Free Nucleic Acids" (MeSH), "Biomarkers" (MeSH), "Sensitivity and Specificity" (MeSH), "Comparative Effectiveness Research" (MeSH), "Diagnostic Imaging" (MeSH), "Creatinine/blood" (MeSH), "Monitoring, Immunologic" (MeSH), "diagnostic tool", "diagnostic performance", "surveillance method", "human subjects", "adult", "renal", "renal transplant", "kidney allograft", "diagnosis", "diagnostic", "AUC", "predictive value", "rejection", "sensitivity", "specificity", "performance", "comparison", "receiver operating characteristic curve", "serum creatinine", "biopsy", "imaging", "donor-specific antibodies", "DSA", "Racial Groups" (MeSH), "Ethnic Groups" (MeSH), "rejection monitoring", "allograft", "kidney".

These keywords were categorized to capture key aspects of the research question :

- **Population:** Keywords of "adult," "human subjects," "kidney transplant," "renal transplant," and "Kidney Transplantation" (MeSH) (including MeSH terms) were used to specify the patient group. "Racial Groups" (MeSH) and "Ethnic Groups" (MeSH) were included to address the "diverse ancestries" element.
- **Intervention:** Phrases such as "dd-cfDNA," "donor-derived cell-free DNA," "cell-free DNA renal transplant," and "Cell-Free Nucleic Acids" (MeSH) specifically targeted the primary diagnostic tool of interest.
- **Outcome:** Terms like "acute rejection," "graft rejection," "allograft rejection," and "Graft Rejection" (MeSH) directly identified the clinical event under investigation.
- **Comparison:** To encompass "standard biochemical and imaging surveillance," keywords like "serum creatinine," "biopsy," "imaging," "donor-specific antibodies," and related MeSH terms were utilized.
- **Diagnostic Role/Timing:** Phrases such as "diagnostic tool," "diagnostic performance," "surveillance method," "sensitivity," "specificity," "AUC," and "predictive value" were crucial for identifying studies assessing the diagnostic utility, accuracy, and particularly the comparative timing of dd-cfDNA.

The search process was restricted to articles published between January 1st, 2020, and May 31st, 2025.

Inclusion Criteria

Inclusion criteria were designed to bring in articles directly aligned with the research question:

English-language studies; Human subjects, specifically adult kidney transplant recipients (age ≥18 years); Use of donor-derived cell-free DNA (dd-cfDNA) as a primary or comparative diagnostic tool for acute rejection; Includes comparison to at least one standard surveillance method (e.g., serum creatinine, biopsy, imaging, donor-specific antibodies); Studies assessing the diagnostic role of dd-cfDNA in acute rejection, ideally reporting metrics (e.g., sensitivity, AUC) or qualitatively describing performance; Studies involving patients from diverse racial/ethnic backgrounds (where available); Studies published between January 1st, 2000 and May 31st, 2025; and original research articles only (literature reviews, systematic reviews, and meta-analyses were excluded).

Articles not meeting the inclusion criteria were excluded.

Selection Process

The resulting set of articles was exported into a Google Sheets document, and the complete selection process is detailed in the PRISMA flow diagram (Figure 2). Following this, a preliminary screening of manual reading of abstracts and titles was conducted to decide whether to include or exclude. The screening included: Reasoning as to why the article was chosen, Full Text Accessibility (only for texts included), Number of Citations, Authors, Title of Article, Year Published, Source, Publisher, Article URL, Cites URL, GS Rank, Query Date, Type of Article, DOI, ISSN, Citation URL, Volume, Issue, Start Page, End Page, ECC, Cites Per Year, Cites Per Author, Author Count, Age, and Abstract.

First Author	Year	Journal	Sample Size	Study Design	AUC for Rejection
Aubert, O.	2024	*Nature Medicine*	2,987	Prospective Cohort	0.87 (ABMR)
Bu, L.	2022	*Kidney International*	1,227	Prospective Cohort	0.81 (Any Rejection)
Halloran, P.F.	2022	*Transplantation*	385	Prospective Cohort	0.93 (ABMR)
Bromberg, J.S.	2024	*Transplantation*	191	Prospective Cohort	0.84 (Subclinical AR)
Stites, E.	2020	*Am. Journal of Transplant.*	231	Retrospective Cohort	0.82 (TCMR ≥1A)
Kim, H.D.	2024	*Frontiers in Immunology*	206	Retrospective Cohort	0.91 (Clinical ABMR)
Parajuli, S.	2024	*Clinical Transplantation*	148	Retrospective Cohort	Not Reported
Benning, L.	2023	*Transplant International*	104	Prospective, Single-Center	0.78 (Any Rejection)
Zhang, H.	2020	*Frontiers in Immunology*	129	Prospective Observational	0.94 (ABMR)
Akifova, A.	2025	*Nephrology Dialysis Transplantation*	184	Randomized Trial	0.85 (ABMR)
Chen, X.T.	2022	*Clinical Chemistry*	102	Prospective Cohort	0.89 (Any Rejection)
Gisch, N.	2025	*Am. Journal of Transplant.*	45 studies	Systematic Review	N/A
Yang, S.	2024	*Bosnian Journal of Basic Medical Sciences*	21 studies	Meta-Analysis	0.88 (ABMR)
Shen, J.	2020	*Clinical Transplantation*	68	Prospective Cohort	0.83 (Acute Rejection)
Mayer, K.A.	2021	*Transplant International*	115	Prospective Cohort	0.81 (ABMR)
Overall Summary	2020-2025	(Multiple)	N=83, Mean=245 Median=148 Range=21-2,987	(Multiple)	N=45, Mean=0.83 Median=0.82 Range=0.58-0.94

Table 1. Characteristics of Selected Included Studies. This table summarizes key information for a representative sample of the 83 studies included in the systematic review. The complete dataset used for this review is available in a Google Sheets document, which can be found in the "Code and Data Availability" section.

Data Extraction and Synthesis

The initial data extraction included standard article metadata such as GS Rank, Query Date, DOI, ISSN, Citations, Age of publication, and Author Count. For included articles, specific variables relevant to the research question were gathered. Gemini (Google LLC, 2025) was used to assist in extracting these variables, including: Title, Journal, Publication Year, Country, Study (single or multicenter), Study Type, Sample Size, Average Age, Proportion Male, Patient Race/Ethnicity, Type of Rejection, Type of cf-DNA, and AUC.

Study Scoring and Relevance Assessment

A scoring system was curated from the literature and specifically tailored to this review's research goals to assess the direct relevance of each included article. Articles were awarded one point (1) if the following criteria were met, and zero points (0) if not. This included: Measured Blood? Analyzes Rejection Probability? Temporal Precedence Explicitly Stated? Longitudinal Monitoring for Timing? Compared to Traditional Biomarkers?

These scoring criteria were curated by the author to assess each study's direct relevance in answering the research question. "Measured Blood?" confirmed the appropriate sample for dd-cfDNA, as blood plasma is the standard medium for this analysis (Wolf-Doty et al., 2021). "Analyzes Rejection Probability?" ensured focus on the primary clinical outcome of rejection (Bu et al., 2022). "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for Timing?" were crucial for evaluating the "earlier identification" aspect of the hypothesis, a key feature of dd-cfDNA's utility (Bromberg et al., 2024). Finally, "Compared to Traditional Biomarkers?" verified that the study included the necessary comparative analysis against standard surveillance methods like serum creatinine to establish the biomarker's superior performance (Mantios et al., 2023). Data extracted from the included articles included study type and design, specifically identifying if the research was a prospective cohort, retrospective, or observational study, along with other standard article metadata.

Data Synthesis

Articles receiving the high scores (3-5 points) were selected for in-depth analysis.

Preliminary conclusions were written from these high-scoring articles after reading them entirely and compiled into a summary document with a brief summary and overview of each article for future data synthesis and reference.

Data Analysis

Data analysis involved a narrative synthesis of findings from high-scoring articles. Extracted data on diagnostic timing, performance metrics, such as AUC and sensitivity, and patient demographics (including ancestry) were noted for future reference and figures. No formal statistical meta-analysis was performed due to the heterogeneity of the study populations and assays. Instead, comparisons between papers were made through a narrative synthesis of their findings, supported by data visualizations (e.g., box plots, scatter plots) that compare reported performance metrics and trends across different study contexts, such as geography and patient demographics.

The primary measure used to evaluate dd-cfDNA's diagnostic performance is the Area Under the Receiver Operating Characteristic (AUROC) curve. The AUROC is a valuable statistical measure that summarizes a test's ability to perform binary classification of patients with and without a condition across all possible thresholds; a value of 1.0 represents a perfect test, while a value of 0.5 represents a test with no discriminatory ability.

The Receiver Operating Characteristic (ROC) curve is a representation of a diagnostic test's overall performance, as described by one number, the Area Under the Curve (AUC). The AUC takes a value between 0.5 (no better than chance) and 1.0 (perfect test). Values between 0.80 and 0.90 may be considered "good". For the purposes of this systematic review, the AUC indicates the extent to which dd-cfDNA separates patients with biopsy-proven rejection (the cases) from patients without biopsy proven rejection (the controls); therefore, a high AUC indicates that the biomarker is quite accurate in identifying patients that are actively rejecting their organ.

Data Visualization

All figures presented in the Results section were generated using Google Colab, a cloud-based computational environment. Data extracted from the included studies

were compiled into a structured dataset and imported into the Colab environment for analysis. The Python programming language, along with the data manipulation library Pandas and the data visualization libraries Matplotlib and Seaborn, were used to create all plots, including histograms, box plots, and scatter plots, to visually synthesize the findings of this systematic review.

Results

In order to fulfill the objectives of this systematic review, the results were synthesized by firstly, describing the studies included, and then quantifying and gaining a general idea about the diagnostic accuracy of dd-cfDNA, and lastly, identifying some important performance trends and knowledge gaps.

Comparison to Serum Creatinine

A key objective of this review was to compare the diagnostic accuracy of dd-cfDNA to the current standard of care, serum creatinine. Across the studies that conducted a direct comparison, dd-cfDNA consistently demonstrated superior performance in detecting allograft rejection. The reported AUC for dd-cfDNA was consistently high, generally above 0.80, whereas the AUC for serum creatinine in the same cohorts was significantly lower, often ranging from 0.55 to 0.65. This indicates that serum creatinine has poor to fair discriminatory ability for rejection, while dd-cfDNA has good to excellent discriminatory ability, supporting the hypothesis that it is a more accurate biomarker.

Identification of Clinically Relevant Subtypes

Other publications were focused on building models to differentiate between AMBR and TCMR (Botella et al., 2024b). The key finding was that dd-cfDNA was a strong potential biomarker for ABMR in patients who are clinically stable and have normal renal function (as measured by serum creatinine) (Botella et al., 2024a). This is a huge advantage, since rejection at the subclinical level of ABMR can cause silent cumulative damage to the graft. Another published study described two patients who developed biopsy-proven ABMR within 60 days of a significantly elevated dd-cfDNA, despite both patients demonstrating normal renal function (Rizvi et al., 2023b).

Dd-cfDNA is particularly useful in certain clinical settings and high-risk

populations. The data suggests that dd-cfDNA is superior to creatinine for quantifying which patients with dnDSA (de novo donor-specific antibodies)—which are harmful immune proteins developed by the recipient that specifically attack the new organ—were definitively rejecting (Botella et al., 2024b) and needs to be considered for treatment decisions and guidance especially 15 days after transplantation (Botella et al., 2024a).

Early Detection with dd-cfDNA

Additional research focused on the ability of dd-cfDNA to detect graft injury significantly earlier than clinical presentation (Bromberg et al., 2024) or elevations in standard biochemical markers like serum creatinine and proteinuria, which are known to be lagging (Kim et al., 2024b). Dd-cfDNA's 30-minute half-life allows for the dynamic, consistent monitoring of ongoing damage and recovery (Wolf-Doty et al., 2021).

Monitoring Frequency in Included Studies

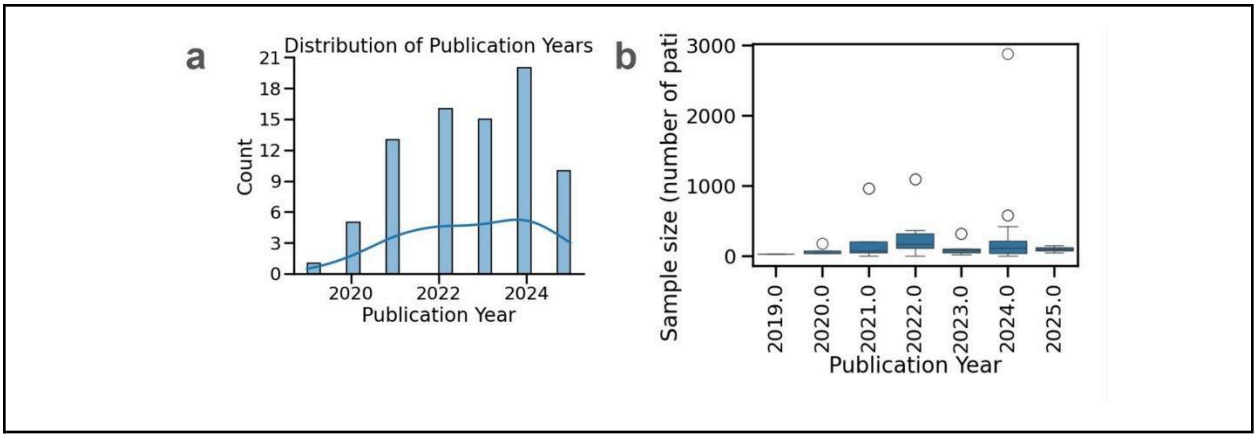
A review of the included studies revealed significant heterogeneity in the dd-cfDNA monitoring protocols used. The frequency of testing varied widely, from daily monitoring in the immediate post-transplant period in some studies, to weekly, monthly, and quarterly testing in others focused on long-term surveillance. This lack of a standardized approach makes it difficult to draw firm conclusions about the optimal monitoring schedule and highlights a key area for future research.

Characteristics of Included Studies and Publication Trends

First, to synthesize the characteristics and findings of the included studies (Goal 2), the compiled list of 83 studies that met the inclusion criteria was examined. Followed by examining the distribution of studies per publication year (Fig. 3a). It was found that publications have been steadily increasing over time, peaking in 2024 with approximately 22 studies. A linear regression analysis confirmed a significant positive trend between publication year and the number of studies (linear regression, slope = 3.5 studies/year, $p<0.05$), which suggests a rapid maturation of the evidence base. Furthermore, the relationship of how sample size changed with time in the field of kidney transplant dd-cfDNA research was interesting (Fig. 3b). In contrast to the number of publications, an analysis of variance (ANOVA) showed that while the median sample size appears to trend

upwards visually, this increase was not statistically significant ($p > 0.05$). This is likely due to high variability and the presence of several large-scale outliers, potentially impacted by the interruption of research during the Covid-19 pandemic. This indicates that while landmark studies are getting larger, the sample size for a typical study in the field has not significantly changed over this period. The evidence is predominantly composed of prospective cohort designs, with these studies forming the largest single category (Fig. 3e). These studies, which follow participants forward in time, minimize bias and provide a strong evidence base.

To address the third goal of identifying trends and knowledge gaps, particularly regarding the biomarker's utility in diverse, underrepresented patient populations, the geographic and demographic distribution of the studies was analyzed. An examination of participant characteristics confirmed that the study focuses on a middle-aged cohort, with an average age of 52 years (Fig. 3c). This is both a useful finding and a commonly observed finding in clinical practice since the average age of transplantation is about 50 years. The analysis also highlighted that study cohorts tended to be predominantly male with a prominent peak, showing a proportion of male participants up to a significant 0.63 (Fig. 3d). To gain a better understanding of the generalizability and nature of the evidence in a global context, a breakdown of the study types by geographic region was conducted. An analysis of the geographic distribution of study types revealed that North American research contributed the most retrospective and observational studies, whereas Europe and Asia contributed a higher proportion of prospective cohort designs (Fig. 3f). A chi-square test of independence confirmed that this difference in the distribution of study types across continents is statistically significant ($\chi 2(6) = 18.98$, $p = 0.004$).



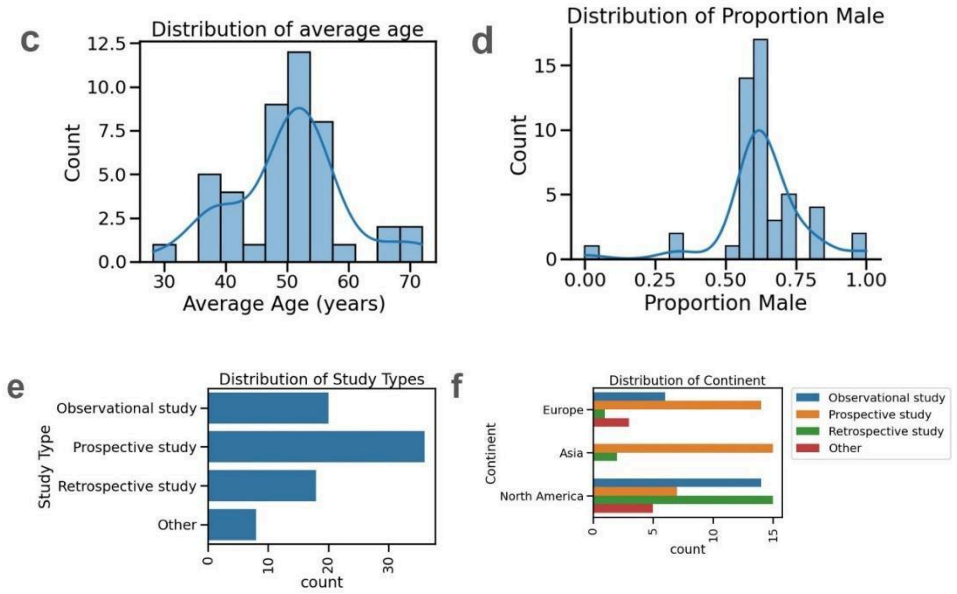


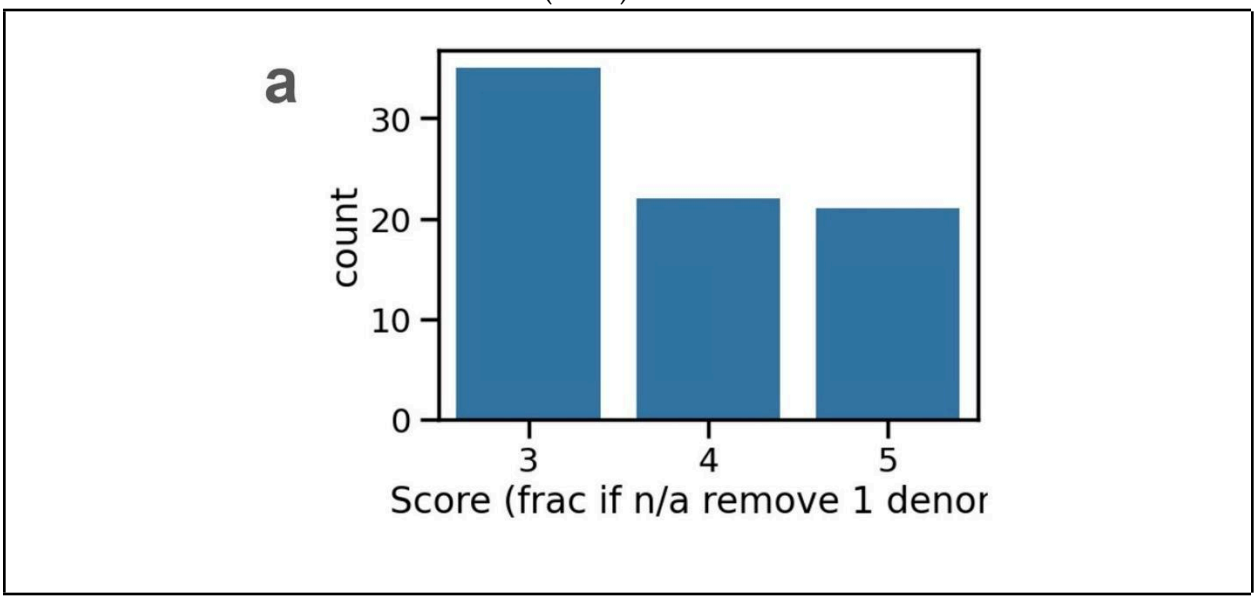
Figure 3: Publication Trends, Demographics, and Methodological Characteristics of Included Studies. (a) A histogram showing the number of included studies published per year, illustrating a rising trend in research volume peaking in 2024. (b) A box plot visualizing the distribution of study sample sizes by publication year, highlighting an increase in the scale of research and several large-scale outliers. (c) A histogram displaying the distribution of the average age of participants, showing a concentration in the middle-aged demographic. (d) A histogram of the proportion of male participants in study cohorts, indicating a predominance of male-majority studies. (e) A horizontal bar chart quantifying the included studies by their primary methodological design, with prospective studies being the most common. (f) A grouped horizontal bar chart illustrating the distribution of study types across the primary geographic regions of North America, Asia, and Europe.

Methodological Quality of the Evidence Base

To assess the methodological soundness of the literature used in the study, a 5-point quality scoring system was enacted. The 5-point quality scoring system was created to determine which sources most fit the hypothesis. Each study achieved one point for fulfilling each of the criteria: (1) Measured Blood?, (2) Analyzes probability of rejection?, (3) Temporal Precedence Explicitly Stated?, (4) longitudinal monitoring for time?, (5) compares to traditional biomarkers?. These scoring criteria were chosen to evaluate the relevance of each study to the specific research hypothesis. The criterion "Measured Blood?" evaluated whether the study

utilized the appropriate sample type to analyze dd-cfDNA (Sawinski et al., 2021a). The criterion "Analyzes rejection probability?" ensured that the study focused on the primary clinical outcome, i.e. rejection (Bu et al., 2022b). "Temporal Precedence Explicitly Stated?" and "Longitudinal Monitoring for Timing?" allowed for the evaluation of the hypothesis, specifically the earlier detection factor, which required identifying studies that monitored the biomarker over time (Parajuli et al., 2024b). Finally, "Compares to Traditional Biomarkers?" validated that the study conducted the comparative clinical analysis of the complete framework, thereby determining the "superior performance" of dd-cfDNA analysis compared to traditional clinical surveillance methods such as serum creatinine (Mantios et al., 2023). The distribution of studies per publication year (Fig. 4a) indicates a rapid maturation of the evidence base, with publications steadily increasing over time and peaking in 2024. Most studies reporting their scores are clustered in the upper range of the scale (e.g. 35 studies scored a 3; 22 studies scored a 4; and 21 studies scored a 5).

To add more specificity to the study characteristics that were contributing to high quality evidence, which criteria were most frequently met across the studies were assessed (Fig. 4b). The breakdown shows that foundational criteria were met across studies, particularly utilizing blood as a sample source (n=79) and comparing dd-cfDNA to traditional biomarkers (n≈78).



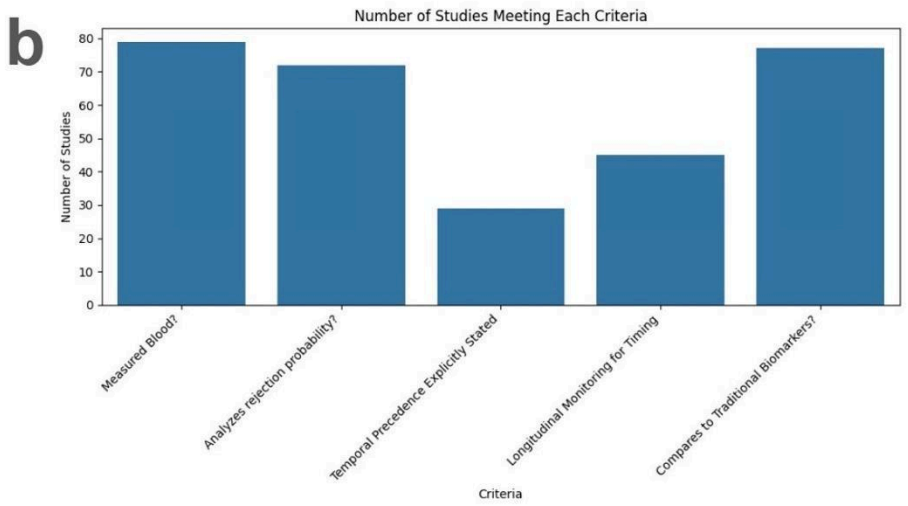


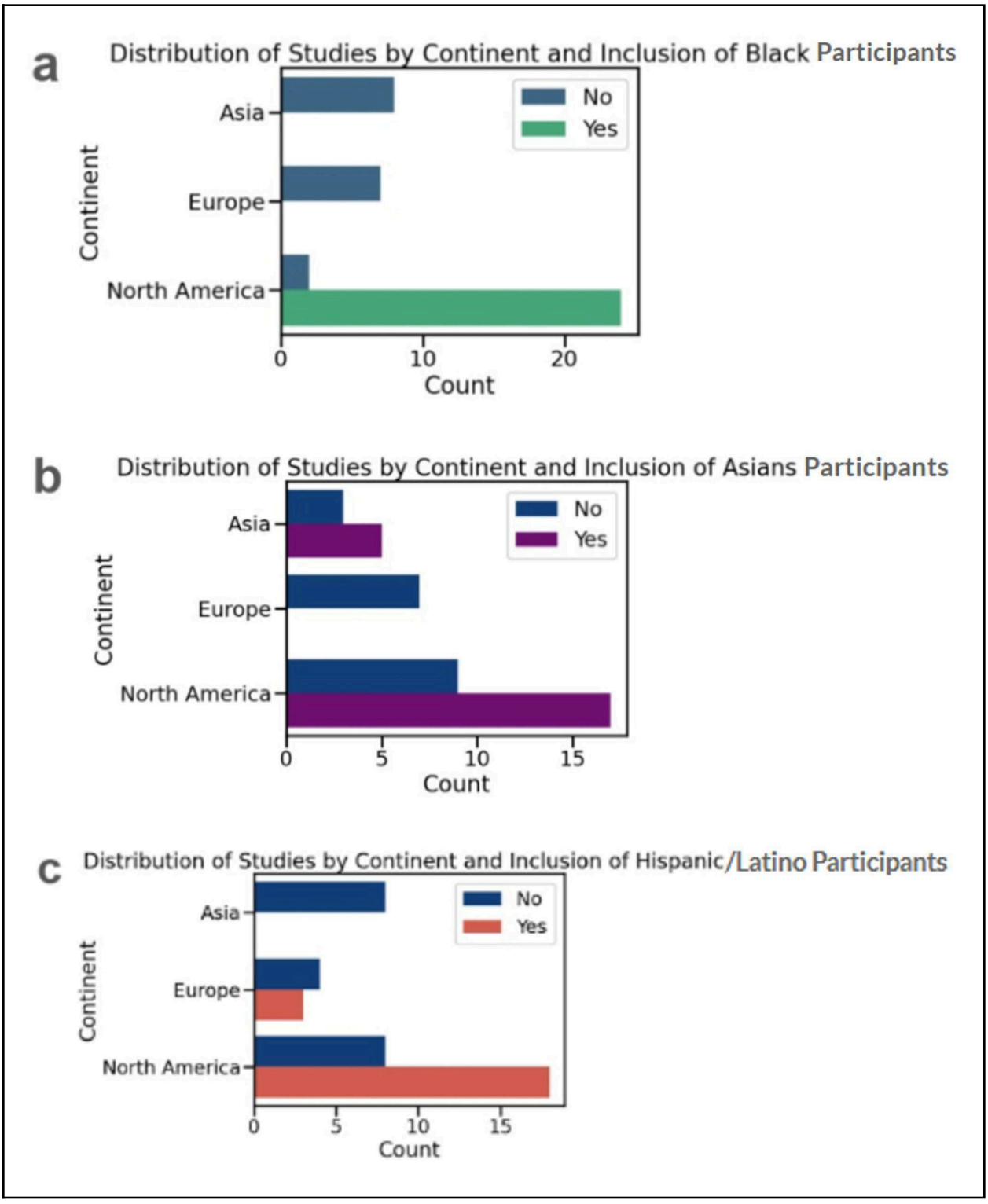
Figure 4: Distribution of Study Quality Scores and Fulfillment of Methodological Criteria. (a) A histogram showing the distribution of the overall 5-point quality scores across all included studies. (b) A bar chart quantifying the number of studies that met each of the five specific methodological quality criteria.

Geographic and Ancestral Distribution of the Evidence Base

In order to understand the efficacy of dd-cfDNA in multiple ancestries, which is an important goal for comprehensive utility of dd-cfDNA in kidney transplant settings, the inclusion of diverse racial and ethnic groups across the primary research continents was examined. To this end, the number of studies covering each self-reported race/ancestry group were examined, which were defined as “Asian,” “Black,” “White,” and “Hispanic/Latino.” The analysis reveals that North American research provides the most diverse evidence base. North American studies were the exclusive source of data for the inclusion of Black participants (n≈24) (Fig. 5a) and Hispanic/Latino participants (n≈18) (Fig. 5c). Furthermore, North American research also contributed the largest number of studies that include Asian participants (n≈17) (Fig. 5b) and White participants (n≈22) (Fig. 5d).

In contrast, the European studies included in this review, while numerous, did not explicitly report the inclusion of Black, Asian, or Hispanic participants, with many not reporting race altogether, though a small number did report including White patients (n≈3). The Asian studies, as expected, contributed evidence on Asian populations (n≈4) but did not report inclusion of other groups. This distribution

highlights a significant gap in the global evidence, suggesting that the generalizability of findings from Europe and Asia to diverse populations is limited.



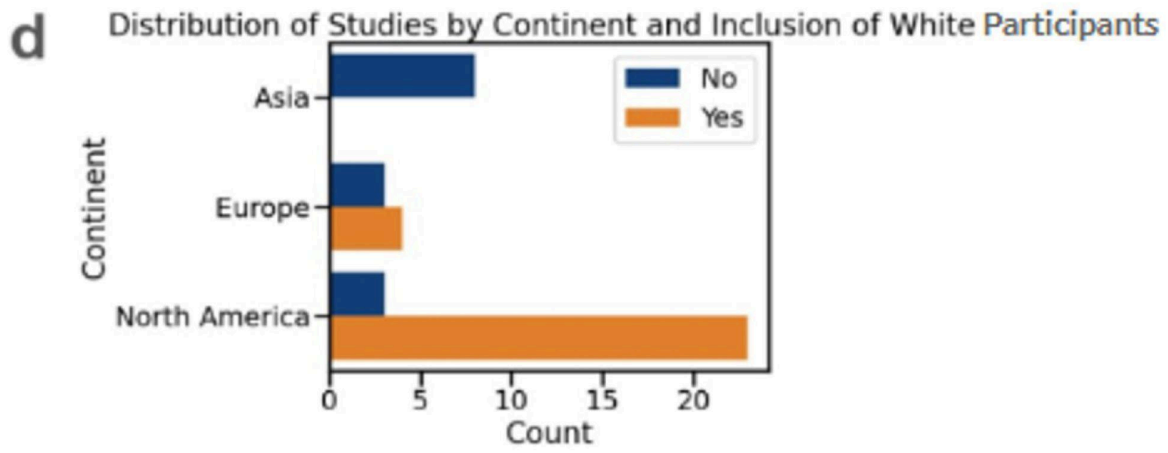


Figure 5: Geographic Distribution of Studies by Patient Ancestry. A series of horizontal bar charts showing the number of studies from North America, Europe, and Asia that did ("Yes") or did not ("No") report the inclusion of specific patient ancestries. Studies that did not report the race/ethnicity of their participants are not included. (a) Inclusion of Black participants. (b) Inclusion of Asian participants. (c) Inclusion of Hispanic/Latino participants. (d) Inclusion of White participants.

Diagnostic Accuracy of dd-cfDNA

A major study goal was to determine how effective dd-cfDNA was at identifying cases of kidney transplant rejection; therefore, AUC values were extracted from studies and analyzed. The diagnostic performance of dd-cfDNA, as measured by AUC, was found to be high across the included studies, as summarized in the histogram of Area Under the Curve (AUC) values (Fig. 6a). The data reveals a consensus towards models using dd-cfDNA having strong performance in the prediction of rejection, with a bimodal distribution showing two primary peaks where the evidence is concentrated. The largest cluster of studies reported an Area Under the Curve (AUC) with a peak centered approximately around AUC=0.80, while a second prominent peak centered even higher around 0.90. It is important to put this in perspective; for a standalone clinical diagnostic test, an AUC of 0.80 is generally considered poor as this indicates significant false positive and false negative rates that may cause indeterminate clinical decisions. An interesting finding is that studies rarely report mediocre performance; if this is influenced by publication bias where negative findings are underreported, it would suggest that the true diagnostic accuracy of dd-cfDNA may be overestimated in the current

literature.

A linear regression analysis was then performed to model the relationship of year of publication and AUC analyses in order to assess whether reported performance of the biomarker has improved as studies have accrued (a primary goal of this review). This does demonstrate an increasing association over time (slope ≈ +0.01 per year of AUC) (Fig. 6b). This positive slope suggests that more 'recent' studies, on average, have reported greater AUC than 'earlier' studies. This potential trend may in part reflect field maturation, perhaps including improved laboratory assays (i.e., could be absolute quantification vs. fractional abundance) and more sophisticated study design (years and capabilities are reflected in the increasing median sample sizes). The evidence for dd-cfDNA has not weakened over time; instead, it seems to have grown and solidified. An alternative explanation might include a growing publication pressure to report findings that best represent previous literature. This study's upward trend in AUC is accompanied by an upward trend in the sample sizes of studies over time.

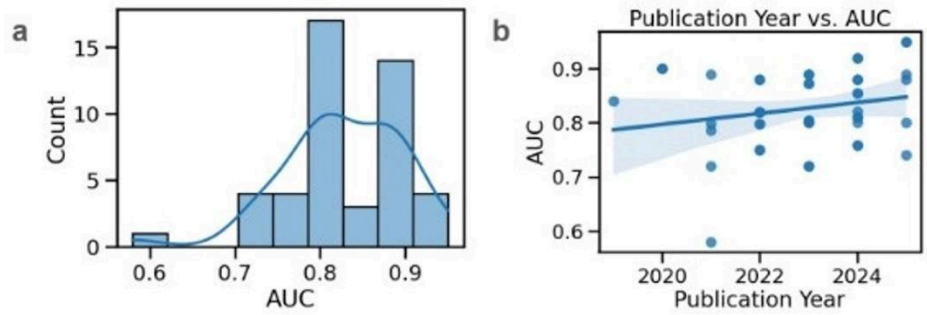


Figure 6: Summary of Diagnostic Accuracy and Performance Trends. (a) A histogram showing the bimodal distribution of reported Area Under the Curve (AUC) values from all included studies. (b) A regression plot illustrating the positive correlation between the publication year of the studies and their reported AUC values.

Performance Across Patient Ancestries

To further assess the generalizability of dd-cfDNA's performance, the AUC values were stratified by studies that reported the inclusion of different racial and ethnic groups. For this analysis, studies that did not report AUC or did not specify the racial demographics of their cohort were excluded, resulting in a subgroup of 45 studies. Overall, the data suggests that dd-cfDNA consistently has accuracy in identifying potential rejection across studies with different cohort demographics.

The median AUC of studies that specifically included Black participants (N=24) was 0.81 (Fig. 7a). This is an important observation supporting the biomarker's utility in this patient population, which has historically experienced higher rates of adverse transplant outcomes due to a combination of social and clinical factors.

It is critical to clarify that the available literature does not permit an analysis of dd-cfDNA performance exclusively within the Hispanic/Latino population. Instead, this review analyzed the group of studies that reported the inclusion of Hispanic/Latino participants within their larger, mixed-ancestry cohorts. For this collection of studies, the median AUC was 0.81, showing a narrow and consistent interquartile range (Fig. 7b). Therefore, while this result indicates that the biomarker performs well in diverse studies that include Hispanic patients, it is impossible to draw conclusions from this data about its specific performance in Hispanic populations alone.

The studies measuring the performance of dd-cfDNA that reported Asian participation strengthen the evidence for applicability of dd-cfDNA across racial and ethnic groups. Interestingly, the analysis suggests that studies including Asian participants demonstrate a trend towards higher and more consistent accuracy, with a median Area Under the Curve (AUC) of approximately 0.82 and a notably narrow interquartile range (Fig. 7c). This enhanced consistency may be attributed to factors like more homogeneous patient cohorts or the use of optimized, population-specific assays, highlighting the potential for dd-cfDNA's performance to be further refined in specific clinical and demographic contexts.

Finally, the analysis of studies including White participants provides further evidence of the biomarker's performance (Fig. 7d). The apparent better performance of studies that did not explicitly report including White participants is likely due to those studies being a heterogeneous mix of smaller or more homogeneous international cohorts where ethnicity is often not specified, whereas the "Yes" group consists primarily of large, diverse North American trials whose real-world complexity results in a more consistent, albeit slightly lower, median AUC.

While there are slight variations in the median AUC across studies that included participants of different self-reported ancestries, an analysis of variance (ANOVA)

confirmed that these differences are not statistically significant ($p > 0.05$). This is consistent with the visual interpretation of the overlapping confidence intervals in the box plots. Therefore, while the available data suggests that dd-cfDNA performs with comparable diagnostic accuracy across cohorts with different ancestral compositions, this conclusion is preliminary. More robust data, such as dedicated studies on Hispanic populations and a broader geographic representation for Asian populations, is needed to definitively confirm these findings and account for potential regional biases.

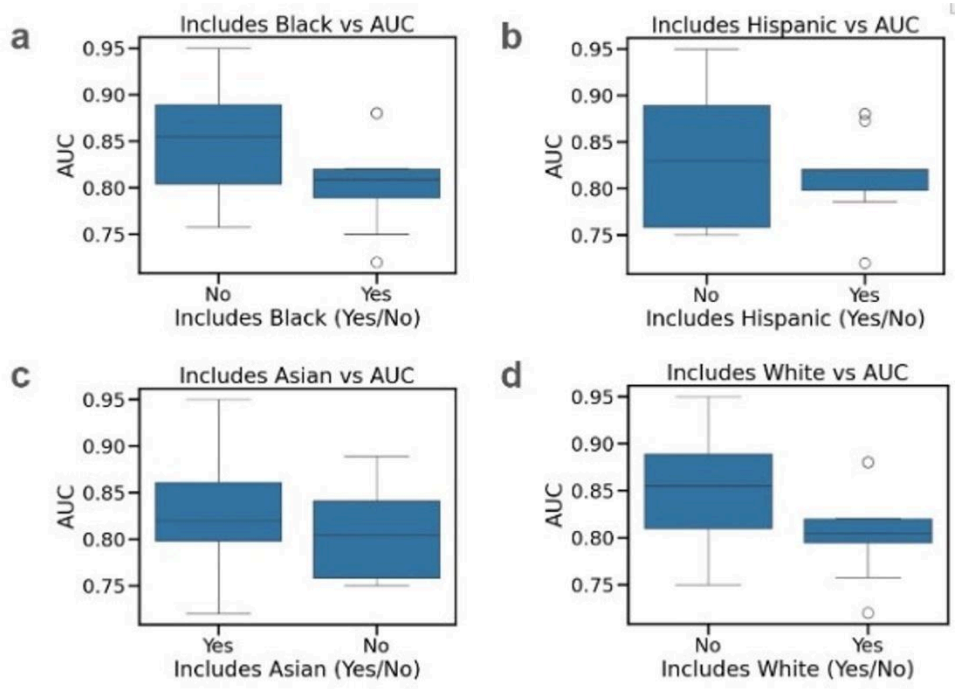


Figure 7: Diagnostic Accuracy (AUC) by Patient Ancestry Inclusion. A series of box plots showing the distribution of reported Area Under the Curve (AUC) values, comparing studies that did ("Yes") or did not ("No") report the inclusion of specific patient ancestries. (a) Distribution of AUC by inclusion of Black participants. (b) Distribution of AUC by inclusion of Hispanic participants. (c) Distribution of AUC by inclusion of Asian participants. (d) Distribution of AUC by inclusion of White participants.

Performance Across Different Study Contexts

Beyond the overall quality of the studies, it is also important to understand how the performance of dd-cfDNA varies across different study contexts, as detailed in Fig. 8. An analysis of performance by geographic region reveals a consistently high level of diagnostic accuracy globally, with the median AUC for studies from Europe, Asia,

and North America all above 0.80 (Fig. 8a). Furthermore, an analysis of study scale shows no significant correlation between a study's sample size and its reported AUC (Fig. 8b); the nearly flat regression line indicates that high performance is not simply an artifact of small studies, a conclusion reinforced by the finding that the largest study with nearly 3,000 patients still reported a strong AUC.

dd-cfDNA's performance is further evaluated by an analysis of demographic and methodological factors. The regression plot of average participant age versus AUC shows a slight downwards trend (Fig. 8c). This indicates that in the age group between ages 30 to 60, the AUC decreased as age increased, thereby signifying the need for more studies to further validate this relationship. Finally, while both single-center and multicenter studies reported similarly high median AUCs, the violin plot reveals that the results from multicenter studies were more consistent, with a tighter distribution and less variability (Fig. 8d). One aspect to note about Fig. 8d is that a violin plot was utilized over a standard box plot because it visualizes the full probability density of the data.

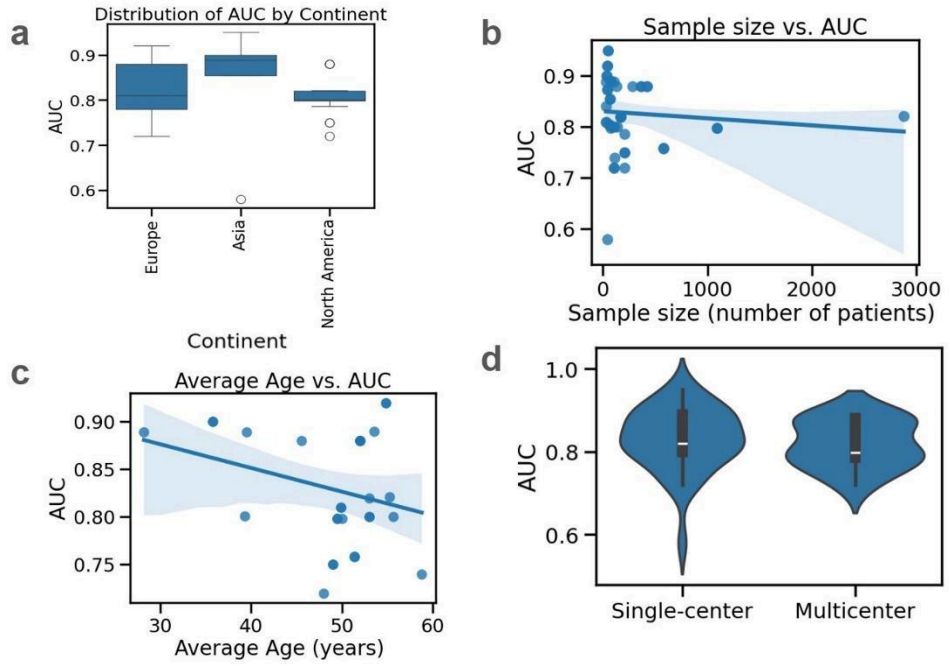


Figure 8: Analysis of dd-cfDNA Diagnostic Accuracy Across Different Study Contexts. (a) A box plot showing the distribution of reported AUC values, stratified by continent. (b) A regression plot illustrating the relationship between study sample size and reported AUC. (c) A regression plot showing the relationship between the average age of study participants and reported AUC. (d) A violin plot comparing the distribution of reported

AUC values between single-center and multicenter studies.

Relationship Between Study Quality and Diagnostic Accuracy

The relationship between the study scores was also analyzed, as determined by my 5-point relevance criteria, and their reported diagnostic accuracy (AUC). The box plot of overall quality scores reveals that studies with higher scores (3, 4, and 5) consistently reported high median AUCs above 0.80 (Fig. 9a). This indicates that the score of a study is not directly related to the reported AUC, as high diagnostic performance was observed across all quality tiers. A more granular analysis shows that studies incorporating more rigorous methodologies, such as longitudinal monitoring, not only reported a high median AUC but also showed more consistent results, evidenced by a tighter distribution of values compared to studies that did not meet this criterion (Fig. 9b).

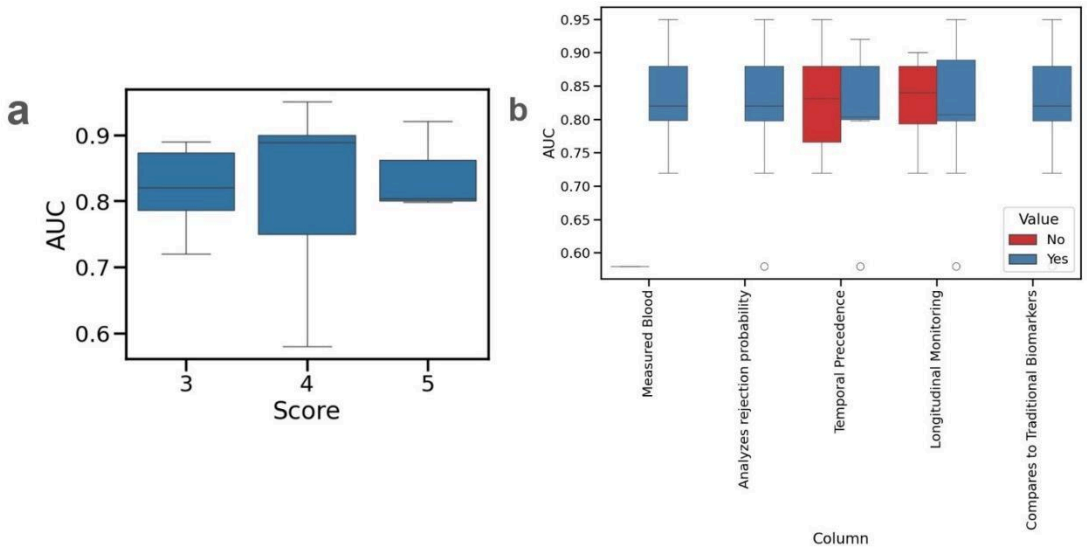


Figure 9: Relationship Between Study Quality and Diagnostic Accuracy (AUC). (a) A box plot showing the distribution of reported AUC values, stratified by the overall 5-point quality score of the studies. (b) A box plot comparing the distribution of reported AUC values for studies that did ("Yes") or did not ("No") meet each of the five specific methodological quality criteria.

Performance in Specific Rejection Subtypes

Lastly, to understand the evidence base for different rejection pathways, the

included studies were analyzed based on their inclusion of either antibody-mediated rejection (ABMR) or T-cell mediated rejection (TCMR). The analysis shows that the literature provides robust coverage of both primary rejection subtypes. A substantial majority of the included studies, approximately 48, investigated patient cohorts that included cases of ABMR (Fig. 10a), while a similarly large number of studies, approximately 45, included cases of TCMR (Fig. 10b), though many included both.

Furthermore, an analysis of the reported diagnostic accuracy (AUC) shows that the high performance of dd-cfDNA is consistent regardless of a study's specific focus on rejection subtypes. The median AUC for studies that included ABMR was approximately 0.82, a value nearly identical to the median AUC for studies that did not specifically focus on ABMR (Fig. 10c). A similar pattern was observed for TCMR, where the median AUC was also consistent at approximately 0.82 in studies that investigated this rejection type (Fig. 10d).

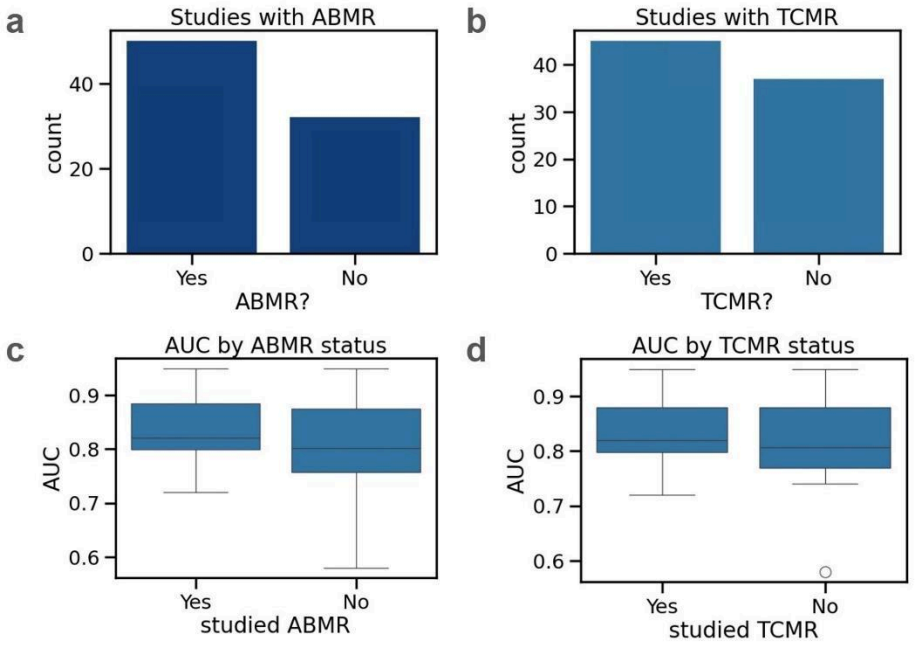


Figure 10: Research Focus on Rejection Subtypes and Corresponding Diagnostic Accuracy. (a) A bar chart showing the number of studies that did ("Yes") or did not ("No") investigate antibody-mediated rejection (ABMR). (b) A bar chart showing the number of studies that did ("Yes") or did not ("No") investigate T-cell mediated rejection (TCMR). (c) A box plot comparing the distribution of reported AUC values between studies that did or did not investigate ABMR. (d) A box plot comparing the distribution of reported AUC values between studies that did or did not investigate TCMR.

Methodological Gaps in Evaluating Rejection Timing

Despite this, there appears to be a large methodological gap in the literature regarding timing of rejection. While a reasonably sized number of studies (n=45) utilized some form of longitudinal monitoring for timing, a much smaller, and specific subset (n≈29) stated the temporal precedence of dd-cfDNA elevation before a rejection diagnosis was confirmed. This indicates a clear opportunity for future research to focus more rigorously on study designs that can quantify this early detection window.

Conclusions

The central finding of this systematic review—that donor-derived cell-free DNA (dd-cfDNA) serves as a robust biomarker for kidney allograft rejection with high diagnostic accuracy—is largely consistent with the conclusions of previous literature reviews and meta-analyses. Multiple reviews have affirmed the biomarker's high potential as a "valuable tool" for non-invasive monitoring (Gisch et al., 2025) and as a reliable liquid biopsy for detecting allograft injury (Oellerich et al., 2021). Furthermore, our findings on diagnostic accuracy align with several recent meta-analyses that have reported high pooled AUC values, such as 0.84 for acute rejection (Zhang et al., 2021), 0.86 for overall rejection (Zhang et al., 2025), and 0.88 for antibody-mediated rejection (Yang et al., 2024). However, the present review builds upon this established consensus by providing a more granular, quantitative synthesis of the evolving evidence base. While prior work has established the biomarker's general utility, our analysis is the first to systematically document the rapid maturation of the field, including the significant increase in study sample sizes and the shift towards more rigorous prospective cohort designs. Most importantly, our review addresses a critical knowledge gap by quantifying the significant geographic and ancestral disparities in the literature. Previous reviews have qualitatively mentioned the need for more diverse cohorts, but our analysis provides concrete evidence that data for Black and Hispanic populations are almost exclusively derived from North American studies, underscoring a limitation in the global generalizability of current findings that was not the primary focus of prior meta-analyses.

This systematic review serves as a thorough synthesis of recent and robust research on donor-derived cell-free DNA that helps solidify its importance in the

post-transplant period (Huang et al., 2023). The overall body of evidence supports that dd-cfDNA is consistently more accurate diagnostically than traditional biomarkers, such as serum creatinine, because it can detect rejection in those with stable creatinine levels (Gupta et al., 2022). The overall implication is that the use of dd-cfDNA represents a shift from reactive to proactive, evidence-based monitoring and screening that enables clinicians to detect allograft injury weeks or months before it presents clinically (Benning et al., 2023a). Although this review did systematically confirm that a number of studies reported a consistently high measure of diagnostic accuracy (e.g., > 0.80 as the AUC; Mantios et al., 2023), it also highlighted an important knowledge gap around the biomarker as most of the data on Black and Hispanic populations came exclusively from North American cohorts. One concrete recommendation based on this analysis is that future meta-analyses should synthesize other core metrics, such as Negative Predictive Value (NPV), in order to clearly define how dd-cfDNA should be used across different clinical scenarios (Aubert et al., n.d.). In summary, this systematic review concludes that dd-cfDNA is a milestone in transplant monitoring that provides a foundation for more personalized and non-invasive patient care (Loupy et al., 2024a).

It is also important to consider the limitations inherent in the designs of the included studies. While this review found a predominance of prospective cohort studies, which is a methodological strength, a notable number of studies were retrospective or single-center designs. Single-center studies may have limited external validity, while retrospective designs can be subject to selection bias and incomplete data. These factors represent a limitation of the overall evidence base and highlight the need for more large, prospective, multicenter trials to confirm the findings synthesized in this review. Furthermore, there was significant variability in diagnostic accuracy reported between different commercial assays used across the studies, representing another key limitation of the current evidence base.

This systematic review confirms that dd-cfDNA is a useful non-invasive biomarker for early allograft injury with long-term prognostic value (Kim et al., 2024a; Tian et al., n.d.; Nie et al., 2025; Loupy et al., 2024). While included studies demonstrated its ability to identify rejection sooner than traditional markers (Bromberg et al, 2024; Parajuli et al., 2024b), this review identified significant limitations, including a lack of ethnic diversity and variable accuracy across different assays (Nguyen et al., 2025). Furthermore, because dd-cfDNA is a dynamic marker whose levels can

become erratic, a single static measurement can be misleading (Nguyen et al., 2025). Therefore, future research must focus on validating performance in diverse populations and establishing standardized protocols for sequential monitoring to best differentiate true injury from transient fluctuations (Stites et al., 2020).

To fully realize the clinical potential of donor-derived cell-free DNA (dd-cfDNA), future research must focus on precision, standardization, and equitable application. This systematic review confirmed the biomarker's high diagnostic accuracy but also highlighted that different commercial assays can have different rejection detection rates for the same rejection subtype. Therefore, robust prospective trials are needed to conduct head-to-head comparisons of existing platforms and establish standardized reporting metrics. Furthermore, to ensure the clinical translatability and applicability of this biomarker, future studies must also assess its cost-effectiveness and the logistical challenges of integrating dd-cfDNA testing into routine clinical workflows. To ensure equitable and effective use, there is an immediate need for future research on underrepresented populations in Europe and Asia, as current evidence for Black and Hispanic populations is almost exclusively from North American studies. The ultimate success of this biomarker will also hinge on its ability to improve long-term patient outcomes, necessitating intervention trials that use dd-cfDNA's early warning signal to guide pre-emptive treatment. This would provide direct evidence on whether a dd-cfDNA-based strategy improves the trajectory of allograft health compared to the current standard of care. This potential for underreporting of negative or inconclusive studies represents a key limitation of the existing evidence base, and future meta-analyses should attempt to quantify the extent of this publication bias.

Code and Data Availability

Data collected as part of this systematic review is freely available at https://docs.google.com/spreadsheets/d/1KAfEUQF-zXAJviuMSKyAEOpJUpMJbEMjBCOCbhbmgj4/edit?usp=sharing. The Google Colab notebook used for analysis and plot generation is freely available at https://colab.research.google.com/drive/1oq564UU3FZyur_mscmA8rdN5-H1EsuY1?usp=sharing.

Acknowledgements

The author would like to express their sincere gratitude to [name redacted] whose insightful feedback and critical comments at every stage were invaluable in shaping the direction and clarity of this systematic review. Her thorough review of all manuscript drafts and figures ensured coherence and factual accuracy throughout the text. The author is especially grateful for her mentorship in data science, as she patiently taught me the fundamentals of programming in Python to transform the raw data into the meaningful figures presented in the Results section. She also provided essential support in the design and final polish of the graphical elements, making sure they clearly and accurately represented the data. This project would not have been possible without her generous investment of time, knowledge, and unwavering support.

Mentor Contribution Statement

[Mentor name redacted] served as the Research Mentor for this systematic review, guiding the author through the scientific research process and offering her expertise in data science methods and the paper structure. Her support was advisory and technical, focusing on providing insightful and critical feedback on all manuscript drafts and figures to ensure logical flow, content clarity, and factual accuracy. She meticulously reviewed all written sections and figures to confirm coherence and truthful representation of the systematic review data. Furthermore, she provided patient instruction on the fundamentals of Python programming required to transform the raw data into the final figures presented in the Results section. She also offered mentorship during the technical execution of data analysis and constructive suggestions during the revision process.

Author Biography

[Author name redacted] is a student researcher at [school name redacted]. He is passionate about the intersections of non-invasive diagnostics, clinical nephrology, and equitable patient care, particularly in transplantation medicine. His academic focus is on monitoring diseases and developing new biomarkers for personalized care strategies. His systematic review aimed to quantify the diagnostic strength of a non-invasive biomarker for kidney allograft rejection and assess the generalizability of the current research, noting gaps in demographic and geographic representation. He is also interested in data science, biostatistics, and the ethical

integration of new technologies into clinical practice. [Author name redacted] aspires to become a transplant surgeon who uses molecular tools like dd-cfDNA to transform clinical decision-making, bridging scientific inquiry with compassionate, technology-enabled patient monitoring to improve long-term graft survival.

References

Akalin, E., Weir, M. R., Bunnapradist, S., Brennan, D. C., Delos Santos, R., Langone, A., Djamali, A., Xu, H., Jin, X., Dholakia, S., Woodward, R. N., & Bromberg, J. S. (2021). Clinical Validation of an Immune Quiescence Gene Expression Signature in Kidney Transplantation. *Kidney*360, 2(12), 1998. https://doi.org/10.34067/KID.0005062021

Akifova, A., Budde, K., Amann, K., Buettner-Herold, M., Choi, M., Oellerich, M., Beck, J., Bornemann-Kolatzki, K., Schütz, E., Bachmann, F., Halleck, F., von Hoerschelmann, E., Koch, N., Schrezenmeier, E., Seelow, E., Waiser, J., Zukunft, B., Eckardt, K.-U., Halbritter, J., … Osmanodja, B. (2025). Donor-derived cell-free DNA monitoring for early diagnosis of antibody-mediated rejection after kidney transplantation: A randomized trial. *Nephrology Dialysis Transplantation*, 40(7), 1384–1395. https://doi.org/10.1093/ndt/gfae282

Aubert, O., Ursule-Dufait, C., & Brousse, R. (n.d.). *Cell-free DNA for the detection of kidney allograft rejection | Nature Medicine*. Retrieved August 11, 2025, from https://www.nature.com/articles/s41591-024-03087-3

Aubert, O., Ursule-Dufait, C., Brousse, R., Gueguen, J., Racapé, M., Raynaud, M., Van Loon, E., Pagliazzi, A., Huang, E., Jordan, S. C., Chavin, K. D., Gupta, G., Kumar, D., Alhamad, T., Anand, S., Sanchez-Garcia, J., Abdalla, B. A., Hogan, J., Garro, R., … Loupy, A. (2024). Cell-free DNA for the detection of kidney allograft rejection. *Nature Medicine*, 30(8), 2320–2327. https://doi.org/10.1038/s41591-024-03087-3

Benning, L., Morath, C., Fink, A., Rudek, M., Speer, C., Kälble, F., Nusshag, C., Beimler, J., Schwab, C., Waldherr, R., Zeier, M., Süsal, C., & Tran, T. H. (2023). Donor-Derived Cell-Free DNA (dd-cfDNA) in Kidney Transplant Recipients With Indication Biopsy—Results of a Prospective Single-Center Trial. *Transplant International*, 36, 11899. https://doi.org/10.3389/ti.2023.11899

Benning, L., Morath, C., Kühn, T., Bartenschlager, M., Kim, H., Beimler, J., Buylaert, M., Nusshag, C., Kälble, F., Reineke, M., Töllner, M., Schaier, M., Klein, K., Blank, A.,

Schnitzler, P., Zeier, M., Süsal, C., Bartenschlager, R., Tran, T. H., & Speer, C. (2022). Humoral response to SARS-CoV-2 mRNA vaccination in previous non-responder kidney transplant recipients after short-term withdrawal of mycophenolic acid. *Frontiers in Medicine*, 9. https://doi.org/10.3389/fmed.2022.958293

Botella, C., Galián, J. A., Jiménez-Coll, V., Fernández-González, M., Morales, F., Martínez-Gómez, G., González-López, R., Alegría, M. J., Moya, M. R., Martinez-Banaclocha, H., Minguela, A., Legaz, I., Llorente, S., & Muro, M. (2024a). Early Monitoring of Donor-Derived Cell-Free DNA in Kidney Allograft Recipients Followed-Up for Two Years: Experience of One Center. *Life*, 14(11), 1491. https://doi.org/10.3390/life14111491

Botella, C., Galián, J. A., Jiménez-Coll, V., Fernández-González, M., Morales, F., Martínez-Gómez, G., González-López, R., Alegría, M. J., Moya, M. R., Martinez-Banaclocha, H., Minguela, A., Legaz, I., Llorente, S., & Muro, M. (2024b). Early Monitoring of Donor-Derived Cell-Free DNA in Kidney Allograft Recipients Followed-Up for Two Years: Experience of One Center. *Life*, 14(11), 1491. https://doi.org/10.3390/life14111491

Bromberg, J. S., Bunnapradist, S., Samaniego-Picota, M., Anand, S., Stites, E., Gauthier, P., Demko, Z., Prewett, A., Armer-Cabral, M., Marshall, K., Kaur, N., Bloom, M. S., Tabriziani, H., Bhorade, S., & Cooper, M. (2024). Elevation of Donor-derived Cell-free DNA Before Biopsy-proven Rejection in Kidney Transplant. *Transplantation*, 108(9), 1994–2004. https://doi.org/10.1097/TP.0000000000005007

Bu, L., Gupta, G., Pai, A., Anand, S., Stites, E., Moinuddin, I., Bowers, V., Jain, P., Axelrod, D. A., Weir, M. R., Wolf-Doty, T. K., Zeng, J., Tian, W., Qu, K., Woodward, R., Dholakia, S., Golovine, A. D., Bromberg, J. S., Murad, H., & Alhamad, T. (2022). Clinical outcomes from the Assessing Donor-derived cell-free DNA Monitoring Insights of kidney Allografts with Longitudinal surveillance (ADMIRAL) study. *Kidney International*, 101(4), 793–803. https://doi.org/10.1016/j.kint.2021.11.034

Chen, X.-T., Qiu, J., Wu, Z.-X., Zhang, H., Chen, T., Yang, S.-C., Zhao, G.-D., He, Y., Shen, X., Luo, J.-Q., Huang, Y., Wang, C.-X., Chen, L.-Z., Wu, C.-L., & Huang, G. (2022). Using Both Plasma and Urine Donor-Derived Cell-Free DNA to Identify Various Renal Allograft Injuries. *Clinical Chemistry*, 68(6), 814–825.

https://doi.org/10.1093/clinchem/hvac053

Cheng, D., Liu, F., Xie, K., Zeng, C., Li, X., Ni, X., Ge, J., Shu, L., Zhou, Y., Shi, H., Liu, H., & Chen, J. (2021). Donor-derived cell-free DNA: An independent biomarker in kidney transplant patients with antibody-mediated rejection. *Transplant Immunology*, 69, 101404. https://doi.org/10.1016/j.trim.2021.101404

Cuadrado-Payán, E., González-Roca, E., Rodríguez-Espinosa, D., Molina-Andújar, A., Montagud-Marrahi, E., Arana, C., González-Rojas, A., Esforzado, N., Torregrosa, V., Ventura-Aguiar, P., Argudo, M., Serrano-Jorcano, D., Ramírez-Bajo, M. J., Bañón-Maneus, E., Casas, S., Broseta, J. J., Puig-Butille, J. A., Revuelta, I., Diekmann, F., & Cucchiari, D. (2025). Implementation of the in-house technique for the determination of donor-derived cell-free DNA in daily clinical practice: Experience from the Hospital Clinic of Barcelona. *Nefrología (English Edition)*, 45(4), 294–301. https://doi.org/10.1016/j.nefroe.2025.03.004

Dreige, Y., Lima, G., Cristelli, M., Nakamura, M., Foresto, R., Pierry, P., Glehn-Ponsirenas, R., Levi, J., Medina-Pestana, J., & Tedesco-Silva, H. (2022). 336.3: Clinical Utility of Donor-Derived Cell-Free DNA During the Period of Recovery of Renal Function After Kidney Transplantation. *Transplantation*, 106(9S), S297. https://doi.org/10.1097/01.tp.0000886964.60815.ca

Halloran, P. F., Reeve, J., Madill-Thomsen, K. S., Kaur, N., Ahmed, E., Cantos, C., Al Haj Baddar, N., Demko, Z., Liang, N., Swenerton, R. K., Zimmermann, B. G., Van Hummelen, P., Prewett, A., Rabinowitz, M., Tabriziani, H., Gauthier, P., Billings, P., & Investigators*, the T. (2022). Combining Donor-derived Cell-free DNA Fraction and Quantity to Detect Kidney Transplant Rejection Using Molecular Diagnoses and Histology as Confirmation. *Transplantation*, 106(12), 2435. https://doi.org/10.1097/TP.0000000000004212

Kidney Disease Statistics for the United States—NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved August 11, 2025, from https://www.niddk.nih.gov/health-information/health-statistics/kidney-disease

Kim, H. D., Bae, H., Kang, H., Lee, H., Eum, S. H., Yang, C. W., Choi, Y. J., Chung, B. H., & Oh, E.-J. (2024a). Donor-derived cell-free DNA predicted allograft rejection and severe microvascular inflammation in kidney transplant recipients. *Frontiers in*

Immunology, 15, 1433918. https://doi.org/10.3389/fimmu.2024.1433918

Kim, H. D., Bae, H., Kang, H., Lee, H., Eum, S. H., Yang, C. W., Choi, Y. J., Chung, B. H., & Oh, E.-J. (2024b). Donor-derived cell-free DNA predicted allograft rejection and severe microvascular inflammation in kidney transplant recipients. *Frontiers in Immunology*, 15, 1433918. https://doi.org/10.3389/fimmu.2024.1433918

Kumar, N., Tandon, A., Rana, R., Rana, D. S., Bhalla, A. K., Gupta, A., Sachdeva, M. P., Huirem, R. S., Chauhan, K., Yashavarddhan, M. H., Basnal, A., Gupta, R., Mallick, P. K., & Ganguly, N. K. (2023). Donor-Derived Cell-Free DNA as a Non-Invasive Biomarker for Graft Rejection in Kidney Transplant Recipients: A Prospective Study among the Indian Population. *Diagnostics*, 13(23), 3540. https://doi.org/10.3390/diagnostics13233540

Lakhani, L., Alasfar, S., Bhalla, A., Aala, A., Rosenberg, A., Ostrander, D., Schollenberger, M. D., Brennan, D. C., & Lipson, E. J. (2021). Utility of Serial Donor-derived Cell-free DNA Measurements for Detecting Allograft Rejection in a Kidney Transplant Recipient After PD-1 Checkpoint Inhibitor Administration. *Transplantation Direct*, 7(2), e656. https://doi.org/10.1097/TXD.0000000000001113

Loupy, A., Certain, A., Tangprasertchai, N. S., Racapé, M., Ursule-Dufait, C., Benbadi, K., Raynaud, M., Vaskova, E., Marchis, C., Casas, S., Hague, T., Bestard, O., Kervella, D., Lefaucheur, C., Viard, T., & Aubert, O. (2024). Evaluation of a Decentralized Donor-Derived Cell-Free DNA Assay for Kidney Allograft Rejection Monitoring. *Transplant International*, 37, 13919. https://doi.org/10.3389/ti.2024.13919

Mantios, E., Filiopoulos, V., Constantoulakis, P., Liapis, G., Vittoraki, A., Casas, S., Marinaki, S., & Boletis, J. N. (2023). Assessment of Donor Derived Cell Free DNA (dd-cfDNA) at Surveillance and at Clinical Suspicion of Acute Rejection in Renal Transplantation. *Transplant International*, 36, 11507. https://doi.org/10.3389/ti.2023.11507

Mayer, K. A., Doberer, K., Tillgren, A., Viard, T., Haindl, S., Krivanec, S., Reindl-Schwaighofer, R., Eder, M., Eskandary, F., Casas, S., Wahrmann, M., Regele, H., & Böhmig, G. A. (2021). Diagnostic value of donor-derived cell-free DNA to predict antibody-mediated rejection in donor-specific antibody-positive renal allograft recipients. *Transplant International*, 34(9), 1689–1702.

https://doi.org/10.1111/tri.13970

Moein, M., Settineri, J. P., Suleiman, H., Sidhu, J., Papa, S., Coyle, S., Dvorai, R. H., Bahreini, A., Leggat, J., & Saidi, R. F. (2024). Application of Combined Donor-Derived Cell-Free DNA and Transcriptome in Diagnosis of Kidney Transplant Rejection. *Transplantation Proceedings*, 56(6), 1259–1263. https://doi.org/10.1016/j.transproceed.2024.02.023

Montgomery, R., & Gelb, B. (n.d.). *Four Things You May Not Know About Kidney Transplants*. NYU Langone News. Retrieved August 11, 2025, from https://nyulangone.org/news/four-things-you-may-not-know-about-kidney-transplants

Nguyen, P. T., Nakaoka, H., Mitsunaga, S., Aoyama, H., Kitamura, H., Saigo, K., & Inoue, I. (2025). *Applications of donor-derived cell-free DNA in kidney transplantation healthcare: View from a prospective single-center study*. Research Square. https://doi.org/10.21203/rs.3.rs-6592444/v1

Nie, W., Wang, Y., Fu, Q., Wu, C., Deng, R., Yu, X., Ye, C., Liu, X., Xu, B., Sun, P., Liu, L., Li, J., Zhang, H., & Wang, C. (2025). Integrating Donor Derived Cell-Free DNA Fraction and Absolute Quantification for Enhanced Rejection Diagnosis in Kidney Transplant Recipients. *Diagnostics*, 15(3), 237. https://doi.org/10.3390/diagnostics15030237

Oellerich, M., Sherwood, K., Keown, P., Schütz, E., Beck, J., Stegbauer, J., Rump, L. C., & Walson, P. D. (2021). Liquid biopsies: Donor-derived cell-free DNA for the detection of kidney allograft injury. *Nature Reviews Nephrology*, 17(9), 591–603. https://doi.org/10.1038/s41581-021-00428-0

Orandi, B. J., Luo, X., Massie, A. B., Garonzik-Wang, J. M., Lonze, B. E., Ahmed, R., Van Arendonk, K. J., Stegall, M. D., Jordan, S. C., Oberholzer, J., Dunn, T. B., Ratner, L. E., Kapur, S., Pelletier, R. P., Roberts, J. P., Melcher, M. L., Singh, P., Sudan, D. L., Posner, M. P., … Segev, D. L. (2016). Survival Benefit with Kidney Transplants from HLA-Incompatible Live Donors. *New England Journal of Medicine*, 374(10), 940–950. https://doi.org/10.1056/NEJMoa1508380

Organ transplants exceeded 48,000 in 2024; a 3.3 percent increase from the transplants performed in 2023—OPTN. (n.d.). Retrieved August 11, 2025, from

https://optn.transplant.hrsa.gov/news/organ-transplants-exceeded-48-000-in-2024-a-33-percent-increase-from-the-transplants-performed-in-2023/

Parajuli, S., Garg, N., Dodin, B., Breyer, I., Zona, E., Patel, S., Pinney, K., & Mandelbrot, D. (2024a). Changes in Donor-Derived Cell-Free DNA Before and After Rejection and De Novo DSA Detection in Primary and Repeat Kidney Transplant Recipients. *Clinical Transplantation*, 38(11), e70019. https://doi.org/10.1111/ctr.70019

Parajuli, S., Garg, N., Dodin, B., Breyer, I., Zona, E., Patel, S., Pinney, K., & Mandelbrot, D. (2024b). Changes in Donor-Derived Cell-Free DNA Before and After Rejection and De Novo DSA Detection in Primary and Repeat Kidney Transplant Recipients. *Clinical Transplantation*, 38(11), e70019. https://doi.org/10.1111/ctr.70019

Publish or Perish. (n.d.). Retrieved August 11, 2025, from https://harzing.com/resources/publish-or-perish

Quick kidney disease facts and stats | American Kidney Fund. (2024, April 19). https://www.kidneyfund.org/all-about-kidneys/quick-kidney-disease-facts-and-stats

Rizvi, A., Faiz, S., Thakkar, P. H., Hussain, S., Gamilla-Crudo, A. N., Kueht, M., & Mujtaba, M. A. (2023a). Kidney Allograft Monitoring by Combining Donor-Derived Cell-Free DNA and Molecular Gene Expression: A Clinical Management Perspective. *Journal of Personalized Medicine*, 13(8), 1205. https://doi.org/10.3390/jpm13081205

Rizvi, A., Faiz, S., Thakkar, P. H., Hussain, S., Gamilla-Crudo, A. N., Kueht, M., & Mujtaba, M. A. (2023b). Kidney Allograft Monitoring by Combining Donor-Derived Cell-Free DNA and Molecular Gene Expression: A Clinical Management Perspective. *Journal of Personalized Medicine*, 13(8), 1205. https://doi.org/10.3390/jpm13081205

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021a). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., &

Jordan, S. C. (2021b). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Sawinski, D. L., Mehta, S., Alhamad, T., Bromberg, J. S., Fischbach, B., Aeschbacher, T., Ghosh, S., Shekhtman, G., Dholakia, S., Brennan, D. C., Poggio, E., Bloom, R. D., & Jordan, S. C. (2021c). Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clinical Transplantation*, 35(9), e14402. https://doi.org/10.1111/ctr.14402

Schenk, K. M., Deutsch, J. S., Chandra, S., Davar, D., Eroglu, Z., Khushalani, N. I., Luke, J. J., Ott, P. A., Sosman, J. A., Aggarwal, V., Schollenberger, M. D., Sharfman, W. H., Bibee, K. P., Scott, J. F., Loss, M. J., Wang, H., Qi, H., Sharon, E., Streicher, H., … Lipson, E. J. (2024). Nivolumab + tacrolimus + prednisone +/− ipilimumab for kidney transplant recipients with advanced cutaneous cancers. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *42*(9), 1011–1020. https://doi.org/10.1200/JCO.23.01497

Sharma, R., Wolff, J., & Wilpula, E. (2022). Novel Technology Has Made It Possible To Diagnose A New Clinical Phenotype Of Rejection called Early Antibody Mediated Rejection, With Significant Therapeutic Implications For Renal Allograft Survival. *Experimental and Clinical Transplantation*, 20(10), 967–969. https://doi.org/10.6002/ect.2021.0489

Shen, J., Guo, L., Yan, P., Zhou, J., Zhou, Q., Lei, W., Liu, H., Liu, G., Lv, J., Liu, F., Huang, H., Dong, W., Shu, L., Wang, H., Wu, J., Chen, J., & Wang, R. (2020). Prognostic value of the donor-derived cell-free DNA assay in acute renal rejection therapy: A prospective cohort study. *Clinical Transplantation*, 34(10), e14053. https://doi.org/10.1111/ctr.14053

Stites, E., Kumar, D., Olaitan, O., John Swanson, S., Leca, N., Weir, M., Bromberg, J., Melancon, J., Agha, I., Fattah, H., Alhamad, T., Qazi, Y., Wiseman, A., & Gupta, G. (2020). High levels of dd-cfDNA identify patients with TCMR 1A and borderline allograft rejection at elevated risk of graft injury. *American Journal of Transplantation*, 20(9), 2491–2498. https://doi.org/10.1111/ajt.15822

Tian, Y., Frischknecht, L., Anna Mallone, Rossler, F., Schachtner, T., & Nilsson, J.

(n.d.). *Evaluation of de novo donor specific antibodies after kidney transplantation in the era of donor-derived cell-free DNA - PMC*. Retrieved August 11, 2025, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11779610/

Wei, L., Zhao, Y., Deng, S., Wu, S., Wang, H., Luo, X., & Yang, H. (2024). Graft-derived cell free DNA: Used for assessment of early graft status and its implications for long-term kidney function. *Frontiers in Physiology*, 15, 1440799. https://doi.org/10.3389/fphys.2024.1440799

Wolf-Doty, T. K., Mannon, R. B., Poggio, E. D., Hinojosa, R. J., Hiller, D., Bromberg, J. S., & Brennan, D. C. (2021). Dynamic Response of Donor-Derived Cell-Free DNA Following Treatment of Acute Rejection in Kidney Allografts. *Kidney360*, 2(4), 729. https://doi.org/10.34067/KID.0000042021

Zhang, H., Zheng, C., Li, X., Fu, Q., Li, J., Su, Q., Zeng, L., Liu, Z., Wang, J., Huang, H., Xu, B., Ye, M., Liu, L., & Wang, C. (2020a). Diagnostic Performance of Donor-Derived Plasma Cell-Free DNA Fraction for Antibody-Mediated Rejection in Post Renal Transplant Recipients: A Prospective Observational Study. *Frontiers in Immunology*, 11, 342. https://doi.org/10.3389/fimmu.2020.00342

Zhang, H., Zheng, C., Li, X., Fu, Q., Li, J., Su, Q., Zeng, L., Liu, Z., Wang, J., Huang, H., Xu, B., Ye, M., Liu, L., & Wang, C. (2020b). Diagnostic Performance of Donor-Derived Plasma Cell-Free DNA Fraction for Antibody-Mediated Rejection in Post Renal Transplant Recipients: A Prospective Observational Study. *Frontiers in Immunology*, 11. https://doi.org/10.3389/fimmu.2020.00342