Can Behavioral Research Improve Transplant Decision-Making? A Mock Offer Study on the Role of Kidney Procurement Biopsies

Darren Stewart¹, Brian Shepard¹, John Rosendale¹, Harrison McGehee¹, Isaac Hall², Gaurav Gupta³, Kunam Reddy⁴, Bertram Kasiske⁵, Kenneth Andreoni⁶, David Klassen¹

¹UNOS, Richmond, VA, ²Univ of Utah, Salt Lake City, UT, ³VCU, Richmond, VA, ⁴Mayo Clinic, Phoenix, AZ, ⁵HCMC, Minneapolis, MN, ⁶Univ of Florida, Gainesville, FL.

Corresponding Author

Darren Stewart
United Network for Organ Sharing
Organ Procurement and Transplantation Network
700 N. 4th St
Richmond, Virginia 23219
United States
Darren.Stewart@unos.org
Can Behavioral Research Improve Transplant Decision-Making? A Mock Offer Study on the Role of Kidney Procurement Biopsies

**Cohort**
- 68 kidney transplant surgeons and nephrologists
- 80% "routinely involved" in offer acceptance decisions
- 18 different kidney programs represented

**Intervention**
Each received the same 20 hypothetical kidney offers

- 4 low serum creatinine (<1.5 mg/dL) donors
  - Experimental variants:
    - "Good biopsy"
    - "Poor biopsy"

- 4 AKI donors
  - Experimental variants:
    - "Good biopsy"
    - "Poor biopsy"

**Analysis of 1300 offer responses revealed...**

- Low serum creatinine (<1.5 mg/dL)
  - "Good" biopsy: 3.07 (CI 1.88-5.02)
  - No biopsy: 2.53 (CI 1.56-4.09)

- Acute Kidney Injury (AKI)
  - "Good" biopsy: 3.67 (CI 2.47-5.46)
  - No biopsy: Reference

**Conclusions**
The current use of biopsies in the US appears skewed toward inducing kidney discard. Offer simulation studies are a viable research tool for understanding decision-making and identifying ways to improve the transplant system.

Abstract

Background: The use of procurement biopsies for assessing kidney quality has been implicated as a driver of the nearly 20% US kidney discard rate. Yet in some contexts, biopsies may boost clinical confidence, enabling acceptance of kidneys that would otherwise be discarded. We leveraged a novel organ offer simulation platform to conduct a controlled experiment isolating biopsy effects on offer acceptance decisions.

Methods: Between November 26 and December 14, 2018, 41 kidney transplant surgeons and 27 transplant nephrologists each received the same 20 hypothetical kidney offers using a cross-over design with weekend “washout” periods. Mini-study 1 included four, low serum creatinine (<1.5 mg/dL) donor offers with arguably “poor” biopsy findings that were based on real offers that were accepted with successful 3-year recipient outcome. For each of the four offers, two experimental variants – no biopsy and “good” biopsy – were also sent. Mini-study 2 included four, acute kidney injury offers with no biopsy, each having an offer variant with “good” biopsy findings.

Results: Among low serum creatinine donor offers, we found approximately 3-fold higher odds of acceptance when arguably “poor” biopsy findings were hidden or replaced with “good” biopsy findings. Among acute kidney injury donor offers, we found nearly 4-fold higher odds of acceptance with “good” biopsy findings compared to no biopsy. Biopsy information had profound but variable effects on decision-making: more participants appeared to have been influenced by biopsies to rule-out, versus rule-in, transplantable kidneys.

Conclusion: The current use of biopsies in the US appears skewed toward inducing kidney discard. Several areas for improvement, including reducing variation in offer acceptance decisions and more accurate interpretation of findings, have the potential to make better use of scarce, donated organs. Offer simulation studies are a viable research tool for understanding decision-making and identifying ways to improve the transplant system.

Introduction

Whether to accept or refuse a particular deceased donor organ being offered to a specific transplant candidate at a certain point in time has been described as one of the most inherently complex tasks in medical decision-making (1). Since organ viability for transplantation begins to deteriorate immediately after recovery (2), the time pressure involved in this decision only amplifies the challenge faced by transplant surgeons and others involved in the process. Furthermore, decisions must be made at unpredictable times – including nights and weekends – whenever a donated organ becomes available.

Transplant hospitals currently receive organ offers through DonorNet®, an online information-sharing system developed in 2007 and maintained by the United Network for Organ Sharing (UNOS). To support decision-making, with each offer DonorNet displays hundreds of clinical data elements about recipients, donors, and organs. System efficiency in the DonorNet era has been challenged (4, 5), and both organ offer refusal and discard rates remain high, particularly for kidneys (6-10).
Insights from behavioral science reveal that the precise way complex information is presented can affect decision-making, through psychological phenomena such as priming, loss aversion, observational learning, and default effects (11-15). To enable scientific study of these and other phenomena such as cognitive burden (“information overload”) (16, 17) and labeling effects (18-20), in 2016 under its UNOS LabsSM initiative (21), UNOS developed a DonorNet simulator, SimUNet, that sends hypothetical kidney offers to participating clinicians and receives their acceptance and refusal decisions for analysis. SimUNet was designed to test a broad variety of potential system changes, including the addition of new data; rearranging of data; removal of data; manipulation of data; and other user interface changes hypothesized to improve decision-making. UNOS conducted its first controlled experiment with SimUNet in 2017 (22).

Information that can play a role in kidney acceptance decision-making includes whether a procurement biopsy was performed, the expertise of the pathologist interpreting the biopsy, as well as biopsy findings themselves. The practice of declining an otherwise clinically acceptable kidney because of arguably “poor” biopsy findings (e.g., glomerulosclerosis>20%) has been called into question, as the evidence for an association between some biopsy results and graft outcomes is suspect (23). Due to the unmistakable link between biopsy findings and kidney discard (24, 25), some have proposed eliminating the routine practice of biopsying kidneys, citing successful kidney transplantation in Europe without this practice (26). Others insist that the biopsy findings are a vital element of their decision-making process, not only to potentially rule out kidneys for transplantation, but also to rule them in (27).

In light of the ongoing controversy in the US over the proper role of procurement biopsies in kidney utilization decisions, we employed SimUNet to conduct a controlled experiment to quantify the effects of biopsy availability and biopsy findings on transplant decision-making.

Methods

Study participants

Based on previously expressed interest during an earlier SimUNet study, leadership from twenty-five kidney transplant programs were recruited to participate in this study, with eighteen agreeing to participate. Participating centers were skewed toward larger transplant volumes: nine (50%) of the participating centers were in the top tertile (70+ transplants) among all US kidney programs in terms of deceased donor kidney transplant volume in 2018; eight (44%) were in the middle tertile (26-69 transplants); one (6%) was a small program, having performed just 16 deceased donor kidney transplants.

From within these eighteen centers, a total of 41 kidney transplant surgeons and 27 transplant nephrologists having authority at their program to make offer acceptance or refusal decisions participated in the study. Nearly 80% (54 of 68) indicated being “routinely” involved in organ offer acceptance decisions at their center. One center had just one participant, two centers each had eight participants, and the remaining centers had between two and six participants. Participants ranged in experience from one to 40 years, with a mean of thirteen years in transplant practice.

SimUNet
We used UNOS’s DonorNet simulator, SimUNet (illustrative screenshots shown in Figure 1), a SQL database application with an internet-based user interface, to conduct the study. SimUNet contains hundreds of data elements describing each hypothetical kidney offer (donor demographics, medical history, serial lab values, serologies, renal anatomy; potential recipient age, demographics, Calculated Panel Reactive Antibodies (CPRA), time on dialysis, etc.) — essentially the same clinical information available in DonorNet.

SimUNet study participants receive hypothetical offers by email, review clinical information electronically, and respond by submitting either an acceptance or refusal decision for each offer. Each offer represents a specific donor kidney for a specific potential recipient. Participants also indicate their confidence that their decision would be the same in an actual clinical setting using a 5-point Likert scale (1=Not at all confident,…, 5=Perfectly confident). For refused offers, participants must select a refusal reason and also have the opportunity to indicate whether they believe the organ is “not suitable for transplant into any patient, under virtually any circumstance.” Participants also have the option of providing a narrative explaining their decision-making process.

Offer characteristics and study design

We hypothesized that the presence of a biopsy and biopsy findings would affect the likelihood of offer acceptance, both to rule-out kidneys ostensibly of transplant quality as well as to rule-in kidneys that might otherwise be discarded. In turn, this study consisted of two mini-studies, as depicted in Table 1.

Table 1 shows all 20 offers used in the study. Mini-study 1 consisted of four different low creatinine donors (donors A-D). Mini-study 2 consisted of four different acute kidney injury (AKI) donors (donors E-H). Table 1 illustrates that each donor in mini-study 1 had three experimental offer variants, for which all clinical factors were identical except for the biopsy findings, which varied between “poor,” “no biopsy,” and “good biopsy.” Each donor in mini-study 2 had two experimental offer variants: “no biopsy” and “good biopsy.” Additional clinical details about each study offer can be found in Supplementary Table 1.

Each participating clinician received all 20 offers at various times during the three week period, with experimental variants for each donor being separated by at least a weekend. Table 1 shows the total number of offers sent, responses received, and overall offer acceptance rates in the study.

The two parts of this study are described in more detail below.

Mini-study 1 (low creatinine donors)

To measure the influence of biopsy findings for ruling-out seemingly transplantable kidneys, we chose four kidney donors (labeled herein “A” through “D”) with arguably “poor” biopsy findings (glomerulosclerosis (GS) ≥ 15%; or, moderate or severe interstitial fibrosis (IF) or vascular changes (VC)) but meeting the following criteria:

- Initial, peak, and recent serum creatinine < 1.5 mg/dL
- Donor age < 65 years
- Kidney Donor Profile Index (KDPI) < 85%
- No anatomical irregularities or surgical damage noted
- Hepatitis C (HCV) antibody (Ab) negative
- HCV nucleic acid test (NAT) negative

1/6/2020
Each of these four “poor” biopsy kidney offers was based on an actual offer sent through DonorNet for a specific candidate that was accepted, transplanted with less than 30 hours of cold ischemic time, and had recipient survival of at least three years with a functioning graft and 3-year eGFR>$30 \text{mL/min/1.73m}^2$.

For each of the four “poor” biopsy offers, two experimental variant offers were created by manipulating biopsy data, as follows:

- No biopsy (biopsy reported as “Not done”)
- “Good” biopsy (biopsy findings manipulated by lowering %GS or setting IF and VC to absent or minimal)

The low creatinine mini-study consisted of a total of 12 offers: 4 donors x 3 variants (“good” biopsy, no biopsy, “poor” biopsy). All three offers for each of the four donors were clinically identical other than biopsy information and randomly generated recipient candidate names.

**Mini-study 2 (AKI donors)**

To measure the influence of biopsy findings for ruling-in kidneys from AKI donors, we chose four kidney donors (labeled “E” through “H”) who met the following criteria:

- Initial serum creatinine increased from below 1.5 to above 2.5 mg/dL (terminal)
- Donor age < 70 years
- KDPI < 95%
- No anatomical irregularities or surgical damage noted
- HCV Ab negative
- HCV NAT negative
- HIV negative

For each AKI donor, two experimental variant offers were generated: one with no biopsy, the other with “good” biopsy findings, as defined by GS<$10\%$, and both IF and VC minimal or absent.

Two of the “no biopsy” AKI kidney offers were based on real offers that were refused and the kidney discarded. Two of the “good” biopsy AKI kidney offers were based on real offers that were transplanted with recipient having a functioning graft at 3+ years with eGFR>$30 \text{mL/min/1.37m}^2$. The AKI mini-study consisted of a total of 8 offers: 4 donors x 2 variants (no biopsy, “good” biopsy). Both offers for each of the four donors were clinically identical other than biopsy information and randomly generated recipient candidate names.

Data were de-identified by (a) removing free-text narratives that could contain person-identifying information, and (b) replacing actual candidate names with randomly generated, ethnicity-consistent names (28).

Participants knew the offers were hypothetical (not actual offers for one of their patients) but were not made aware of study goals (i.e., focus on biopsy findings) or design in advance. The study received IRB approval through Advarra (Columbia, MD, USA).

**Study design (offer timing)**
Each study participant received each of the 20 offers over the three-week period November 26 – December 14, 2018. The 12 low creatinine offers were sent via a 3-period (weeks 1, 2, and 3), 6-sequence crossover design, and the 8 AKI offers were sent via a 2-period (weeks 1 and 3), 2-sequence crossover design (29). Participants were randomized independently to one of the 6 low creatinine and 2 AKI sequences. Weekends served as “washout periods” to minimize participants’ memory of specific offers by the time a variant offer was received. Offers were sent on weekdays at random times of day (8am-5pm Monday-Thursday; 8am-12pm on Fridays) in participants’ local time zone. Participants were encouraged but not required to respond within one hour of receiving the offer and had until Sunday at midnight of each week to respond.

Statistical methods

Donor-specific p-values shown in the figures were derived using McNemar’s test (mcnemar.test function in R’s stats package). Hierarchical, mixed effects logistic regression modeling was employed (PROC GLIMMIX in SAS/STAT 12.1) to estimate apriori hypothesized fixed and random effects (donor, center, and participant nested within center) potentially associated with odds of offer acceptance. Fixed effects inference was based on the Kenward-Rogers degrees of freedom approximation; random effect confidence intervals were derived via the profile likelihood method. To assess the influence of biopsies at an individual doctor level, we used a mixed-effect logistic regression model with random participant by treatment interaction effects to estimate each participant’s odds ratio for accepting low creatinine and AKI kidneys based on biopsy findings. These odds ratios are empirical Bayes estimates that leverage Bayesian shrinkage to improve reliability of subject-specific estimates (30).

Data manipulation and table and figure creation were conducted in R version 3.5.

The clinical and research activities being reported are consistent with the Principles of the Declaration of Istanbul as outlined in the ‘Declaration of Istanbul on Organ Trafficking and Transplant Tourism’.

Results

A total of 1,360 simulated offers were sent during the three week study period. Participants submitted responses to 1,300 (96%) of offers, with 55 of 68 doctors responding to all 20 offers. Just over half (669, 52%) of responses were acceptances. For 19% of responses, participants indicated they were “perfectly confident” that they would have made the same decision in an actual clinical setting; 61% were either “very confident” or “confident”; 18% were “somewhat confident” and just 2% were “not at all confident.”

Table 1 shows the offer acceptance rates for each of the eight donors. For the four low serum creatinine (<1.5 mg/dL) donors (Figure 2, left side), acceptance rates rose dramatically for offer variants with experimentally manipulated, ‘good’ biopsy findings compared to the offers having arguably “poor” biopsy findings. Effects were highly statistically significant. For three of the four donors, acceptance rates rose sharply when biopsy findings were hidden (reported as “not done”) compared to the “poor” biopsy offers.

“Organ biopsy results” was submitted as the refusal reason for 97 (64%) of the 152 “poor” biopsy offer refusals, including all 34 (100%) of the Donor A (GS 29%) refusals. By contrast, “organ biopsy results” was cited for just 38% of Donor D (mild-moderate IF) “poor” biopsy offer refusals. For the Donor D
variant with IF displayed as “absent,” zero offers were refused due to biopsy results. By contrast, despite “good” biopsy findings (A: GS 6%, B: GS 3%, C: minimal VC), 28 (42%) of 67 refusals for the other three low creatinine donors were cited as due to “organ biopsy results.”

In 19 cases (6%) of the 152 low-creatinine donor refusals, the respondent indicated the organ was “not suitable for transplant in any patient” -- 16 of the 19 cases were offers with arguably “poor” biopsy findings.

As an example of the apparent influence of biopsy findings in the context of a low creatinine donor offer, one participant confidently accepted the “no biopsy” Donor A offer, explaining “reasonable donor kidney, even with the cyst, for this preemptive recipient.” However, when the otherwise clinically-identical offer was displayed with glomerulosclerosis of 29%, the participant confidently refused the offer, citing “organ biopsy results,” and explaining “5/17 glomerulosclerosis is high but could be sampling error, esp given donor Cr is only 1.0 and no HTN/DM. Nevertheless, hard to accept a donor kidney with that much glomerulosclerosis."

When biopsy findings were experimentally manipulated for AKI kidneys from “not done” to “good,” acceptance rates more than doubled for donors E, F, and G and rose by over 50% for donor H. Effects were highly statistically significant. (Figure 2, right side)

For the non-biopsied AKI donor variants, 120 (60%) of the 199 refusals were reported due to “organ function (creatinine, ejection fraction, pump parameters, etc.).” Of the 19 (10%) refusals due to “other, specify,” 17 (89%) explicitly stated needing biopsy results as their refusal reason.

Also among non-biopsied AKI kidney offers, 48 (19%) responses indicated that the kidney was “not suitable for transplant in any patient,” compared to 35 (14%) of “good” biopsy AKI donor kidney offers.

As an example of the apparent influence of biopsy findings in the context of an AKI donor offer, one participant very confidently refused the “no biopsy” Donor E offer. However, when the otherwise clinically-identical offer was displayed with a “good” biopsy (GS of 2%), the participant confidently accepted the offer, explaining "although donor was on HD due to ATN, overall the biopsy is encouraging, and this recip should do well with organ (although she already has 6.2 yrs of wait time)."

Figure 3 illustrates nearly-identical relationships between biopsy findings and acceptance rates for surgeons compared to nephrologists.

Statistical modeling quantified the odds of acceptance with “good” biopsy findings as being 3.07-fold greater (p<0.0001) than with arguably “poor” biopsy findings among low creatinine donor kidneys (Table 2). Similarly, estimated odds of acceptance were 2.53-fold greater (p=0.0002) with no biopsy versus arguably “poor” biopsy findings. Among the four AKI donors in the study, odds of acceptance were 3.67-fold greater (p<0.0001) with a “good” biopsy compared to no biopsy.

Though nephrologists had an estimated 28% higher odds of acceptance than surgeons, this finding was not statistically significant (p=0.25) (Table 2). More experienced doctors tended to accept more often (22% increased odds per 10 years), but this result was of questionable statistical significance (p=0.06). Similarly, though not statistically significant (p=0.23), a possible time-of-day effect emerged, with the greatest odds of acceptance in the early morning (8-10am) and lower odds in the late afternoon or night.
Among random effects, variation in acceptance rates was most influenced by the donor (variance of 0.91, p<0.0001). Notably, the variance component for participants (0.30, p=0.0001) was three-fold greater than for centers (0.10), for which the variance was not statistically different from zero (p=0.17). (Table 2) This large doctor-to-doctor variability, even for doctors at the same center, is illustrated in Figure 4. Among doctors responding to all 20 offers, acceptance rates ranged from 25% to 100%.

The effect of biopsy findings varied among individual participants, in the context of both low creatinine (participant by treatment interaction variance component, p=0.03) and AKI (p=0.03) donor kidney offers. Approximately 12% of participants appeared to have been highly influenced by biopsies both to rule-out low creatinine donors as well as to rule-in AKI donors. Among participants highly influenced primarily in one direction, more (31%) appear to have been influenced by biopsy findings for ruling-out low creatinine donors, as compared to ruling-in AKI donors (15%). About 43% of participants appear to have been uninfluenced or only moderately influenced by biopsy findings (Figure 5).

Statistical modeling results were largely unchanged in a sensitivity analysis that excluded those participants (N=14, 21%) who acknowledged only being “occasionally” or “rarely” involved in offer decision-making.

The four “low creatinine” donors selected for the study were based on actual deceased kidney donor offers, as previously described. Among all deceased donor kidneys recovered for transplant in 2018 in the US, 730 (3.7%) met our low creatinine donor criteria and underwent procurement biopsy with GS ≥ 15%, IF reported as “moderate” or “severe,” or VC reported as “moderate” or “severe.”

Among these actual procured kidneys, 365 (50.0%) were transplanted and 365 (50.0%) were discarded. Of those discarded, 147 (40.3%) had “list exhausted / no recipient located” as the recorded discard reason and 137 (37.5%) were recorded as “biopsy findings.”

In total, 979 (26.1%) of the 3,753 kidneys recovered in 2018 with the intent to transplant but ultimately not utilized were reportedly discarded due to biopsy findings.

With regard to AKI donors, among all deceased donor kidneys recovered for transplant in 2018 in the US, 997 (5.1%) met our AKI study criteria.

Among these actual procured kidneys, 101 (10.1%) were not biopsied, of which 65 (64.4%) were transplanted and 36 (35.6%) were discarded. The most common reason cited for discard was “diseased organ.” Among the 896 AKI kidneys that were biopsied, 489 (54.6%) had GS<10% with both IF and VC were reported as absent or minimal.

Discussion

In a novel, randomized controlled experiment to study the effect of clinical parameters on transplant decision-making in a “laboratory” environment, we found that each of the three central biopsy parameters used to assess chronic kidney damage – glomerulosclerosis, interstitial fibrosis, and chronic vascular changes – independently had a profound influence on offer acceptance decisions.

Despite the low creatinine, arguably “poor” kidney offers used in the study having been accepted and transplanted in real life with successful 3+ year outcomes (as of 10/2/2019, all recipients are still alive
with a functioning graft according to the OPTN database), the odds of acceptance was approximately three times lower compared to no biopsy or “good” biopsy versions of the same kidney offers. In 16 responses to these kidney offers, participants indicated the organ was “not suitable for any patient, under virtually any circumstance,” suggesting belief that such organs should have been discarded rather than transplanted. These results bolster the claim that transplant-quality kidneys are being declined, and possibly discarded, due to the use of procurement biopsies for assessing organ quality (24, 31).

Among actual kidneys recovered for transplant in 2018 that were similar to our study’s four “low creatinine, poor biopsy” kidneys, half (365) were discarded, many explicitly due to biopsy findings and many others due to inability to find an accepting center/candidate (for which biopsy findings could be a contributing factor). In total, nearly 1,000 kidneys were discarded in 2018 with the reason for discard reported as “biopsy findings.”

Conversely, we found that in the clinical context of a kidney donor with rising serum creatinine reflective of AKI, the presence of a “good” biopsy is associated with a nearly four-fold increased odds of acceptance compared to no biopsy. However, we found that fewer participants relied on biopsy findings to rule-in high risk of discard kidneys compared to those that seemingly used biopsies to rule-out kidneys ostensibly of transplant quality. Clearly, clinical scenarios exist in which biopsies can boost confidence for clinicians hesitant to accept a less-than-ideal but transplantable kidney (27, 32). Among kidneys recovered for transplant in 2018 from donors with AKI as defined in our study, we found 36 were not biopsied and discarded, suggesting dozens of kidneys with high creatinine might be salvaged each year had biopsies been performed and revealed absence of chronic renal damage.

Though biopsies have the potential to both rule-in and rule-out kidneys for transplant, the current use of biopsies in the US appears skewed toward inducing kidney discard. But given the longstanding and widespread reliance on procurement biopsies for organ quality evaluation, it seems unlikely that US procurement and transplant practice will diametrically change to entirely avoid the use of biopsies for acceptance decisions. Since biopsies clearly have the potential to rule-in transplant-quality kidneys that might otherwise be discarded, a more realistic future state worth pursuing might include the following:

- Improving and standardizing biopsy sampling and preparation techniques (33)
- Improving and standardizing biopsy data reporting and interpretation (34)
- Improved understanding of the associations between biopsy findings, donor demographic and clinical factors, and post-transplant outcomes to better inform decision-making (35)
- Reducing the substantial variation across Organ Procurement Organizations (OPOs) in choosing which kidneys to routinely biopsy (24)
- Narrowing the routine use of biopsies to AKI, high KDPI, and other kidney donors in which biopsies could be used to help rule-in and avoid discard (36)
- Risk-adjustment for biopsy findings in post-transplant survival metrics to reduce center risk aversion, a change forthcoming to program specific reports (PSRs) in 2020 (37)
- Use of automation (e.g., machine/deep learning) (38-42) to avoid human subjectivity and bias (43) in interpreting biopsy images

In addition to the influence of biopsies, we found substantial heterogeneity in decision-making, even among doctors within the same center, a finding echoed in a prior offer simulation study (22). In both studies, the doctor-to-doctor variance in acceptance rates was three-fold greater than the statistically insignificant center effect. This suggests that transplant clinicians at the same center may not be making decisions grounded in the same overarching philosophy. It also highlights the potential value in the
adoption of well-vetted, understandable, and trusted clinical decision-making tools (44-47) that boost surgeon self-efficacy (48) and support the patients’ role in decisions (49). While some degree of heterogeneity in acceptance decisions is appropriate, for example since not all surgeons are equally experienced in transplanting the same types of less-than-ideal kidneys, exceedingly high doctor-to-doctor variation is likely associated with suboptimal decision-making (e.g., declining organs that would yield the candidate a survival benefit (50); elevated odds of waitlist death associated with offer declines (6)) and may contribute to inequities in transplant access (7).

Analysis of simulated organ offers has limitations. Though results from a “laboratory” environment study of acceptance behaviors can be insightful and help point toward possible system and practice improvements, the decision to accept a hypothetical offer may differ from the decision that would be made in actual clinical practice. In fact, we have found that simulation study acceptance rates are substantially higher than real offer acceptance rates. This may be because a study using hypothetical offers may have difficulty fully reflecting the influence of exogenous factors such as transplant center finances (51), surgeon scheduling and availability issues (including possible weekday vs. weekend effects (52)), concerns about center performance monitoring (53), and involvement of other staff (e.g., OPO and transplant center coordinators) and patients in the decision-making workflow.

Thus, a key assumption in leveraging the findings from such studies is that laboratory-estimated effects would manifest as similar effects (in terms of relative magnitude or at least direction) on actual offer acceptance rates, albeit relative to a much lower baseline level of acceptance. This assumption could be empirically validated through controlled experiments on real offers, where feasible, and also complemented by rigorous studies of actual offer acceptance patterns (8).

The strengths of this study include the ability to isolate the effect of biopsy findings in a controlled setting, thus avoiding concerns about unmeasured covariate confounding extant in the study of real kidney offers. Importantly, offer simulation studies also allow for center and person-level analyses, whereas actual offer data are currently limited to the former. Our study benefited from a large number of participants (68) and an exceptionally high offer response rate of 96%. Finally, by using clinical data from real offers as the basis for simulated offers, we are able to link laboratory-related acceptance patterns with the hindsight of knowing the outcomes for the real transplant recipients. Of course, the offers with successful three-year outcome selected for the study are not necessarily representative of all low creatinine, “poor biopsy” transplants, some of which may have had a poor outcome, possibly associated with histology.

This study has demonstrated the power of offer simulation research to isolate factors that influence decision-making and identify highly significant effects. Not only can such studies help highlight key areas of clinical practice that deserve increased scrutiny and improvement, such as the role of procurement biopsies, but they can also help guide the implementation of user-interface enhancements, for example to DonorNet. Offer simulation also has the potential to be used as a quality improvement tool for centers to understand and reduce heterogeneity in decision-making among those involved in fielding organ offers. Other anticipated directions of this work include expanding to organs beyond kidney, involving patients and others involved in decision-making, linking simulated offers more closely with long-term recipient outcomes, and supporting educational initiatives for surgeons and multidisciplinary transplant teams (Figure 6).
Disclosures
The authors have nothing to disclose.

Acknowledgements
Executing this study successfully would not have been possible without vital input from those who provided clinical, behavioral science, and experimental design guidance on our early SimUNet studies, including Lloyd Ratner, Carl Berg, Robert Stratta, Kristin MeKeel, Richard Formica, Jan Willem Lindemans, Megan Cline, Judd Kessler, Colin Sullivan, and Jesse Schold. UNOS Labs work would not be possible without UNOS Research and IT leadership from Ryan Ehrensberger, Bob Carrico, Alex Tulchinsky and Marty Wilson. The UNOS IT core SimUNet development and study execution team has consisted of Michael Ghaffari, Ben Howard, Josh Steward, Richard Shipmon, and Matt Stevens. Former UNOS data scientist Subhash Jaini’s early programming work and UNOS Research Analyst Amanda Robinson’s statistical power analyses are also noteworthy contributions.

Author Contributions
Darren Stewart: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Supervision; Validation; Visualization; Writing - original draft; Writing - review and editing
Brian Shepard: Conceptualization; Funding acquisition; Resources; Writing - review and editing
John Rosendale: Investigation; Methodology; Writing - review and editing
Harrison McGehee: Data curation; Investigation; Visualization; Writing - review and editing
Isaac Hall: Investigation; Methodology; Writing - review and editing
Gaurav Gupta: Investigation; Methodology; Writing - review and editing
Kunam Reddy: Investigation; Methodology; Writing - review and editing
Bertram Kasiske: Investigation; Methodology; Writing - review and editing
Kenneth Andreoni: Investigation; Methodology; Writing - review and editing
David Klassen: Investigation; Methodology; Supervision; Writing - review and editing
References


44. Wey A, Salkowski N, Kremers WK, Schaffhausen CR, Kasiske BL, Israni AK, et al. A kidney offer acceptance decision tool to inform the decision to accept an offer or wait for a better kidney. American


<table>
<thead>
<tr>
<th>Mini-study</th>
<th>Donor type</th>
<th>Donor ID</th>
<th>Donor age</th>
<th>KDPI</th>
<th>Experimental offer variant</th>
<th>Chronic biopsy findings displayed in SimUNet</th>
<th>Total Offers, Responses, and Acceptance Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Glomerulosclerosis</td>
<td>Interstitial Fibrosis</td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>54</td>
<td>76%</td>
<td>'Poor' biopsy</td>
<td>19%</td>
<td>Mild</td>
<td>Mild</td>
<td>68</td>
</tr>
<tr>
<td>B</td>
<td>54</td>
<td>76%</td>
<td>No biopsy</td>
<td>6%</td>
<td>Mild</td>
<td>Mild</td>
<td>68</td>
</tr>
<tr>
<td>C</td>
<td>64</td>
<td>83%</td>
<td>'Good' biopsy</td>
<td>3%</td>
<td>Mild</td>
<td>Mild</td>
<td>68</td>
</tr>
<tr>
<td>D</td>
<td>47</td>
<td>80%</td>
<td>'Poor' biopsy</td>
<td>5%</td>
<td>Mild</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>D</td>
<td>47</td>
<td>80%</td>
<td>'Good' biopsy</td>
<td>5%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>E</td>
<td>59</td>
<td>77%</td>
<td>No biopsy</td>
<td>29%</td>
<td>Mild</td>
<td>Mild</td>
<td>68</td>
</tr>
<tr>
<td>E</td>
<td>59</td>
<td>77%</td>
<td>'Good' biopsy</td>
<td>2%</td>
<td>Mild</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>F</td>
<td>32</td>
<td>47%</td>
<td>No biopsy</td>
<td>6%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>F</td>
<td>32</td>
<td>47%</td>
<td>'Good' biopsy</td>
<td>3%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>G</td>
<td>57</td>
<td>91%</td>
<td>No biopsy</td>
<td>6%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>G</td>
<td>57</td>
<td>91%</td>
<td>'Good' biopsy</td>
<td>6%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>H</td>
<td>35</td>
<td>51%</td>
<td>No biopsy</td>
<td>6%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
<tr>
<td>H</td>
<td>35</td>
<td>51%</td>
<td>'Good' biopsy</td>
<td>6%</td>
<td>Absent</td>
<td>Absent</td>
<td>68</td>
</tr>
</tbody>
</table>

**Table 1.** Experimental Design Framework Showing the 20 Study Offers and Summary Results

1/6/2020
Table 2: Statistical Modeling Results for Odds of Offer Acceptance

<table>
<thead>
<tr>
<th>Kidney donor type</th>
<th>Fixed Effects</th>
<th>Odds Ratio (OR)</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creatinine &lt; 1.5 mg/dL</td>
<td>No biopsy (ref: 'poor' biopsy)</td>
<td>2.53</td>
<td>1.56 - 4.09</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>'Good' biopsy (ref: 'poor' biopsy)</td>
<td>3.07</td>
<td>1.88 - 5.02</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>AKI (rising Cr &gt; 4.0 mg/dL)</td>
<td>'Good' biopsy (ref: no biopsy)</td>
<td>3.67</td>
<td>2.47 - 5.46</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Nephrologist (ref: surgeon)</td>
<td>1.28</td>
<td>0.83 - 1.97</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Tenure (per 10 years)</td>
<td>1.22</td>
<td>0.99 - 1.51</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Response day/time</td>
<td>1.22</td>
<td>0.99 - 1.51</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Weekday early morning (ref: weekend)</td>
<td>1.70</td>
<td>0.96 - 3.01</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Weekday late morning (ref: weekend)</td>
<td>1.53</td>
<td>0.88 - 2.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weekday early afternoon (ref: weekend)</td>
<td>1.28</td>
<td>0.77 - 2.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weekday late afternoon or night (ref: weekend)</td>
<td>1.13</td>
<td>0.69 - 1.85</td>
<td></td>
</tr>
</tbody>
</table>

Variance

<table>
<thead>
<tr>
<th>Random effects</th>
<th>Component (VC)</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transplant center (n=18)</td>
<td>0.10</td>
<td>0.00 - 0.51</td>
<td>0.17</td>
</tr>
<tr>
<td>Participant (n=68) within transplant center</td>
<td>0.30</td>
<td>0.11 - 0.64</td>
<td>0.0001</td>
</tr>
<tr>
<td>Donor (n=8)</td>
<td>0.91</td>
<td>0.37 - 3.01</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1: SimUNet’s “Donor Summary” screen contains basic medical and demographic information such as height, weight, BMI, age, gender, KDPI, and ethnicity, as shown in Panel A. By scrolling down, participants can view an extensive array of additional medical information about the donor, including biopsy findings (Panel B). SimUNet’s “Match Results” screen information about the matched potential recipient and provides participants with the opportunity to respond to each offer (Panel C). Participants can either refuse or accept each offer; must indicate their level of confidence and, for refused offers, select a refusal reason; and have the option of indicating that the organ is not suitable for transplantation.

Figure 2: The left panel shows acceptance rates for the three offer variants (“poor” biopsy, no biopsy, “good” biopsy) for the four low creatinine donors in the study. Compared to the “arguably poor” biopsy offers, acceptance rates rose dramatically when biopsy findings were hidden or replaced with “good” results. The right panel shows that acceptance rates increased sharply for each of the four AKI donors with “good” biopsy results.

Figure 3: The left panel shows acceptance rates for the three offer variants (“poor” biopsy, no biopsy, “good” biopsy), averaged across the four, low creatinine study donors. For both surgeons and nephrologists, compared to the “arguably poor” biopsy offers, acceptance rates rose dramatically when biopsy findings were hidden or replaced with “good” results. The right panel shows that for both surgeons and nephrologists, acceptance rates increased sharply for each of the four AKI donors with “good” biopsy results.

Figure 4: Each data point represents an individual participant’s acceptance rate among the (up to) 20 study offers. Among participants responding to all 20 offers, individual acceptance rates ranged from 25% to 100%. Participants are grouped vertically with other participants affiliated with the same transplant center. Horizontal bars represent transplant center acceptance rates, which ranged from 35% to 80%. One center having only a single participant is not shown.

Figure 5: Mixed effect logistic regression modeling with random participant by treatment interaction effects was used to estimate each participant’s odds ratio for accepting low creatinine and AKI kidneys based on biopsy findings. Among the 68 study participants, 8 (12%) appear to have been highly influenced (OR>1.75) by biopsy findings both to rule-out low creatinine kidney offers as well as to rule-in AKI kidneys. By contrast, 29 (43%) appear to have been uninfluenced or at most moderately influenced by biopsy findings. Also, 21 (31%) participants appear to have been highly influenced by biopsy findings in the context of low creatinine kidneys but not necessarily for AKI kidneys. By comparison, 10 (15%) participants appeared to have been highly influenced by biopsy findings for AKI kidneys, though not necessarily for low creatinine kidneys.

Figure 6: Organ offer behavioral research has a number of potential applications, from evaluating potential system improvements to DonorNet, training surgeons, and facilitating introspective center-level assessments of offer evaluation processes. The present study highlights a number of possible opportunities for improvement in the practice of obtaining and utilizing procurement biopsies in the context of kidney transplantation.
Figure 1. Illustrative Screenshots Showing a Sampling of Clinical Information Displayed in UNOS’s Organ Offer Simulation Tool, SimUNet, and its Decision-Making Interface.

A) Donor Summary:
- **Donor:** Female / 53 years
- **Provider Information:** OPO: Local offer
- **Donor Information:**
  - **Height:** 5 ft 7 in / 172 cm
  - **Weight:** 126 lbs / 57 kg
  - **BMI:** 19.3
  - **Age:** 53 years
  - **Gender:** Female
- **Current KPD:** 81% Graft Survival Rates by KPD
- **Ethnicity/Race:** Black

B) Organ Specific Data:
- **Kidney Laterality:** Left
- **Dual Kidney Transplant Available:** No
- **Biopsy Findings:**
  - **Kidney Biopsied:** Yes (read by local pathologist)
  - **Biopsy Type:**
    - Number of glomeruli visualized: 25
    - Number of glomeruli sclerosed: 2
    - Glomerulosclerosis: 0%
    - Interstitial Fibrosis: None
    - Vascular Changes: Minimal
    - Acute Changes:

C) Offer Response for this specific candidate:
- **Refuse**
- **Accept**
- **Level of confidence in your response:**
  - Select level of confidence
- **Refusal Reason:**
  - 502 Organ Biopsy Results
  - In my opinion, this organ is not suitable for transplantation.

*If possible, please respond within 1 hour. Otherwise, please respond before the final response deadline of 8/20/2020 12:59:59 AM EST.
Figure 2. Association of Offer Acceptance Rates with Experimental Manipulation of Biopsy Findings for the Eight Study Donors

Low Serum Cr (<1.5) Donors

- **Donor A**: Glomerulosclerosis (GS) 29%
- **Donor B**: GS 19%
- **Donor C**: Severe vasc. changes (V/C)
- **Donor D**: Mild-moderate interstitial fibrosis (I/F)

Acute Kidney Injury (rising Cr>2.5) Donors

- **Donor H**: P=0.0003
- **Donor G**: P=0.0009
- **Donor F**: P=0.0008
- **Donor E**: P=0.0003

P-values reflect arguably ‘poor’ vs. ‘good’ biopsy among participants responding to both variants, per McNemar’s test.
Figure 3. Association of Offer Acceptance Rates with Experimental Manipulation of Biopsy Findings, Surgeons vs. Nephrologists
Figure 4. Variation in Offer Acceptance Rates among Participants, Grouped by Transplant Center
Figure 5. Classification of Study Participants Based on Apparent Influence of Biopsies on Decision-Making
Figure 6. Variation in Offer Acceptance Rates among Participants, Grouped by Transplant Center

<table>
<thead>
<tr>
<th>Potential uses of organ offer simulation research</th>
<th>Recommendations for improved biopsy practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Isolate the influence of clinical factors (e.g., biopsy) on decision-making</td>
<td>• Standardize biopsy sampling and preparation</td>
</tr>
<tr>
<td>• Assess the impact of DonorNet® user-interface changes</td>
<td>• Standardize biopsy reporting and interpretation</td>
</tr>
<tr>
<td>• Understand and mitigate so-called “labeling effects” (e.g., PHS increased risk)</td>
<td>• Improved understanding of association between biopsy findings and outcomes</td>
</tr>
<tr>
<td>• Educate surgeons and physicians on fielding organ offers</td>
<td>• Reduce variation in OPO decisions to biopsy kidneys</td>
</tr>
<tr>
<td>• Reduce inter- and intra-doctor variation</td>
<td>• Narrow the routine use of biopsies to maximize opportunities to <em>rule-in</em> kidneys and avoid discards</td>
</tr>
<tr>
<td>• Test the impact of clinical decision support (CDS)</td>
<td>• Add biopsy findings to program specific reports (PSRs) to reduce risk-aversion (in progress)</td>
</tr>
<tr>
<td>• Identify ways to support shared decision-making with patients</td>
<td>• Use of automation/machine learning to avoid human subjectivity in interpreting biopsy images</td>
</tr>
<tr>
<td>• Facilitate center-specific quality assurance and process improvement (QAPI)</td>
<td></td>
</tr>
<tr>
<td>• Conduct multi-center collaborative improvement</td>
<td></td>
</tr>
</tbody>
</table>