Donor Derived Cell Free DNA: is it all the same?

Joseph K. Melancon\textsuperscript{1}

\textsuperscript{1}Department of Surgery, The George Washington University Hospital, Washington, D.C., U.S.A.

Corresponding Author:
Joseph K. Melancon, MD
2131 K Street NW, Washington, D.C., 20037, U.S.A.
Email address: jmelancon@mfa.gwu.edu
Abstract

Background: Clinical utility of donor derived cell-free DNA (dd-cfDNA) in transplantation has been extensively reviewed, supporting its use as a surveillance tool for the early and accurate detection of allograft injury. Yet studies comparing different assay methods have been lacking.

Methods: Paired sampling of commercially available dd-cfDNA (AlloSure and Prospera) was compared and examined against histology and manufacturer guidance. 76 patients were prospectively assessed, with 11 biopsy proven rejections (ABMR, n=2, TCMR, n=9).

Results: Prospera demonstrated larger measurements of dd-cfDNA in comparison to AlloSure, but this was not significant p=0.1233. AlloSure results supported its published guidance, with no significant differences between diagnostic test characteristics observed while Prospera was discordant against its published guidance, missing two TCMR1A cases. AlloSure demonstrated a significantly shorter turnaround time (p=0.0001) from blood draw to patient result.

Conclusion: Although dd-cfDNA’s are similar they are not the same. Extensive evidence for dd-cfDNA interpretation remains key to building clinical utility when considering clinical implementation and remaining consistent to a single platform is important when creating data comparisons.

Introduction

The clinical utility of donor derived cell-free DNA (dd-cfDNA) in transplantation has been demonstrated in the management of patients, where it has been widely used as a surveillance tool for the early and accurate detection of allograft injury. The validity and utility of this important new biomarker has been extensively reviewed. However studies comparing different assay methods have been lacking, therefore we compared commercially available test services assessing whether differences are clinically meaningful.

Two tests commercially available for patient management were evaluated, AlloSure (CareDx, Inc., Brisbane, CA) and Prospera (Natera, Inc., San Carlos, CA). Both have published analytical validation studies but have different degrees of clinical validation. AlloSure has large prospective, multicenter data, while Prospera was substantiated using a single-center, retrospective biobank. Both techniques use proprietary library preparation methods and standard next-generation sequencing (NGS) to quantify dd-cfDNA.

NGS technologies enable sequencing small amounts of DNA in a more quantifiable and robust manner, with acceptable time to return diagnostic tests. For NGS assays to discriminate dd-cfDNA from recipient cell-free DNA, polymorphisms between donor and recipient are exploited. Single-nucleotide polymorphisms (SNPs) are the choice for building scalable, high-quality tests to quantify dd-cfDNA, however there is no evidence to show more SNPs assessed is associated with superior outcomes.

Current publications around dd-cfDNA are discordant, with some suggesting superiority and use of second generation technology. Yet both AlloSure and Prospera use Illumina based NGS platforms and SNP methodology, and so, it remains unclear whether these claims are true or not. The objective of this study is to provide an early real-life single center experience comparing results using the commercially available tests to assess these claims using paired values in renal allograft recipients.

Materials & Methods

Sample Collection and dd-cfDNA quantification
Between November 2019 and March 2020, 76 single renal transplant patients from our center had simultaneous AlloSure and Prospera dd-cfDNA tests prospectively performed comparing dd-cfDNA levels
and reported association with clinical outcomes. Consecutive post-transplant patients seen for surveillance biopsy appointments at either 3 or 12 months were included in the study. The sample size was an unpowered pilot, determined by the number of kits that were available without causing dual billing problems. Prospera kits were donated by the manufacturer and internal research funds were used to perform this work. At the time of study AlloSure was reimbursed by Medicare and being used under a surveillance and for-cause protocol (months 1,2,3,4,6,9,12 within the first year, and every 3 months thereafter) at our site, while the Natera test was not reimbursable and used only paired with an AlloSure. Dd-cfDNA samples were collected and associated with surveillance biopsy which is performed at our institution either month 3 or month 12 post-transplant (Figure 1). Local IRB approval was granted. Interpretation and performance of the tests was based on the manufacturers’ guidance and published validity data for the respective assays. Prospera uses the 1% cut off, where AlloSure uses a gradient, considering relative change above 0.5%. Turnaround time (TAT) from blood draw to the returned test result was recorded. Table 1 compares the two assays.

Patient Demographics and Management
Our typical patient population for our institution is highly sensitized, with PRA>20%, larger AA population, more frequent HLA-mismatches with the primary causes of end stage renal disease caused by hypertension and diabetes (Supplemental Table 1). All patients had thymoglobulin at 3mg/kg and 1 dose of Simulect at 20mg. Patients were maintained on extended release tacrolimus and mycophenolate mofetil for maintenance and weaned from steroids by 1-week post-operative, with DSA testing done for cause.

Histological reads - Biopsy-proven rejection
Biopsies were read and scored by the local institutional pathologist according to Banff 2017 classification. Results were classified as antibody mediated rejection (ABMR, n=2), T Cell mediated rejection (TCMR, n=9), acute tubular necrosis (ATN, n=4), BK nephropathy (BKVAN, n=1) or normal (n=60). 52 dd-cfDNA tests were ordered as part of surveillance, with 24 were ordered due to a change in clinical presentation (Formation of de-novo donor specific antibody (DSA), change in creatinine, change in Tacrolimus level, Figure 1).

Statistical Analysis
The Kruskal-Wallis test was used to evaluate the distributions of the difference between the paired dd-cfDNA percentages (AlloSure – Prospera) across the categories of biopsy occurrence. The Wilcoxon signed rank test was used to evaluate the difference between the paired turnaround times between the assays. The method outlined by Hanley and McNeil was used to derive sample size estimates based upon comparing AUC ROC for paired data. All statistical tests reported are two-sided, where statistical significance is defined as a p-value less than 0.05.

Results
A total of 76 samples were assessed with both AlloSure and Prospera, which included 11 (14%) biopsy proven rejections. Measurements of dd-cfDNA between the testing methods for the paired samples (excluding 1 paired sample with AlloSure score of > 16 and Prospera score of 19.61) tend to demonstrate larger measurements for Prospera in comparison to AlloSure as shown in Figure 2A (0%-1%) and Figure 2B (0%-20%), (Kruskal-Wallis test chi-square=2.5916, df=2, p=0.2737). This trend was most apparent for samples with associated biopsy confirmed rejections (Figure 2C).

Dd-cfDNA as a molecular marker of injury may be elevated by other causes, these include BKVAN. The 1 patient with BKVAN had SV40 staining on biopsy with blood-based PCR load of 100,000 copies. AlloSure
has evidence suggesting the degree of injury can differentiate viremia from nephropathy, but the sample is small with active trials undergoing. Prospera has no data on its utility on BK.\textsuperscript{11}

No significant differences amongst diagnostic test characteristics were observed, despite AlloSure trending higher performance. This is based upon the estimates obtained and their associated 95% confidence intervals (Table 2, Table 3A-3B and Table 4). A prohibitively large number of samples would be required to delineate the nominal differences, which are not clinically relevant. For example, the estimated AUC ROC for AlloSure and Prospera are 0.7343 and 0.7483, respectively. To demonstrate AUC ROC superiority of Prospera relative to AlloSure, assuming a significance level of 5%, power to detect a difference of 80%, correlation among the paired samples of 96% and 15% of the biopsies performed resulted in rejection, the estimated number of samples that would need to be collected, tested across both diagnostic platforms, and have biopsies performed to confirm rejection is more than 2300. Differences between the assays matter most for accurately detecting all types of rejection. Prospera missed 3 cases of TCMR1A using the 1% threshold which were detected by AlloSure using its published guidance of 0.5%. Even at 0.5%, Prospera missed two TCMR1A cases, which were identified by AlloSure.

The most meaningful difference captured in this study is the amount of time from sample collection to the availability of the reported results. Figure 3 displays the distribution and cumulative distribution of the difference in turnaround time between the paired samples. 75% of the paired samples had a difference of at most -1 day, indicating 75% of the AlloSure tests were reported at least 1 day earlier than Prospera tests (Figure 3). With the largest difference being 5 days for 2 samples (3%). There were 3 samples (4%) where the AlloSure samples took longer to report than the paired Prospera samples, with a difference of a single day.

**Discussion**

While AlloSure and Prospera both use the novel clinical parameter of measuring the percentage of dd-cfCDNA to great precision, our data remains consistent with the literature surrounding AlloSure and differs from the claims for the reported performance of Prospera. As suggested by Grskovic et al\textsuperscript{12}, this may be due to the study design of the Prospera validation study. In addition, results demonstrated AlloSure has a superior turnaround time (Wilcoxon signed rank test, $S=1463$, $n=76$, $p<0.0001$) from blood draw to patient result.

The concept of more being better is one we are all programmed to believe, yet this report supports the focused approach of AlloSure, confirms it is not outperformed by the more complex methodology used by Prospera and in fact is more accurate in clinical interpretation and significantly faster to gain patient results. Further to this, both tests were drawn at the same time, both have similar workflows for ordering and both dependent on courier-based service. In this regard, the superior AlloSure TAT makes it much more accessible in terms of making clinical decisions that guide patient care.

Assessing the overall performance of both, the AUC of the receiver operator curves were not statistically different, although more reflective of the published AlloSure AUC (77%) than the Prospera AUC (87%) which is likely due to ascertainment bias. This is important, as diagnostic differences between assays would need very large numbers to be able to differentiate, yet the clinically relevant implications show the data supporting AlloSure is much more accurate and robust, with more clinical utility.
Prospera missed 3 cases of TCMR1A using the 1% threshold which were detected by AlloSure using its published guidance of 0.5%. Even at 0.5%, Prospera missed two TCMR1A cases, two of which were caught by AlloSure. Although not statistically significant; the clinical interpretation of TCMR1A is critical to clinical management of patients. The data published to guide interpretation of TCMR1A is much more robust for AlloSure and is consistent with our experience.9

This series has limitations due to an underpowered sample size, single center pilot experience and single time point used in each patient, which was done to preserve patients from additional and unnecessary blood work. A larger prospective, multi-center study is needed to validate these findings; however, these results suggest there may be little to gain. It is evident as more data is generated, that cross walking published dd-cfDNA data across different platforms may be ineffective and that although dd-cfDNA’s are similar they are not the same. This is increasingly important when managing patient populations, to allow consistent comparison using a single assay.

We have utilized surveillance biopsies to follow our patients since most of our kidney transplant recipients are highly sensitized and therefore have an increased risk of acute rejection. However, considering growing evidence, we are re-evaluating our protocol and believe dd-cfDNA monitoring has a place to support patients as part of their post-transplant surveillance. However, with the wide adoption of dd-cfDNA and the potential for further assays entering the field, a clear understanding of the technology and evaluation of clinical validation data supports the importance to remain consistent to a single platform.

More importantly, the published supporting evidence, use of real-life data, and the need for rapid turnaround to guide patient management is critical when considering clinical adoption.

Conclusions

Both dd-cfDNA platforms are useful and provide important adjuncts in the transplant clinician’s armamentarium. We believe that the use of dd-cfDNA monitoring will continue to expand clinically and that its future accuracy and precision will increase, but this needs to be supported by extensive clinical evidence and data when considering its implementation into a clinical program.

Disclosures

J Melancon is on the speaker’s bureau of Natera and CareDx. I declare no conflict of interest and performed this study independent of any company involvement.

Acknowledgements/Funding

No external funding was received.

Author Contributions

J Melancon: Conceptualization; Data curation; Formal analysis; Investigation; Writing - original draft; Writing - review and editing

Supplemental Data

Supplement Table 1. Patient Cohort Demographics
References

Table 1. Side-by-side characteristics of Prospera vs. AlloSure

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Prospera</th>
<th>AlloSure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limit of Detection</td>
<td>0.15%</td>
<td>0.12%</td>
</tr>
<tr>
<td>Claimed Sensitivity for TCMR</td>
<td>100% at 1%</td>
<td>RCV above 0.5%</td>
</tr>
<tr>
<td>Claimed AUC for all rejection</td>
<td>87%</td>
<td>77%</td>
</tr>
<tr>
<td>Number of SNPs</td>
<td>13962</td>
<td>405</td>
</tr>
<tr>
<td>Precision (CV)</td>
<td>4.29%</td>
<td>4.5% and 7.7% for samples above and below 2% dd-cfDNA respectively</td>
</tr>
<tr>
<td>Lowest input material needed</td>
<td>15ng</td>
<td>3ng</td>
</tr>
<tr>
<td>Blood Tubes Needed</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2. Side-by-side performance characteristics between AlloSure and Prospera for all rejection at 1% threshold. NPV (negative predictive value); PPV (positive predictive value).

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>AlloSure</th>
<th>Prospera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity [95% Exact CI]</td>
<td>0.4545 [0.1675, 0.7662]</td>
<td>0.5455 [0.2338, 0.8325]</td>
</tr>
<tr>
<td>Specificity [95% Exact CI]</td>
<td>0.8462 [0.5455, 0.9808]</td>
<td>0.6923 [0.3857, 0.9091]</td>
</tr>
<tr>
<td>PPV [95% Exact CI]</td>
<td>0.7143 [0.2904, 0.9633]</td>
<td>0.6000 [0.2624, 0.8784]</td>
</tr>
<tr>
<td>NPV [95% Exact CI]</td>
<td>0.6471 [0.3833, 0.8579]</td>
<td>0.6429 [0.3514, 0.8724]</td>
</tr>
</tbody>
</table>
Table 3A. Side-by-side performance characteristics between AlloSure and Prospera for cellular rejection at 0.5% threshold. NPV (negative predictive value); PPV (positive predictive value).

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>AlloSure</th>
<th>Prospera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity [95% Exact CI]</td>
<td>0.7778 [0.3999, 0.9719]</td>
<td>0.5556 [0.2120, 0.8630]</td>
</tr>
<tr>
<td>Specificity [95% Exact CI]</td>
<td>0.5333 [0.2659, 0.7873]</td>
<td>0.5333 [0.2659, 0.7873]</td>
</tr>
<tr>
<td>PPV [95% Exact CI]</td>
<td>0.5000 [0.2304, 0.7696]</td>
<td>0.4167 [0.1517, 0.7233]</td>
</tr>
<tr>
<td>NPV [95% Exact CI]</td>
<td>0.8000 [0.4439, 0.9748]</td>
<td>0.6667 [0.3489, 0.9008]</td>
</tr>
</tbody>
</table>

Table 3B. Side-by-side performance characteristics between AlloSure and Prospera for cellular rejection at 1.0% threshold. NPV (negative predictive value); PPV (positive predictive value).

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>AlloSure</th>
<th>Prospera</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity [95% Exact CI]</td>
<td>0.4444 [0.1370, 0.7880]</td>
<td>0.4444 [0.1370, 0.7880]</td>
</tr>
<tr>
<td>Specificity [95% Exact CI]</td>
<td>0.8000 [0.5191, 0.9567]</td>
<td>0.6000 [0.3229, 0.8366]</td>
</tr>
<tr>
<td>PPV [95% Exact CI]</td>
<td>0.5714 [0.1841, 0.9010]</td>
<td>0.4000 [0.1216, 0.7376]</td>
</tr>
<tr>
<td>NPV [95% Exact CI]</td>
<td>0.7059 [0.4404, 0.8969]</td>
<td>0.6429 [0.3514, 0.8724]</td>
</tr>
</tbody>
</table>
Table 4. Side-by-side performance characteristics between AlloSure and Prospera – Area Under the Receiver Operating Characteristic Curve (AUC ROC)

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>AlloSure [95% CI]</th>
<th>Prospera [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC ROC [95% CI]</td>
<td>0.7343 [0.5215, 0.9471]</td>
<td>0.7483 [0.5429, 0.9536]</td>
</tr>
</tbody>
</table>
Figure 1. Workflow of patients investigated with AlloSure and Prospera under surveillance or for-cause dd-cfDNA testing.
Figure 2. Three-part panel showing comparison of the difference in dd-cfDNA % across all sample paired dd-cfDNA percentages stratified by result (A) 0% to 1% (B) 0% to 20%, and (C) biopsy occurrence (excludes 1 sample above the limit of quantification)
Figure 3. Two-part panel showing the turnaround time and cumulative distribution of the Difference (AlloSure – Prospera) in Turnaround Time (TAT) between the assays.