Liver and Multi-Organ Exchange

John P. Dickerson and **Tuomas Sandholm**
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Kidney exchange, where candidates with inoperable organ failure trade incompatible but willing donors, is a life-saving alternative to the deceased donor waitlist, which has inadequate supply to meet demand. Computer scientists and economists study this problem because it represents an opportunity to field a real barter exchange that improves sustainability of life and creates significant societal value. In this paper, we first consider extending large-scale organ exchange to liver lobes. Liver exchange has its own interesting aspects that differ from kidney exchange. We describe and implement a general, and to our knowledge the most realistic, organ exchange *compatibility graph* generator and show that it addresses important weaknesses in the most commonly used kidney exchange generator. We show liver exchange clearing results on this demographically accurate data, and optimally clear liver exchanges at the estimated nationwide size using an enhancement to the fielded kidney exchange clearing algorithm. We then explore cross-organ donation where a candidate needing a kidney and a candidate needing a liver can swap donors, and show that this increases the total number of lives saved. We conclude with thoughts regarding the fielding of a nationwide liver or joint liver-kidney exchange from a legal and computational point of view.

Introduction

The transplantation of organs from a deceased donor to a needy living candidate first occurred nearly sixty years ago, but only became popular in the 1970s due to the introduction of immunosuppressants that help prevent the rejection of foreign organs in a patient's body. Since then, the majority of transplantation has occurred through a deceased donor waiting list consisting of needy patients who wait for any willing donors to die, resulting in the harvesting and subsequent transfer of a compatible organ from the donor's cadaver to the living patient. There is a great supply shortage of cadaveric organs in most societies (including the US), and the imbalance between supply and demand keeps growing. As of October 2012, there were 93,860 patients waiting for a kidney, 16,075 waiting for a liver, and 5,902 for another organ (e.g., pancreas, heart, lung, intestine) in the US alone.

In recent years, *live* donation of organs has significantly increased the total number of organ transplants. In live donation, a donor gives one of his two kidneys, one of his two liver lobes, or a part of an intestine, etc., to the patient so both the donor and patient can live. The effect of live donation has been most prominent in kidney donation, where a recent advance—kidney *exchange* (Roth *et al.* 2004)—has provided renewed hope to even "hard to match" patients. In kidney exchange, patients bring willing but incompatible donors to a large waiting pool. Patients can then swap incompatible donors with other patients. Matching a candidate to a donor is difficult for a variety of reasons, including blood (ABO) type, tissue (HLA) type, age, and—due to the limitations of current medical knowledge—unknown exogenous factors. Nevertheless, kidney exchanges on the regional and national scale have seen marked success over the last few years.

In this paper, we explore the creation of a living donor *liver* exchange. This is similar to kidney exchange in many ways, but remains unexplored[1] and different in three critical ways. While a donor and candidate must be blood type (ABO) compatible, (a) they need not be HLA-compatible (explained later), (b) the age of the donor and candidate makes a significant difference in transplant success (Egawa *et al.* 2004), and (c) the donor must be heavier than the candidate (or else the donor's liver, which must be cut in two before transplantation, will not be large enough to support the donor and candidate). This provides an interesting twist to both the current computational methods used for solving real nationwide kidney exchange (based on integer programming techniques proposed by Abraham, Blum, and Sandholm (2007)), as well as the theoretical matching and mechanism design results in the area. This paper provides, to our knowledge, the first foray into the computational methods and implementation necessary to set the groundwork for a fielded nationwide liver exchange. It is clear that a liver exchange would be highly beneficial for sustaining life and creating value into society. We then propose *multi-organ* exchange—where candidates in need of both kidneys and livers can swap donors in the same pool—and show that it saves significantly more lives in aggregate than running separate kidney and liver exchanges.

[1] A notable exception is that in Korea, 16 swaps of willing donors 0 in a single hospital over the course of six years. All swaps were arranged by hand. This shows the feasibility of the idea at a small scale (Hwang *et al.* 2010).

Preliminaries

In order to develop a nationwide liver or multi-organ exchange, we must first accurately model the realities of such an exchange and design optimal, scalable clearing algorithms for it. In this section, we describe the creation of a *compatibility graph* representing the space of possible swaps among n candidate-donor pairs, based on traits of the candidates and donors. We then describe the *clearing problem*, a formalization of the process used to determine an optimal set of swaps.

Compatibility Graph

We begin by encoding an n-patient organ exchange as a directed graph. Construct one vertex for each incompatible candidate-donor pair. Add an edge e from one candidate-donor vertex v_i to another v_j, if the candidate at v_j can take a liver lobe or kidney from the donor at v_i. This process creates a compatibility graph for the general concept of barter exchange, where participants can swap items with each other. Within the compatibility graph, a cycle c represents a possible swap, with each vertex in the cycle obtaining the item of the next vertex. A matching is a collection of *disjoint* cycles; no vertex can give out more than one item (e.g., more than one kidney or liver lobe). Cycles ensure that donors give items if and only if their patients receive organs.

The Clearing Problem

The *clearing problem* is that of finding a maximum-cardinality matching consisting of disjoint cycles of length at most some small constant L. The cycle-length constraint is crucial since all operations in a cycle have to be performed simultaneously. Were this not the case, a donor might back out after his incompatible partner has received a liver. This backing out is legal because, in nearly all countries including the US, it is illegal to form a binding contract over the exchange of organs. The availability of operating rooms, doctors, and staff causes long cycles to be unexecutable. As is the practice in the US-wide kidney exchange and most other real kidney exchanges, we let $L = 3$.

Denote the set of all cycles of length no greater than L by $C(L)$. Then, given binary indicator variables $\forall c \in C(L)$, we must solve the following integer linear program:

$$\max \sum_{c \in C(L)} c \qquad s.t. \qquad \sum_{c: v_i \in c} c \leq 1 \quad \forall v_i \in V$$

The clearing problem with any fixed $L > 2$ is NP-complete (Abraham *et al.* 2007). (The cases $L = 2$ and $L = \infty$ can be solved in polynomial time.) Significantly better (i.e., higher cardinality) results are found with $L = 3$ over $L = 2$, so solving the NP-complete version of the problem is necessary in practice (Roth *et al.* 2007). The problem, at least with respect to kidneys, can be solved optimally in practice at the steady-state nationwide scale using a specialized tree search algorithm based on the branch-and-price framework for integer programming (Abraham *et al.* 2007). We will later discuss this algorithm in more detail as well as enhancements to it for liver exchange and multi-organ exchange.

A Parameterized, Realistic Compatibility Graph Generator

In order to create an at-scale nationwide liver or multi-organ exchange, we first have to develop a compatibility graph generator with which we can run simulations. First, we draw data from reliable sources (here, specific to the US). Second, this data is fed into a graph creation algorithm that probabilistically determines the existence of compatible and incompatible candidate-donor pairs, as well as compatibility constraints between different candidate-donor pairs. In the large, with high probability, graphs generated by this algorithm will mimic the demographics that would prevail in a large-scale fielded exchange in the US. (Plugging different raw data (e.g., age, weight, blood type distributions) into the generator algorithm would provide realistic generation of non-US compatibility graphs.) We then conclude the section with a comparison of liver exchange graphs generated by our algorithm to kidney exchange graphs generated by the standard generator of Saidman et al. (2006). Our generator is a generalization of (i.e., more powerful than) that current standard.

Sampling from Real-World Data

Current medical knowledge is incapable of exactly predicting the compatibility of a particular donor and candidate. However, many attributes are known that can guide doctors—and algorithms—toward a realistic quantification of the chance of organ rejection. In this section, we describe these factors and the open source data sets that our algorithm uses to realistically sample the US population. In the discussions ahead, we use "OPTN" to refer to the data available from the Organ Procurement and Transplantation Network.[2] All OPTN data is current as of November 11, 2011.

Gender While a donor of one gender can donate an organ to a candidate of another gender, we must take gender into account during graph generation. This is because other traits that affect the probability of a transplant's success (e.g., weight or age) depend on a person's gender. We draw candidate genders from the OPTN data set, and donor genders from the greater US population through the 2010 US Census report.[3] Figure 1 shows the distributions of liver-needing candidates and the natural US population as donors. Men are very over-represented in the candidate pool. (Note that similar distributions can be obtained for kidney-needing candidates, and used in a multi-organ generator.)

	Male	Female
Candidate	61.71	38.29
Donor	48.53	51.47

Figure 1: Distribution of (liver) candidate and donor genders, drawn from OPTN and 2010 US Census data, respectively.

[2] http://optn.transplant.hrsa.gov/data/

[3] http://www.census.gov/compendia/statab/cats/population.html

Blood Type A candidate and donor must be ABO blood type compatible (e.g., an A-type donor is compatible with A- and AB-type candidates), although blood type suppression through drugs is a recent advance that has the potential to remove this constraint (Takahashi 2007). We draw candidate blood types from the OPTN distribution (dependent on gender), and donor blood types from the overall US.[4] The OPTN distribution is roughly equal across genders, and both distributions are roughly equal to each other. Nevertheless, it is important to have this parameterized capability in the generator in the event that, for instance, some "harder" blood type (e.g., AB) gets over-represented in the candidate pool. Figure 2 shows the exact distribution and the ABO-compatibility matrix, with percentages shown for liver-needing candidates.

Donor	Candidate			
ABO	O	A	B	AB
O	x	x	x	x
A		x		x
B			x	x
AB				x

	Male		Female	
ABO	Cand.	Donor	Cand.	Donor
O	47.83	44	48.91	44
A	38.39	42	37.08	42
B	11.37	10	11.41	10
AB	2.40	4	2.58	4

Figure 2: Top: ABO blood type compatibility matrix. Marks indicate a donor (row) as ABO-compatible with a candidate (column). Bottom: ABO percentages for candidates and donors.

Age Age plays a role in transplantation, but we were unable to find any specific quantification of the amount by which increased donor or candidate age (or, in the case of children, decreased candidate age) affects this success rate. Even without this information, age is important to model because it will allow us to generate a realistic distribution of candidate and donor *weights*, a trait whose effect is easily quantified. We sample ages (dependent on gender) for candidates from the OPTN pool and for the donors from the 2010 US Census at a granularity level of one year. To save space, Figure 3 does not separate the population into one-year segments as rows, while our generator does. In our generator we also take into account the constraint that organ donors must be 18 years old, and we normalize the distributions accordingly.

Weight Unlike in kidney exchange, the physical weight of both the candidate and donor play an enormous role in the feasibility of liver transplantation.[5] Intuitively, the size of a liver is generally proportional to the size of the person who grew it. In live liver donation, the donor's liver is cut in two

[4] http://bloodcenter.stanford.edu/about_blood/blood_types.html

[5] Large weight differences between donor and candidate can factor into kidney exchange as well, but this has not been taken into account in either the current state of the art generator or the weighting algorithms used in the fielded US-wide kidney exchange.

	Male		Female	
Age	Candidate	Donor	Candidate	Donor
<1	0.259	–	0.465	–
1–5	0.837	–	1.220	–
5–10	0.568	–	1.075	–
11–17	0.717	–	1.444	–
18–34	4.193	31.883	5.554	29.357
35–49	14.851	27.798	14.976	26.617
50–64	64.851	25.066	57.079	25.053
≥ 65	13.725	15.252	18.186	18.972

Figure 3: Probability distribution of ages, respective of candidate and donor gender.

(one lobe is removed). For both donor and candidate to remain healthy, the slice of liver left in the donor must be large enough to maintain her life, and the slice of liver given to the candidate must be large enough to maintain his. Thus, a general rule of thumb that the donor must weigh as much as (or more than) the candidate is in place in live liver donation. We adopt that convention for liver exchange.

Given the age and gender (generated separately from OPTN data for candidate and US Census data for donors, as described earlier), we sample from a fine-grained table of weights recently released by the Center for Disease Control (McDowell *et al.* 2008). This data, given on a by-year basis until age 20 and in increments of 5 years thereafter, includes mean weights, sample errors, and sample sizes. From this, we calculate a standard deviation and sample from a normal distribution with this mean and standard deviation. While there are issues with this method—most notably that the candidate weights may be drawn from a different distribution than the general US public, and that human weights are not distributed normally but are skewed toward weighing more—we feel that this sampling approach provides a reasonable starting point for future generation techniques. The full table of weights is omitted due to space.

HLA Antibodies and Antigens In kidney exchange, tissue type (HLA antibodies and antigens) are another very important determinant of compatibility. A candidate and donor sharing antigen encoding on the same locus possibly results in a positive *virtual crossmatch* across antigens. A positive virtual crossmatch means that the system can detect incompatibility. In kidney exchange graph generation, this is quantified by the probability that the candidate is not tissue-type compatible with a randomly drawn donor. This probability is called %PRA for panel reactivity antibody (Saidman *et al.* 2006). Furthermore, tissue type can change over time, resulting in the need for contingency plans after the time of algorithmic matching but before the surgery. For example, if the candidate comes down with a cold or flu days before surgery, the surgery may need to be rescheduled or permanently canceled.

In liver exchange, %PRA plays less of a role due to the use of suppressant drugs. As such, while the generator supports %PRA (and can use sampled data from the OPTN databases[6]), we exclude %PRA in our liver experiments.

[6] The relationship (e.g., spousal, parent-child) between candi-

However, %PRA is included in our multi-organ experiments for kidney candidates.

Generator Algorithm

We now give the method for generating the compatibility graph from data sampled from the sources given in the previous section. Note that the probability distributions from the previous section (and the organs to which they pertain) can be swapped without affecting the correctness of the algorithm beyond the "is compatible" checks described below.

Algorithm 1: Compatibility graph generator

Input: Integer n, real number f, real-world data
Output: Compatibility graph $G = (V, E)$ s.t. $|V| = n$
begin
  $G := (V = \emptyset, E = \emptyset)$
  while $|V| < n$ **do**
    $c =$ candidate, $d =$ donor
    $c.drawOrganType()$
    $\{c, d\}.drawGender()$
    $\{c, d\}.drawBlood(gender)$
    $\{c, d\}.drawAge(gender)$
    $\{c, d\}.drawTissueType(gender)$
    $\{c, d\}.drawWeight(gender, age)$
    if $\neg isCompatible(c, d)$ **then**
      $V = V \cup \{v^{c,d}\}$
  for $v_i, v_j \in V$ *s.t.* $V_i \neq V_j$ **do**
    if $isCompatible(v_j^c, v_i^d)$ *and* $x \in U[0, 1] > f$ **then**
      $E = E \cup \{(v_i, v_j)\}$
  return *directed compatibility graph* G

Algorithm 1 gives a two-step process for generating a compatibility graph $G = (V, E)$, given a number n, such that $|V| = n$. First, sample from real-world data until n incompatible candidate-donor pairs are generated. When generating a liver exchange, one would set the algorithm to sample from the liver data given above; however, when generating a multi-organ exchange consisting of livers and kidneys, one would include the proper proportions of kidney and liver candidates and sample from the appropriate real-world data per organ. As of the writing of this paper, the kidney waitlist is 5.84 times longer than the liver waitlist, which would be reflected in this algorithm.

If needed, the algorithm can easily be augmented to keep track of any compatible candidate-donor pairs generated. As is common practice in kidney exchange, these pairs are assumed to match on their own, and do not enter the pool.[7] Other additions could be made to the algorithm as data becomes available (e.g., correlating donor and candidate characteristics under the assumption that a donor may likely come from the candidate's family).

date and donor can yield information on HLA compatibility, and is supported by the generator of Saidman et al. and our generator.

[7]Recent kidney exchange research suggests that incentivizing even compatible pairs to join a nationwide exchange could result in better matchings (Rees *et al.* 2009; Ashlagi and Roth 2011).

After n incompatible candidate-donor pairs are generated, the algorithm steps through each pair v_i, v_j of candidate-donor pairs and, if the latter's candidate v_j^c is compatible with the former's donor v_i^d, then a directed edge is added from v_i to v_j. Note the inclusion of an exogenous "failure factor" $f \in [0, 1]$ that, if prescribed, randomly determines an edge failure even in the case of a compatibility success. This factor is common in the kidney literature (Ashlagi *et al.* 2011), and is used to account for incompleteness of medical knowledge and, during simulation, temporal fluctuations in candidate-donor compatibility.

Algorithm 1 calls a function *isCompatible(c,d)*. In the liver case, this checks whether two patients are ABO-compatible and whether the donor's weight is greater than or equal to the candidate's weight. In the kidney case, this checks whether two patients are ABO-compatible and whether a virtual crossmatch based on tissue type returns negative. As better medical knowledge and data become available, this function can be generalized to take new compatibility aspects into account.

Comparison to Kidney Exchange

We now compare our generator to the current state of the art (kidney) exchange generator (Saidman *et al.* 2006). While the generators and data are similar in spirit, the medical differences between kidney and liver compatibility create distinctly different compatibility graphs both at the small and large scale. We will discuss those differences below.

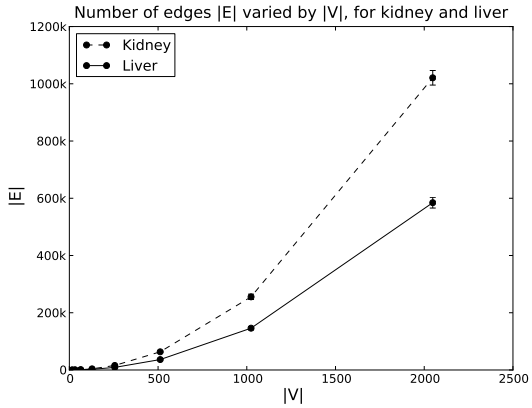

Figure 4: #Edges (in thousands) in generated liver and kidney compatibility graphs (100 graphs per $|V|$). The kidney graphs are denser than the liver graphs.

Figure 4 plots the average number of edges in the liver-only compatibility graphs, using the generator in this paper, against the average number of edges in the kidney compatibility graphs generated by the state of the art, as the number of candidate-donor pairs increases. The kidney compatibility graphs are, for graph sizes above 64, denser than comparably-sized liver compatibility graphs. This is interesting because it shows that, even though the liver exchange graphs do not need to take %PRA (i.e., HLA incompatibility) into account, their sensitivity to age and weight distri-

butions proves to be more constricting than HLA sensitivity! Regardless, neither the liver nor the kidney graphs are sparse in the classical sense of the word: at $|V| = 1024$, the number of edges in the liver graph is 26% of the total possible edges in a 1024-clique. This lack of sparsity drives the experimental computational complexity of solving the real-world clearing problem.

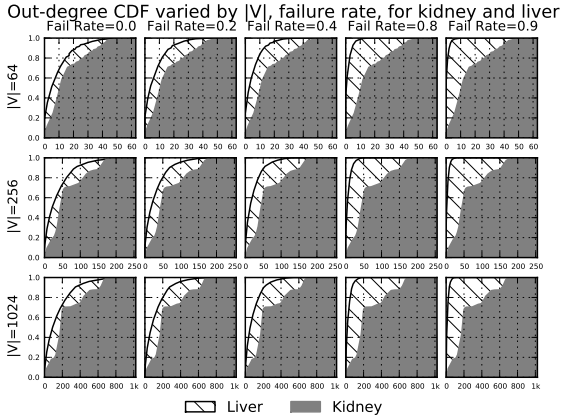

Figure 5: Cumulative distribution functions of the out-degree of vertices as we increase $|V|$ (varies per row) and exogenous failure rate f (varies by column), shown for the liver graphs (in white) and kidney graphs (in gray). Note the divergence between kidney and liver graph as the exogenous failure rate increases, as well as the three qualitative sections in the kidney graphs due to the three different %PRA classes.

Figure 5 enumerates the differences in the out-degree of the vertices in compatibility graphs for liver-only exchange generated using our algorithm (shown in white) and compatibility graphs for kidney exchange from the Saidman et al. generator (shown in gray). The size of the graph, $|V|$, is held constant along the rows, while the exogenous failure rate (f) between two otherwise compatible candidates and donors is held constant in each column. We vary $|V|$ and $f \in \{0.0, 0.2, \ldots, 0.9\}$. Note that there is no notion of an exogenous failure rate in the kidney graphs (although the %PRA virtual crossmatch simulation is similar to an exogenous failure rate, but not parameterized); as such, the kidney exchange graphs vary only in terms of cardinality.

The cumulative distribution functions over the out-degrees of vertices, shown in Figure 5, exhibit interesting behavior. For example, there are more vertices with low degree in the liver exchange graphs than in the kidney exchange graphs. More interesting is the behavior exhibited by the kidney exchange graphs as $|V|$ increases. For instance, when $|V| = 1024$, we see three distinct out-degree sections in the kidney exchange graphs. These are an artifact of the somewhat ad-hoc method of doing %PRA virtual crossmatch tests in the Saidman et al. generator. The generator groups pairs into three sensitivity levels ("high", "medium", and "low"). As $|V|$ increases, those patients who are highly sensitized tend toward very few edges, while those at the medium and low sensitivity levels tend toward a medium and high number of edges, respectively. We believe that this is an artifact of the generator by (Saidman *et al.* 2006) and is not representative of the real kidney exchange data. Our generator (even if used for kidneys) does not have such coarse artifacting because it can bucket sensitivity into finer-grained classes.

The Clearing Algorithm

Now that we have an instance generator to enable testing, we will experiment on clearing liver and multi-organ exchanges to determine whether this is viable at the nationwide scale.

We first briefly discuss the most scalable optimal kidney exchange clearing algorithm (Abraham *et al.* 2007), which is also used in the US-wide kidney exchange; we adapt that algorithm for our liver exchange experiments. At a high level, given a compatibility graph $G = (V, E)$, the algorithm (intelligently) enumerates cycles of length at most L and chooses the optimal disjoint set of these cycles according to the objective function of maximizing cardinality of the matching.

In reality the number of cycles is prohibitively large (cubic in $|E|$ for $L = 3$) to write down in memory. Therefore, solving this problem hinges on a technique called *branch-and-price* (Barnhart *et al.* 1998), a method for incrementally generating only a small part of the model during tree search, yet guaranteeing optimality by proving that all the promising variables have been incorporated into the model. The actual solver uses several additional techniques to make kidney exchange clearing scalable for memory and time (Abraham *et al.* 2007). It uses empirically and theoretically motivated heuristics to seed the initial cycle (i.e., variable) set used on the model, and then incrementally brings cycles into the model depending on their shadow price, a quantitative estimate of a cycle's utility given the current model. Optimality is proven when no cycles can possibly increase the objective. The algorithm also uses specific branching heuristics and primal heuristics to construct feasible initial integral solutions at each branch. If these integral solutions match the (restricted, possibly fractional) LP solution, then the subtree can be pruned and optimality potentially proven.

A Liver-Specific Cycle Seeding Heuristic

The selection of the initial seed columns—representing individual cycles—is a heuristic process. The prior algorithm uses the cycles from two heuristically-generated feasible solutions (very few such cycles) and hundreds of thousands of randomly selected cycles from $C(L)$. Since enumerating $C(L)$ in its entirety is a costly ordeal, their sampling relies on a series of random walks. Starting at a randomly chosen vertex, a random walk takes steps to new vertices. At each step, if an edge exists leading back to the initial vertex, the corresponding cycle is added to the set of seed cycles and a new start vertex is chosen. This results in a randomized, but not uniformly random, sampling of all cycles.

We define a different sampling method for the cycle seeding problem. Liver compatibility graphs tend to have a large number of vertices with low out-degree (Figure 5). These candidates are difficult to match. With this in mind, we conduct a biased random walk sampling in the same spirit as

the prior algorithm, except weighting the selection of the randomized start vertex inversely proportional to its out-degree. This can be done efficiently through an initial sorting of the vertices by out-degree, a process whose one-time $O(|V| \log |V|)$ runtime is overshadowed by the NP-hard clearing problem.

Experimental Results

We now provide some preliminary computational results for a hypothetical nationwide liver or multi-organ exchange, using the realistic data generated by Algorithm 1. First, we describe timing and matching results in the *static* case, where the algorithm sees the problem in its entirety up front. Second, we describe results for the *dynamic* case, where candidate-donor pairs arrive in the pool over time and are either matched or die waiting. We show results at sizes mirroring an estimated steady-state size of a US-wide liver exchange. Finally, we explore the possibility of a multi-organ exchange, where both liver- and kidney-needing candidates can swap donors in the same pool. This results in more lives being saved than were the nation to run separate liver and kidney exchanges.

Static Liver Exchange Experiments

In the static case, the generator outputs a single graph and the optimization engine solves the clearing problem on this graph exactly once. Figure 6 shows timing results on liver exchange graphs of various sizes $|V|$ and exogenous failure rates f. Intuitively, when f is low (or zero), the optimizer must consider many more edges than when f is high, resulting in longer runtimes for denser graphs. As expected, the computation time increases drastically with the size of the graph.

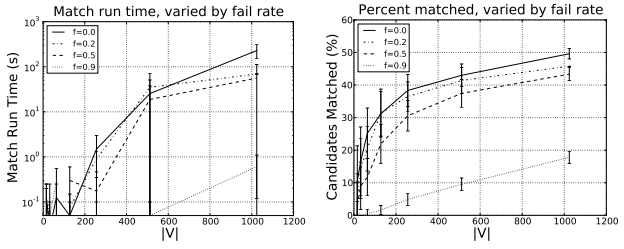

Figure 6: Match runtime (left) and percentage of candidates matched (right), varying failure rate f and graph size $|V|$. Lower failure rates result in denser graphs and longer runtimes (as well as higher variance).

Figure 6 also shows the percentage of candidates matched (the number of candidates matched by the algorithm divided by the total number of candidates in the pool) as a function of compatibility graph size $|V|$ and exogenous failure rate f. Intuitively, when f is held low, the percentage of candidates matched is higher than when the failure rate is high. Of interest is the match behavior as $|V|$ increases. Regardless of f, the percentage of candidates matched increases with the size of the underlying compatibility graph. This behavior is similar to that seen in kidney exchange and motivates the need for a large (i.e., nationwide) liver exchange.

Addressing the needs of society. The estimated steady-state monthly size of the nationwide kidney exchange is 10,000 candidate-donor pairs. Given the current waitlist sizes for kidneys and livers in the US, we provide a rough estimate of the steady-state size of a nationwide liver exchange. With 93,860 candidates waiting for a kidney and 16,075 candidates waiting for a liver, the steady-state for the liver exchange can be estimated at approximately 16,075 / 93,860 $\approx$ 17% of 10,000, or roughly 1,700 candidates. So, our clearing algorithm should be able to handle monthly batch runs of a nationwide liver exchange.

Dynamic Liver Exchange Experiments

In the dynamic case, a variable number of candidates enter and leave the pool over a period of multiple time units. While the fielded nationwide kidney exchange currently operates under the static paradigm described earlier, recent work in the kidney exchange community has shown that optimizing in the dynamic setting leads to both more realistic and higher cardinality matchings over time (Awasthi and Sandholm 2009; Ünver 2010; Dickerson *et al.* 2012a). Because of this, the dynamic setting in general organ exchange cannot be ignored.

We start with a pool of $|V| = 1024$ candidates assumed to be highly sensitized patients who built up in the system over time. These are matched myopically. Given a matched cycle by the algorithm, we then simulate that transplant actually succeeding in real life via an exogenous parameter set to $f = 0.5$. If any edge in a cycle fails, that entire cycle fails, and all candidates are returned to the pool (with the failed edge removed). We simulate candidates leaving the pool (either through finding a transplant or dying). On expectation $|V_{new}| = 213$ new candidates arrive in the pool per month, and the algorithm continues. We test over 24 months.

Figure 7 shows the number of candidates matched at each time period. This is the number of candidates matched *by the algorithm*, but before the virtual failures are taken into account. On the left, 12% of candidates will be alive after 10 years, corresponding to the expected lifetime of a kidney patient on dialysis waiting for a kidney (USRDS 2007). On the right, the probability of a candidate dying is set to an expected life of 1–2 years. This mimics the urgency of needing a liver transplant. While dialysis can be used to keep a patient with failed kidneys alive, no such treatment exists for livers. This corresponds to a significant drop in the number of candidates matched, due to the decreased number of candidates in the pool at each time period. (Note that a large number of candidates are matches per month in the beginning when the exchange goes live because there is a large pool that has accumulated. Soon thereafter a steady state is reached.)

Figure 8 shows matching time (as a function of both optimization and graph evolution time, although graph evolution time is 0) over the 24 time periods. After the initial solution to the large $|V| = 1024$ starting compatibility graph, solution time hits a steady-state on the order of a few minutes. We see that the 0 graph with candidates' expected lifetime set to the short lifetime of a liver patient solves more quickly than those set to the expected lifetime of a kidney patient on

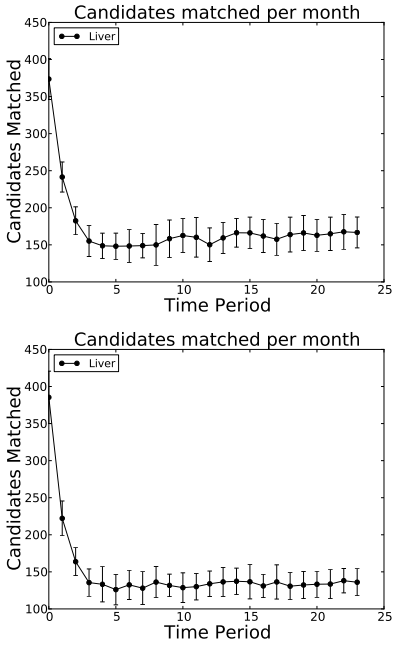

Figure 7: Number of candidates matched per time period in a dynamic setting over $T = 24$ months, with an expected lifetime per candidate of 10 years (top) or 1–2 years (bottom).

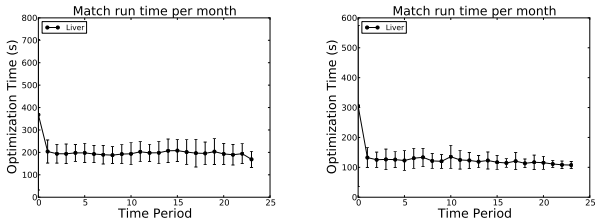

Figure 8: Match run (i.e., optimization only) runtime for a dynamic setting over $T = 24$ months, for a graph with $|V_s| = 1024$ initial candidates in the pool and an expected number of 213 new candidates entering the pool per time period, with an expected lifetime per candidate of 10 years total (left) or 1–2 years (right).

dialysis; this is a function of the smaller matchings shown in Figure 7, again showing the urgency of liver transplantation.

Dynamic Bi-Organ Exchange Experiments

In this section, we expand beyond simulating a dynamic liver exchange to the novel concept of *multi-organ* exchange. In the long run, one could imagine exchanges of multiple different kinds of organs. However, to our knowledge, only kidneys and livers have ever been swapped (and only separately). In any case, kidneys and livers are by far the most common organ transplants. Therefore, in this section we will focus on kidneys and livers. We show that combining an independent nationwide liver exchange with a nationwide kidney exchange into a joint kidney-liver exchange results in a statistically significant increase in the number of organ transplants.

Using our parameterized generator described in Algorithm 1, we simulate a bi-organ exchange featuring candidates in need of either a kidney or a liver who can swap donors in a *combined* candidate-donor pool. Approximately 85% of the candidates in the simulated pool need kidneys, while the other 15% need livers, as determined by the most recent OPTN waitlist data. We mimic the experiments in the previous section, with a starting pool size of $|V| = 1024$ candidates who are highly sensitized and are assumed to have built up in the pool over time. We use the same exogenous transplant failure parameter ($f = 0.5$) as in the previous section, and simulate candidate-donor pairs entering and exiting the pool in a similar fashion. To generate the candidates, we draw from the two different US distributions based on whether the candidate needs a kidney or a liver. Naturally, donors are drawn from the same US distribution in the two cases. We test over 24 months.

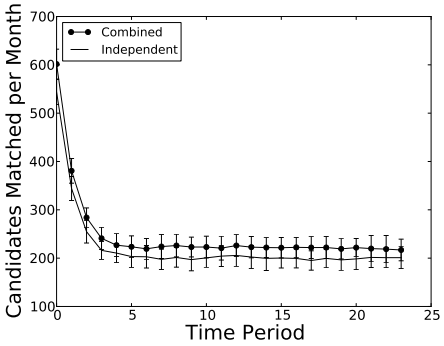

Figure 9: Number of candidates matched in independent liver and kidney exchanges and a combined multi-organ exchange, per time period, in a dynamic setting over $T = 24$ months.

Figure 9 shows the number of candidates matched each month in the combined bi-organ exchange, as well as the aggregate number of candidates matched while keeping both liver- and kidney-needing candidates in separate pools. Clearly evident is the loss of life resulting from keeping both the liver and kidney pools independent, with the bi-organ exchange matching roughly 20–30 more candidates per month when compared to the two independent exchanges.

When we compare the *total* number of matches made over the entire two-year period simulated above, the difference in lives saved between two independent pools and the combined bi-organ pool is even more stark. In our experiments, the combined bi-organ pool produced 10.6% more matches than the sum of the two independent organ pools. An independent samples t-test revealed that the difference between the aggregate number of lives saved using independent, simultaneous liver and kidney exchanges and using a combined multi-organ exchange was significant, $t(52) = 19.43$, $p \ll 0.0001$.

Conclusions and Future Work

We explored the possibility of extending large-scale organ exchange to liver lobes. We developed a general, and to our knowledge the most realistic, organ exchange compatibility graph generator and showed that it addresses important weaknesses in the most commonly used kidney exchange generator. Liver exchange has its own interesting aspects that differ from kidney exchange. We showed liver exchange clearing results on demographically accurate data, and optimally cleared liver exchanges at the estimated nationwide size using an enhancement to the state-of-the-art kidney clearing algorithm. We explored the prospect of multi-organ exchange, where candidates needing either a liver or kidney can swap willing donors in the same pool. We showed that this combination of donor pools results in significantly more lives saved.

This paper is intended as a first foray into automated liver and multi-organ exchange. As such, there is much room for future research (much of which is applicable to other organ exchange and even to barter exchanges beyond organs), and is motivated by experiences fielding the nationwide kidney exchange. One direction of future work is to take on the slow and politics-laden task of founding a liver exchange, or including livers in currently fielded kidney exchanges. Another is to develop scalable computational methods for the dynamic problem. Even for kidneys, the best current techniques are for simplified models (Ünver 2010; Ashlagi *et al.* 2013) or face computational challenges (Awasthi and Sandholm 2009; Dickerson *et al.* 2012a).

Even for the static problem, scalability problems tend to get worse with the inclusion of a recent innovation in kidney exchange—donation chains started by altruistic donors. The cycle length cap L no longer applies to chains since they do not require simultaneous execution. Recent work explores this innovation, and hits computational limits experimentally with long chains (Ashlagi *et al.* 2012; 2011; Dickerson *et al.* 2012a; 2012b; Gentry and Segev 2011; Gentry *et al.* 2009). We do not expect altruistic donors in liver exchange because a liver donation is significantly riskier for the donor than a kidney donation, complicating the ethical considerations of even allowing altruistic donors in the pool (Woodle *et al.* 2010). However, that remains to be seen. In any case, one could include chains started by kidney-donating altruists into a bi-organ exchange—if the scalability challenges of chains can be adequately addressed.

Finally, this paper (and most papers on kidney exchange) deals with optimizing algorithmic organ matches; in reality, most algorithmic matches in fielded kidney exchanges do not result in an actual transplant. We expect this would be the case in liver and multi-organ exchange as well, although the exact failure rates for liver and multi-organ exchanges would be different than the observed failure rates in currently fielded kidney exchanges due to the medical and logistical differences in the organs and the transplant processes. Making organ exchange failure-aware is a critical step toward improving yield; recent work explores this notion (Blum *et al.* 2013; Dickerson *et al.* 2013) to both theoretically and empirically maximize the expected number of actual transplants stemming from an algorithmic match. Recent work by Glorie (2012) is an initial foray into learning a better estimate of the probability of a transplant failure between a patient and a donor, but much is left to be done.

Regardless, the urgent societal need for liver exchange is there today, and we hope to be able to address it through a dedicated or combined liver- or multi-organ exchange.

References

David Abraham, Avrim Blum, and Tuomas Sandholm. Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges. In *ACM Conference on Electronic Commerce (EC)*, 2007.

I. Ashlagi and A.E. Roth. Individual rationality and participation in large scale, multi-hospital kidney exchange. In *ACM Conference on Electronic Commerce (EC)*, pages 321–322, 2011.

I. Ashlagi, D.S. Gilchrist, A.E. Roth, and M.A. Rees. Nonsimultaneous chains and dominos in kidney-paired donation—revisited. *American Journal of Transplantation*, 11(5):984–994, 2011.

I. Ashlagi, D. Gamarnik, M. Rees, and A.E. Roth. The need for (long) chains in kidney exchange. NBER Working Paper No. 18202, July 2012.

Itai Ashlagi, Patrick Jaillet, and Vahideh H. Manshadi. Kidney exchange in dynamic sparse heterogenous pools. In *ACM Conference on Electronic Commerce (EC)*, 2013.

Pranjal Awasthi and Tuomas Sandholm. Online stochastic optimization in the large: Application to kidney exchange. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 405–411, 2009.

Cynthia Barnhart, Ellis Johnson, George Nemhauser, Martin Savelsbergh, and Pamela Vance. Branch-and-price: column generation for solving huge integer programs. *Operations Research*, 46:316–329, 1998.

Avrim Blum, Anupam Gupta, Ariel D. Procaccia, and Ankit Sharma. Harnessing the power of two crossmatches. In *ACM Conference on Electronic Commerce (EC)*, 2013.

John P. Dickerson, Ariel D. Procaccia, and T. Sandholm. Dynamic matching via weighted myopia with application to kidney exchange. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 1340–1346, 2012.

John P. Dickerson, Ariel D. Procaccia, and T. Sandholm. Optimizing kidney exchange with transplant chains: Theory and reality. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 711–718, 2012.

John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Failure-aware kidney exchange. In *ACM Conference on Electronic Commerce (EC)*, 2013.

H. Egawa, F. Oike, L. Buhler, AM Shapiro, S. Minamiguchi, H. Haga, K. Uryuhara, T. Kiuchi, S. Kaihara, and K. Tanaka. Impact of recipient age on outcome of ABO-incompatible living-donor liver transplantation. *Transplantation*, 77(3):403, 2004.

S. E. Gentry and D. L. Segev. The honeymoon phase and studies of nonsimultaneous chains in kidney-paired donation. *American Journal of Transplantation*, 11(12):2778–2779, 2011.

S.E. Gentry, R.A. Montgomery, B.J. Swihart, and D.L. Segev. The roles of dominos and nonsimultaneous chains in kidney paired donation. *American Journal of Transplantation*, 9(6):1330–1336, 2009.

Kristiaan M. Glorie. Estimating the probability of positive crossmatch after negative virtual crossmatch. Technical report, Erasmus School of Economics, 2012.

Shin Hwang, Sung-Gyu Lee, Deok-Bog Moon, Gi-Won Song, Chul-Soo Ahn, Ki-Hun Kim, Tae-Yong Ha, Dong-Hwan Jung, Kwan-Woo Kim, Nam-Kyu Choi, Gil-Chun Park, Young-Dong Yu, Young-Il Choi, Pyoung-Jae Park, and Hea-Seon Ha. Exchange living donor liver transplantation to overcome ABO incompatibility in adult patients. *Liver Transplantation*, 16(4):482–490, 2010.

M.A. McDowell, C.D. Fryar, C.L. Ogden, and K.M. Flegal. Anthropometric reference data for children and adults: United States, 2003–2006. *Nutrition*, 10(10):1–45, 2008.

Michael Rees, Jonathan Kopke, Ronald Pelletier, Dorry Segev, Matthew Rutter, Alfredo Fabrega, Jeffrey Rogers, Oleh Pankewycz, Janet Hiller, Alvin Roth, Tuomas Sandholm, Utku Ünver, and Robert Montgomery. A nonsimultaneous, extended, altruistic-donor chain. *New England Journal of Medicine*, 360(11):1096–1101, 2009.

Alvin Roth, Tayfun Sönmez, and Utku Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119(2):457–488, 2004.

Alvin Roth, Tayfun Sönmez, and Utku Ünver. Efficient kidney exchange: Coincidence of wants in a market with compatibility-based preferences. *American Economic Review*, 97:828–851, 2007.

S. L. Saidman, Alvin Roth, Tayfun Sönmez, Utku Ünver, and Frank Delmonico. Increasing the opportunity of live kidney donation by matching for two and three way exchanges. *Transplantation*, 81(5):773–782, 2006.

K. Takahashi. Recent findings in ABO-incompatible kidney transplantation: classification and therapeutic strategy for acute antibody-mediated rejection due to ABO-blood-group-related antigens during the critical period preceding the establishment of accommodation. *Clinical and Experimental Nephrology*, 11(2):128–141, 2007.

U. Ünver. Dynamic kidney exchange. *Review of Economic Studies*, 77(1):372–414, 2010.

United States Renal Data System (USRDS), 2007. http://www.usrds.org/.

Steve Woodle, John Daller, Mark Aeder, Ron Shapiro, Tuomas Sandholm, Vincent Casingal, David Goldfarb, Richard Lewis, Jens Goebel, and M Siegler. Ethical considerations for participation of nondirected living donors in kidney exchange programs. *American Journal of Transplantation*, 10:1460–1467, 2010.