A Data Science Approach for Quantifying Spatio-Temporal Effects to Graft Failures in Organ Transplantation

Diego Pinheiro[‡], Farag Hamad[†], Martin Cadeiras[§], Ronaldo Menezes[‡], Nezamoddin Nezamoddini-Kachouie[†]

[‡]BioComplex Laboratory, Department of Computer Science, Florida Institute of Technology
dsilva@biocomplexlab.org, rmenezes@cs.fit.edu

[†]Department of Mathematical Science, Florida Institute of Technology
fhamad2014@my.fit.edu, nezamoddin@fit.edu

[§]Department of Medicine, UCLA
mcadeiras@mednet.ucla.edu

***Abstract*— The transplantation of solid organs is one of the biggest accomplishments of modern medicine and although shortage of organs is a major public health issue, nearly 8,000 people died while waiting for an organ in 2014. Meanwhile, the allocation system currently implemented can lead to organs being discarded and the medical community still investigates factors that affects early graft failure such as distance and ischemic time. In this paper we investigate early graft failure under a spatio-temporal perspective using a data science unified approach for all six organs based on complementary cumulative analysis of both distance and ischemic time. Interestingly, although distance seems to highly affect some organs (e.g. liver), it appears to have no effect on others (e.g. kidney). Similarly, the results on ischemic time confirm it affects early graft failure with higher influence for some organs such as (e.g. heart) and lower influence for others such as (e.g. kidney). This poses the question why allocation policies should be individually designed for each organ in order to account for their particularities as shown in this work.**

I. Introduction

Transplantation of solid organs is a life-saving approach in medicine and in many cases the only viable approach to patients with end-stage organ failure. In 2014, nearly 30,000 solid organ transplants were performed in the USA. Despite this encouraging number, 22 people continue to die every day because they do not have access to an organ [**?**]. Shortage of solid organs (livers, hearts, kidneys, lungs) is well known nevertheless organs are often discarded after being recovery for transplant. In a 2012 article, the New York Times discussed that many organs that could be used to save people are being discarded and they pointed out that 18% of kidneys collected for transplant have been discarded in 2012.

When an organ becomes available from a deceased donor, the allocation policies such as medical urgency, expected benefit and geographical constraints (distance between donor and recipient) are applied to people in the waiting list to select a match. Allocation policies regard the survivability of the organ outside the human body, namely, the *cold ischemic time*, as an important factor since it is associated with the quality degradation of the organ. Besides, the distance is an important factor on these decisions given that the farther the distance from the donor hospital to the transplant center, the worse might be the quality of the organ. However, to our knowledge, no one looked at the effect of distance to the survivability of the organ in a more unified way.

In this paper we investigate the importance that distance really plays in this process. We look at the correlations between distance traveled by organs and graft failure rates. We find that distance seems to not play a major role in the failure rates except for the case of heart and liver. Our findings is a first move towards the argument that allocation policies should be custom made for each organ and that perhaps even the division of the country into Organ Procurement Organizations (OPO) could also be done independently for each organ in order to improve the allocation process. The country division into OPOs and hence the direct consideration of these areas in the allocation process is likely to be stopping possible better allocation of organs such as kidneys where the distance is not so relevant.

The National Organ Transplant Act (NOTA) was passed by Congress in 1984 stating that organs (from the deceased) are to be treated as national resources and their allocation is to be done using a fair process. Factors such as waiting time, medical urgency, histocompatibility match, preservation time, and procurement cost are to be considered when trying to understand the allocation system [**?**]. In order to implement/enforce allocation processes, the US has been divided into 11 administrative UNOS (United Network for Organ Sharing) regions which are further divided into 69 OPOs.

The major criteria used for allocating organs to patients are the severity status levels and geographical location where the organ became available. Organs are allocated to local OPOs based on the severity levels. When the severity status levels within the local OPOs are exhausted (no match is present at the local level), the organs are allocated to the other OPOs at the regional level (same UNOS region), and lastly to the OPOs at the national level (other UNOS regions). This process is poorly understood by the population at large which leads on an unwillingness to become organ donors [**?**].

Recent studies have looked at the effect of geography (distance) in organ allocation. One of the most related to our study is the work of Ghaoui et al. [**?**]. The authors focused on UNOS Region 1 and the identification of disparities using GIS (Geographical Information System) visualizations. The noticed that spatial clustering in the UNOS Region 1 and that the emergence of such clustering can be predicted from the distance to transplant center or density of population.

Massengill [**?**] argues that the distance between organ donor and recipient should be given a lower weight during the allocation process. However, our findings here does not support Massengill's sugestion. We find that while in some organs, the distance is not so relevant, in others it appears to be intrinsically to graft failure rates.

The truth is that the importance of organ allocation has led many scientists into discussing specific allocation strategies for different kind of organs, e.g. livers [**?**], kidneys [**?**], hearts [**?**]. Here we provide a holistic in which the effect of distance is gauged against every solid organ in the UNOS database. The data analysis we performed confirms that ischemic time is a better measure than distance traveled to the probability of graft failure. Moreover, distance is not so relevant to organs whose viability is longer than hours such as kidneys.

This study uses the UNOS transplant data which contains information of every organ transplant performed in the United States ($N = 373870$) related to the six major solid organ transplants: heart (13.21%), intestine (0.52%), lung (5.54%), liver (27.07%), kidney (51.86%) and pancreas (1.81%). This data covers the period from October 1$^{\text{st}}$, 1987 to approximately December 31$^{\text{st}}$, 2010 and considers transplants from deceased donors and excludes transplants from living donors. The selected variables are graft status GS indicating whether the graft failed or not, the distance D between donor hospital and the transplant center, graft failure time T indicating the days from transplant to failure and the ischemic time I in hours.

II. Methods

There are several factors contributing to organ failure after transplantation including acute and chronic rejection, patient disease history and other existing conditions such as hepatitis, and infection. The chronic rejection is long term loss of functionality whereas acute rejection is pretty severe reaction as early as one week after transplantation. There is a real time system that allocates available organs (through donors) to recipients who are ordered in a waiting list. The organ-recipient allocation is based on several clinical and non-clinical factors. An organ might be allocated to a recipient within the same state as the donor or to a recipient in another state. As a result a donated organ may travel from less than 100 miles to thousands of miles for transplantation. Browsing through available data for transplanted organs over past few decades since 1980's reveals that at least 13% of the transplanted organs failed within the first year for kidney and it reaches to almost 29% for intestine. The failure rate within two years of transplantation is at least 18% for kidney and reaches to 38% for intestine, and just 55% of transplanted intestines survived more than three years. To improve organ survival after transplantation, our general goal is to find factors that may contribute to organ failure. To this end, in this paper we perform statistical analysis to test whether organ failure is correlated with the distance that donated organs traveled. The unilateral hypothesis is

$$\begin{cases} H_0 : P_{D>D_0} = P_{D\leq D_0} \\ H_a : P_{D>D_0} > P_{D\leq D_0}, \end{cases} \tag{1}$$

where $P_{D>D_0}$ is the proportion of the failed organs that traveled more than D_0 and $P_{D\leq D_0}$ is the proportion of the failed organs that traveled less than D_0. Spatial conditional cumulative proportion of failed organs is estimated by:

$$P_{D>D_0} = P(T \leq t | D > D_0) = \frac{P(T\leq t, D>D_0)}{P(D>D_0)}, \tag{2}$$

where $P(D > D_0)$ is empirical spatial complementary cumulative distribution.

This hypothesis is tested for different organs separately. We also perform statistical analysis to test whether organ failure is correlated with the ischemic time. The unilateral hypothesis for this test is

$$\begin{cases} H_0 : P_{I>I_0} = P_{I\leq I_0} \\ H_a : P_{I>I_0} > P_{I\leq I_0}, \end{cases} \tag{3}$$

where $P_{I>I_0}$ is the proportion of the failed organs with ischemic higher than I_0 and $P_{I\leq I_0}$ is the proportion of the failed organs with ischemic below I_0. Temporal conditional cumulative proportion of failed organs is estimated by:

$$P_{I>I_0} = P(T \leq t | I > I_0) = \frac{P(T\leq t, I>I_0)}{P(I>I_0)}, \tag{4}$$

where $P(I > I_0)$ is empirical temporal complementary cumulative distribution. All the population proportions were estimated building confidence interval using the wilson score:

$$\tilde{p} \pm z_{\alpha/2} \frac{\sqrt{\hat{p}\hat{q}/n + z^2_{\alpha/2}/4n^2}}{1 + z^2_{\alpha/2}/n}, \; \tilde{p} = \frac{\hat{p} + z^2_{\alpha/2}/2n}{1 + z^2_{\alpha/2}/n} \tag{5}$$

where $\hat{q} = 1 - \hat{p}$, $z_{\alpha/2}$ is the z-score for the significance level $\alpha = 0.05$.

Finally, the spatiotemporal pearson correlation $\hat{\rho}_{(D,I)}$, i.e. correlation between distance and ischemic time is estimated using the distance that each transplanted organ traveled before transplantation vs. its ischemic time.

III. Spatio-Temporal Effect on Graft Failures

The medical community has published works indicating that ischemic time and geographical distances separating donor and recipient lead to increased risks of graft failure [**?**], but at the same time the medical community also argued that in some cases, ischemic time could not account for differences between success in organ transplantation [**?**] and neither to the risk of acute rejection of the organ. E.g a case with lung transplant was published by Fiser et al. [**?**] Given that the issue has not been completely settled we took an approach founded in Data Science to understand

for each organ the effect of distance and ischemic time. For our purposes, ischemic time refers to *cold ischemic time* as this is a larger variability than warm ischemic time.

The first experiments regard spatial information as measured by distance D. Although one could argue that ischemic time has a direct related to distance travel by the organ, this is not the case for all organs. It is generally the case that for organs with shorter viability time (e.g. heart) the ischemic time has a high correlation to the distance traveled ($\rho = 0.6$); these organs are so critical that they cannot be ischemic for a long time, hence most of the ischemic time is due to transportation. In organs such as kidney for instance the correlation is not so high ($\rho = 0.3$) because the organs can be ischemic for sometime in a health facility waiting to be transported. For this reason we also decided to analyze the spatial effect of distance traveled to graph failures.

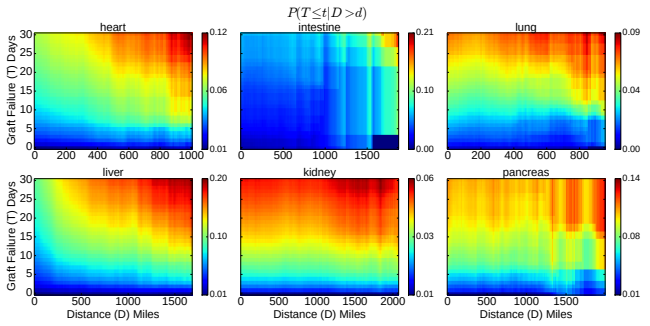

Fig. 1. The spatial effect of distance on early graft failures for the six major transplanted solid organs. The distance D (x-axis) ranges from 0 to 99$^{\text{th}}$ distance percentile $d_{99^{\text{th}}}$ for each organ using 10 miles bins. The graft failure time T in days (y-axis) ranges from 0 (i.e. grafts that failed in the same day they were transplanted) to 30 days. Each cell represents $P(T \leq t|D > d)$, i.e., the proportion of transplants that traveled distance D greater than d miles for which the graft failed in less than or equal to t days divided by the total number of transplants that traveled distance D greater than d miles. The spatial effect appears to be higher on hearts and livers when compared to intestine, kidney and pancreas.

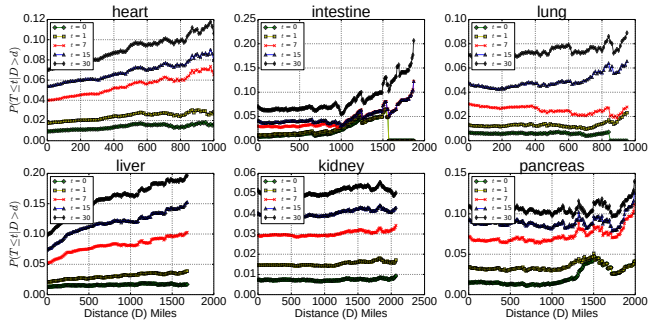

Fig. 2. The spatial effect of distance on early graft failures for the six major transplanted solid organs. The x-axis represents the distance D in miles and the y-axis represents $P(T \leq t|D > d)$ for 5 different values of t {0, 1, 7, 15, 30}. Note that each curve is a chosen row in the heatmap of Figure **??** and also that in the y-axis t is accumulative. For instance, 10.62% (95% CI, 7.91-13.33) of hearts that travel more than 1000 miles failed in less than 1 month when compared to all hearts 7% (95% CI, 6.76-7.24). This spatial effect also impact livers as well but seems to not affect the other organs. Table **??** presents more more information.

The spatial results (Figures **??** and **??**) demonstrate that distance does not seem to be a good predictor of graft failure except for two organs: heart and liver. It is unclear why that is the case but one could argue as before that ischemic time is more relevant to graft failure because it is a more accurate metric of graft failure; maybe a better predictor of such failures.

The Figure **??** shows the influence of ischemic time (I) on early graft failures for each organ. The ischemic time range is different for each tissue due to each different viability. For instance, a heart highly degrades when the its ischemic time cross the 4-5 hour limit.[1]

Our results in Figure **??** reproduce quite well what is known from the literature. More importantly however is that the heatmaps allow us to argue that ischemic time is not relevant until a certain threshold is reached. For instance, if you take heart as an example, the chart seems to indicate that the data does not show any significant increased risk if the organ is not ischemic for more than 2 hours. After that, we see that the probability of failure rapidly increases. The same can be see from other organs. As another example, the threshold in which ischemic time becomes relevant for kidneys seem to be around 10-14 hours; below that time, the ischemia is not so relevant.

The proportion of early graft failures increases with i with a higher influence for some organs (e.g. heart and liver) and with a lower influence for others (e.g. kidney and pancreas). Figure **??** makes it more evident what we can see already in Figure **??**. We took the specific times (t) in each of the heat map to clearly present the effect of ischemic time. In Figure **??** it is also clearer the argument we made about ischemia not being so relevant below a certain threshold for each organ. This is seen because the curves maintain themselves constant until they reach the threshold. For lungs, the point in which the ischemic time start to matter mostly starts at 4 hours.

In Figure **??**, the experiments used the entire UNOS dataset. Note however that the number of intestine and pancreas transplants are significantly less than for the other organs and hence the behavior of the heatmaps and curves cannot be said to be statistically reliable. Yet, they are included here because the UNOS dataset contains six solid organs and we have performed our approach in all six.

IV. Discussions

The understanding of issues that may influence organ failure is crucial to the medical community. Our paper separately looks at the effect of ischemic time and distance traveled by the organs. Due to the fact that the UNOS data is analyzed aggregated one would probably be curious to look at disaggregating the data to understand it further.

According to our data-science experiments, we had two main findings: *(i)* ischemic time appears to be more relevant as a predictor of graft failures than distance traveled by the organ. Moreover, *(ii)* neither variable reported is a

[1] http://www.dcids.org/facts-about-donation/frequently-asked-questions/

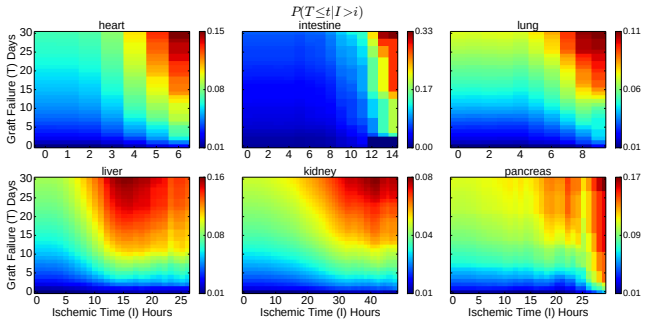

Fig. 3. The temporal effect of ischemic time on early graft failures for the six major transplanted solid organs represented as heatmaps. The x-axis represents the ischemic time I in 1 hour bins ranging from 0 to the 99^{th} ischemic time percentile of each organ $i_{99\text{th}}$. The y-axis represents the graft failure time T in days ranging from 0 (i.e. grafts that failed in the same day they were transplanted) to 30 days. The x-axis might differ for each organ since each graft is not necessarily associated with the same tissue viability. Each cell in the heatmap represents $P(T \leq t | I > i)$, i.e., the proportion of transplants with ischemic time I greater than i hours for which the graft failed in less than or equal to t days divided by the total number of transplants with ischemic time I greater than i.

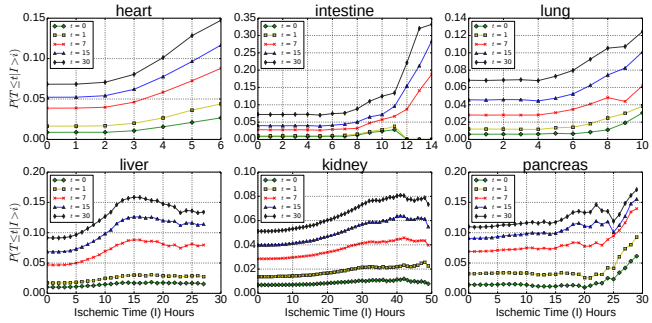

Fig. 4. The temporal effect of ischemic time on early graft failures for the six major transplanted solid organs. The x-axis represents the distance D in miles ranging from 0 to the 99^{th} Distance percentile for each organ $d_{99^{th}}$. The y-axis represents $P(T \leq t | I > i)$ for 5 different values of t {0, 1, 7, 15, 30}. Similarly to the case of D, some organs are more influenced than others by the increase of ischemic time. Note that in the y-axis t is accumulative. What is important to notice in these lines is that they have little change (they do not increase) with the increase of ischemic time on the x-axis until they reach a certain critical point (threshold). The exceptions in these charts are intestine and pancreas for which the amount of transplants in the UNOS dataset is small and hence the results are statistically not reliable. For instance, 14.75% (95% CI, 11.71%-17.77%) of hearts with ischemic time grater than 6 hours failed in less than 1 month when compared to all hearts 6.83% (95% CI, 6.60%-7.06%); 7.68% (95% CI, 7.24%-8.12%) of kidneys with ischemic time greater than 34 hours failed in less than 1 month when compared to all kidneys 5.13% (95% CI, 5.02%-5.23%). Consult Table **??** for more information.

good predictor for organs with high viability times such as pancreas and kidney.

For future work, other external factors not considered in this work need to be assessed such as the age of donor and recipient, contributing causes of graft failure (i.e. infection, acute rejection and chronicle rejection), the donor cause of death as well as whether the patient was non compliant with possible prescribed treatments.

The question whether organ allocation policies need to be individually designed for some organs still seems to remain open. Interestingly, the modest framework built in this work

TABLE I

WILSON SCORE 95% CONFIDENCE INTERVALS FOR THE POPULATION PROPORTION $\hat{P}_{T \leq t | D > d}$ AND $\hat{P}_{T \leq t | I > i}$ FOR $t = 30$ DAYS.

	$\hat{P}_{T \leq 30 \mid D > d}$				$\hat{P}_{T \leq 30 \mid I > i}$			
Graft	d	$N_{D>d}$	$N_{T \leq 30}$	95% CI (%)	i	$N_{I>i}$	$N_{T \leq 30}$	95% CI (%)
heart								
	0	42872	3000	7.00 (6.76, 7.24)	0	46364	3164	6.83 (6.60, 7.06)
	500	3686	346	9.43 (8.49, 10.37)	3	21816	1754	8.05 (7.69, 8.41)
	1000	485	50	10.62 (7.91, 13.33)	6	518	75	14.74 (11.71, 17.77)
liver								
	0	88720	8645	9.75 (9.55, 9.94)	0	90755	8313	9.16 (8.97, 9.35)
	500	10099	1499	14.86 (14.16, 15.55)	9	31043	3717	11.98 (11.62, 12.34)
	1000	3263	534	16.40 (15.14, 17.67)	18	2274	347	15.32 (13.84, 16.80)
	1500	1364	256	18.86 (16.78, 20.93)				
kidney								
	0	161748	8364	5.17 (5.06, 5.28)	0	178257	9135	5.13 (5.02, 5.23)
	500	29932	1482	4.96 (4.71, 5.20)	17	104560	5969	5.71 (5.57, 5.85)
	1000	14036	694	4.96 (4.60, 5.32)	34	13919	1067	7.68 (7.24, 8.12)
	1500	6976	372	5.36 (4.83, 5.88)				
	2000	2568.0	128	5.05 (4.21, 5.90)				

to analyze spatial-temporal aspects impacting graft failure could be further generalized towards an unified framework taking into account other factors as aforementioned. Such unified framework could possibly serve as a tool to assess organ allocation policies as well as to support the design of new policies for different organs.