An Axiomatic and Empirical Analysis of Mechanisms for Online Organ Matching

Nicholas Mattei, Abdallah Saffidine, and Toby Walsh

Abstract

Matching donations from deceased patients to patients on the waiting list account for over 85% of all kidney transplants performed in Australia. As algorithms are given responsibility to make decisions that impact our lives, there is increasing awareness of the need to understand the properties, e.g., efficiency, fairness, stability, and strategyproofness, of these algorithmic decisions. Deceased organ matching is an unusual two-sided market as both sides are dynamic and at each time step. In addition, both sides of the market share identical preferences over the other side: each prefers the healthiest organ or patient available. However, these preferences are dynamic as the health and age of the patients change over time. We propose a simple mechanism to perform matching in this dynamic, two-sided market and compare this new mechanism with the more complex algorithm currently under consideration by the Organ and Tissue Authority in Australia. We study axiomatic properties of these online mechanisms including the robustness measures of stability and strategyproofness. We identify a number of different types of fairness, such as to patients, to regions and to blood types and consider how they can be achieved. We also perform experiments using real world data provided by the Organ and Tissue Authority of Australia. Through both the axiomatic and empirical analysis we find that our simple mechanism is more efficient, more stable, and fairer compared to the other mechanism currently under consideration.

1 Introduction

Kidney disease costs the Australian economy billions of dollars every year. Over ten thousand people in Australia are on dialysis, each costing hundreds of thousands of dollars in medical and welfare costs. On 1st February 2016, there were 1083 people on the waiting list for a kidney transplant. Australia is especially challenged in this area as kidney disease is a major problem within the indigenous population. The incidence of end stage kidney disease in the indigenous population in remote areas of Australia is 18 to 20 times higher than that of comparable non-indigenous peoples.[1] Over the course of 2016, 1448 people received a kidney from a deceased donor, whilst a further 265 people received a kidney from a living donor. Just 44 of these came from paired exchanges. Paired exchange represents less than 3% of all transplants.

A significant trend in Australia (as in other developed countries) is that age plays a major role in kidney disease. It is impacting both the demand and supply sides of the kidney transplant market. On the demand side, the age of patients in Australia waiting to receive a kidney has increased significantly in recent years. In 2010 just 11% of the waiting list were 65 years or older while in 2015, this had increased to 15%. Over the next 30 years, the proportion of the population of Australia aged over 65 years is predicted to double to around 25 per cent. This aging demographic will likely further increase the age of people on the waiting list for a kidney transplant. On the supply side of the market, the age of donated kidneys has also increased significantly. In 1989, the mean age of donated kidneys in Australia was just 32 years old. In 2014, this had increased dramatically to 46 years old. Surgeons are now able and willing to transplant older kidneys into older patients. In 1989, the oldest transplanted kidney came from a donor aged 69 years. In 2014, this has increased to an 80 year old donated organ. A number of factors including increasing life expectancy, medical advances, and improved road safety have been driving these changes on one or both sides of the market.

[1] These (and subsequent) statistics about kidney disease and transplant in Australia are taken from [5].

Matching in Australia currently uses a mechanism based on first-come, first-served. Simply put, the longer one waits, the closer one moves to the top of the waiting list. An arriving organ is offered to the highest person on the list who is a compatible match. Given the limited supply of organs, whilst this mechanism is "fair" from a procedural perspective, it is now no longer viewed by many in the medical profession as "efficient" in terms of best use of the limited supply of organs. A 70-year-old patient may receive a kidney from a 30-year-old donor, and live 10 or even 15 years. But the organ might have lasted decades longer in a younger body according to UNOS statistics. Critics of the current system have argued that the organ's full potential for giving life is "wasted" on an older person. In contrast, an organ from a 60-year-old donor transplanted into a 30-year-old patient may fail before the patient reaches old age, thereby creating the need for an additional organ. Worse still, the patient may be sensitized by the immune suppression drugs and so require an even closer match. Transplanting young organs into old patients, and old organs into young patients might therefore be considered less than optimal [23].

Care must be exerted when designing new mechanisms as we must understand how certain game theoretic properties could arise. In 2012, a huge scandal was uncovered in Germany when it was discovered that a number of doctors were fabricating patient data and doctoring patient samples in an attempt to improve the ranking of their patients waiting for liver transplants in the Eurotransplant scheme.[2] The scandal resulted in several criminal prosecutions, and an approximately 30% drop in donation rates in the subsequent two years as public confidence in the scheme fell drastically.

We have therefore started to work with the Organ and Tissue Authority, the statutory body in Australia that allocates organs, to develop a new mechanism that explicitly takes account of the age of the organs and of the recipients for the first time. There is thus a desire for a new mechanism that matches the age of patients to that of the organs and does so in a fair and efficient manner while avoiding opportunities for manipulation. In this paper, we consider what it means for a deceased organ matching mechanism to be fair when it takes account of features like age. When reducing the pool of candidate matches by age, other concerns come into focus like geography and blood type. For instance, matching nationally rather than at the state or even hospital level improves the quality of matches possible. However, there are concerns that organs will flow out of the less populated states and territories to the larger states where demand is highest. We suggest that fairness needs to be considered on multiple levels: patient, region, and blood type to name just three. We argue that a simple mechanism that matches the age of organs to age of patients gives an allocation that is stable under some rather natural preferences for the two sides of the market. There is thus no "incentive" to deviate from the outcome it returns. This simple mechanism also offers patients a uniform time on the waiting list; so it is procedurally "fair" like the current mechanism which ignores ages.

Contribution. We compare mechanisms for the *online* organ matching problem using both an axiomatic and experimental analysis to study issues including manipulation, stability, and fairness. Incorporating these concerns into allocation schemes is an important step in adopting these algorithms and an important area of research [17, 28]. This work offers a case study in how to efficiently and fairly solve *online* allocation problems [2] and it fits into a broader research program to develop models and mechanisms for resource allocation problems that reflect the richness and complexity of the real world [34, 7]. Through both the axiomatic and empirical analysis we find that our simple mechanism is more efficient, more stable, and fairer compared to the other mechanism currently under consideration.

2 Mechanisms for Kidney Matching in Australia

We are working closely with the Organ and Tissue Authority of Australia to develop a new mechanism for allocating donated kidneys to patients. Their charter requires the allocation of organs to be fair and efficient, though it is not formally defined what this means. The current mechanism in use

[2] https://www.theguardian.com/world/2013/jan/09/mass-donor-organ-fraud-germany

in Australia offers organs sequentially to patients on the waiting list based on their compatibility and time on list. Old organs will be offered to young patients, who will likely decline the organ, causing the transplant to be delayed as the organ is now offered to the next patient on the waiting list. And young organs will be offered to old patients, which may not get the best use out of a healthy organ. Hence, there is a desire for a new mechanism that takes account the age of patients and organs.

Following the US lead the new mechanism will take account of these ages using two measures: the Kidney Donor Patient Index (KDPI) and a patient's Expected Post-Transplant Survival (EPTS) score. The Kidney Donor Patient Index (KDPI) is an integer from 0 to 100 that is calculated from the age of the donor, their diabetic status, cause of death and other factors. A donated kidney with a KDPI of X has an expected risk of graft failure greater than X% of all donated kidneys. Expected Post-Transplant Survival (EPTS) is also an integer from 0 to 100 that is calculated from the age of the recipient, their diabetic status, the number of prior organ transplants, their time on dialysis and a number of other factors. A patient with an EPTS of Y receiving a high quality donated kidney has an expected survival time that is greater than 100-Y% of all patients. For a detailed discussion of these measures see `www.srtr.org` and the work of Bertsimas *et al.* [9].

Whilst the Organ and Tissue Authority have decided to use KDPI and EPTS in their new mechanism, they have yet to decide on the precise details. Currently they collect KDPI and EPTS but do not use it when proposing a match. One candidate under consideration by the Organ and Tissue Authority is the BOX mechanism. This favors those matches that fit in the box bounded by KDPI ≤ 50 and EPTS ≤ 25 and this ordering can be viewed graphically in Figure 1. It favors matches in a rectangular box to the bottom left, then above the upper left diagonal, and then towards the lower right diagonal. The BOX mechanism has some similarity to the current US mechanism that preferentially offers organs with KDPI ≤ 20 to patients with an EPTS $\leq$20.

However, it would seem that the first objective should be to allocate organs so that the KDPI of an organ is close to the EPTS of the patient receiving the organ. Therefore we propose a simple MIN mechanism [27]. This picks a compatible patient for an arriving organ that minimizes |KDPI-EPTS|, tie-breaking by time on waiting list. Thus we are left with two mechanisms to compare.

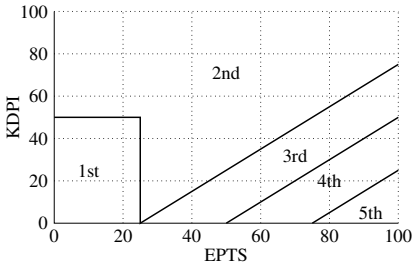


Figure 1: The BOX mechanism with a lexicographical scoring function with five key regions.

MIN**:** Allocate an arriving organ to a compatible patient that minimizes $|KDPI - EPTS|$, tie-breaking by time on the waiting list and then randomly.

BOX**:** This mechanism ranks patients according to a lexicographical scoring function. The most important terms in the scoring function ensure a match between compatible types. The least important terms tie-break according to features like time on the waiting list. The middle term orders matches as: (1) $KDPI \leq 50$ and $EPTS \leq 25$; (2) $EPTS - 25 < KDPI$; (3) $EPTS - 25 \leq KDPI < EPTS - 50$; (4) $EPTS - 50 \leq KDPI < EPTS - 75$; (5) $EPTS - 75 \leq KDPI$.

3 A Formal Model for Online Organ Matching

Allocated organs from deceased donors to patients is inherently an online problem. Online problems are well studied in computer science especially in areas like scheduling [2, 30]. In an online organ matching problem, we do not know in advance when organs will arrive, and must match them almost immediately. We propose a formal model for online organ matching in which, at each time step, one of three actions occurs: (1) a set of organs arrives, (2) a set of patients arrive, (3) or a set of patients depart. Organs are matched as soon as they arrive, whilst new patients are simply added to the waiting list. In practice, organs tend to arrive in pairs as each deceased donor typically donates both their kidneys. Patients depart the waiting list when they are matched to an arriving organ, receive a transplant elsewhere, or become too sick for transplant or die. In Australia, the size of the waiting list remains roughly constant over time.

We employ a model of organ matching where each arriving organ has an associated blood type, KDPI and state while each patient has an associated blood type, EPTS and state. The EPTS of a patient slowly increases over time as they spend longer on dialysis, in addition to other jumps such as the possible onset of diabetes. We only permit matches of a compatible type, e.g. organs coming from a donor of blood type O can be matched to recipients of any blood type, whilst organs coming from a donor of blood type AB can be matched only to recipients of blood type AB. We focus on compatibility by blood type, but in the real world there are a host of additional factors, e.g. HLA type, for a more nuanced discussion see the data and simulators from `www.srtr.org`.

We define the quality of a matching of an organ to a patient in two ways. At the level of the individual, the *utility* of a match is simply $100 - KDPI$. A patient wants the youngest possible organ. At the level of society, the *welfare* of a match is $100 - |KDPI - EPTS|$. To maximize the benefit to society of a limited supply of organs, we want the *KDPI* to be as close as possible to *EPTS*. As we consider randomized mechanisms, we can compute such measures as utility and welfare both ex post and ex ante.

4 Desirable Axioms for Allocation

An axiomatic study of algorithms is a corner stone of game theory [25] and computational social choice [11]. For instance, both the MIN and BOX mechanisms are **anonymous**: identical patients added to the waiting list at the same time are treated identically. A mechanism satisfies **participation** if a patient cannot increase their expected utility by joining the waiting list at a later date. Unfortunately, the online nature of organ matching means that no mechanism can satisfy participation.

Theorem 1 *No mechanism for online organ matching satisfies participation.*

Proof. Suppose two patients join the waiting list at the first time step, an organ with a KDPI of 50 arrives at the second time step, and one with a KDPI of 0 at the third. Both patients have an incentive to wait till after the first organ is allocated before joining the waiting list. □

A fundamental and desirable property of an allocation of organs is efficiency. In this setting all agents have the same utility value over the organs. If you lower the KDPI of the organ matched to one patient, you must inevitably increase the KDPI of an organ matched to some other patient. Therefore all allocations are efficient in terms of utility. We consider also efficiency in terms of welfare. An online allocation is **welfare efficient** if and only if there is no other online allocation in which one patient has greater welfare and none of the other patients have lower welfare. A mechanism is welfare efficient if and only if it only returns online allocations that are welfare efficient. Note that we limit our discussion to online allocations. Organs can only be matched to patients that are actually present on the waiting list as we cannot match to a patient yet to join the waiting list.

Example 1 *Suppose we have one patient on the waiting list with an EPTS of 100, and an organ with a KDPI of 0 arrives and is matched to this patient. Suppose a new patient now arrives with an*

EPTS of 0, followed by an organ with a KDPI of 100, that is again matched with the only patient on the waiting list. This is the only possible online allocation so it is welfare efficient. But in an offline setting, this allocation is not welfare efficient. We could match the first patient with the second organ, and the second patient with the first organ, increasing the welfare of both patients.

Another desirable property of online mechanisms is that they do not lead to **starvation**. In online settings this is typically handled by giving higher priority to agents the longer they wait. In all the proposed mechanisms, wait time is only used for tie-breaking between agents. It is possible in the worst case for patients never to be matched. However, we do not observe this in practice with the historical distribution of patients and organs. Both mechanisms under consideration may never match patients on the list.

5 Stable Allocations

We can view the allocation of organs from deceased donors as a two-sided matching problem. On one side of the market, we have the patients on the waiting list. To maximize their post transplant survival time, each patient simply wants to receive the best quality organ, i.e., the organ with the lowest KDPI possible. Hence, the patients have identical preferences over the organs. On the other side of the market, we have the organs with preferences over the patients. Of course, organs don't *actually* have preferences. They are just organs. And the donors are deceased so also arguably don't have preferences for their organs at this point. This preference of the organs could be seen as a societal or medical preference, i.e., to ensure the maximum good from limited supply. We suppose then that the preference of the organ side of the market is to minimize |KDPI -EPTS|. Roughly speaking, the preference of this side of the market is to match age of organ to age of patient.

A fundamental notion in a two-sided matching market is stability [31]. We do not want an allocation where a patient and organ not currently matched to each other would both prefer to be matched to each other rather than their current matching. That is, we want an allocation where there is no incentive to deviate. This is, however, a somewhat unusual two-sided matching market compared to the usual two-sided market in which we have a static set of men and women with static preferences over each other [18]. One side of our market is dynamic as patients arrive and depart over time. In addition, the preferences of the other side of the market are dynamic as the EPTS of the patients change over time. In the case of deceased organ matching, both sides of the market have identical preferences, so rather than a possible lattice of solutions [18], it turns out that under modest assumptions there is an unique stable allocation, ignoring the permutation of patients with identical EPTS and of organs with identical KDPI. This stable allocation is the one in which KDPI=EPTS.

Theorem 2 *In the two-sided online organ matching market, despite EPTS scores changing over time and patients arriving and departing, there is always a stable allocation.*

We suppose that the market is large so that there are enough compatible patients to ensure that KDPI can be made equal to EPTS in each match. There is then an unique stable allocation. In this allocation, all organs with a KDPI of 0 are matched with patients with an EPTS of 0. The patients receive organs with lowest possible KDPI so cannot be happier. And the organs gets matched to patients so that KDPI=EPTS so again could not be happier. All the organs with a KDPI of 1 are matched with patients with an EPTS of 1. The organs get matched to patients so that KDPI=EPTS so could not be happier. And the patients could only be happier if they were matched with an organ with a KDPI of 0. But none of these organ would prefer such a match. Similarly, the organs with a KDPI of 2 are matched with patients with an EPTS of 2, and so on. The unique stable allocation has KDPI=EPTS in every match.

Constructing this stable allocation required us to have complete knowledge of the organs to arrive, as well as the patients to arrive and depart. In practice, this is an online problem and we do

not have such knowledge when proposing matches. Indeed, we can show an impossibility result that no mechanism for online organ matching can be *guaranteed* to return a stable allocation because of the uncertainty about the future.

Theorem 3 *No mechanism for online organ matching is guaranteed to return a stable allocation.*

Proof. Suppose we have two patients on the waiting list, one with a constant EPTS of 10, and another with a constant EPTS of 20, and an organ arrives with a KDPI of 30. There are two cases. In the first case, the mechanism matches this first organ to the patient with an EPTS of 10. In this case, the next organ to arrive has a KDPI of 20. This is matched to the remaining patient with an EPTS of 20. This is unstable as the organ in this matching with a KDPI of 20, and the patient in the first matching with an EPTS of 10 would both prefer to be matched with each other. In the second case, the mechanism matches this first organ with the patient with an EPTS of 20. In this case, the next organ to arrive has a KDPI of 40. This is matched with the remaining patient with an EPTS of 10. This is unstable as this patient, and the organ in the first matching with a KDPI of 30 would both prefer to be matched with each other. □

Although we cannot guarantee that an online mechanism will return a stable allocation, the MIN mechanism is likely to return an allocation that is stable in a large market as the unique stable allocation is the one where KDPI=EPTS for all pairs. In the experimental section, we will demonstrate that the MIN mechanism does indeed return an allocation close to stable, even taking into account the dynamic features of the market.

6 Strategyproof Allocations

Could we have avoided the Germany transplant scandal? Could we design a mechanism where patients (or their doctors) cannot manipulate their EPTS to improve the probability of being matched? As in the German case, we focus on manipulations that improve the probability of a patient being matched. We could, however, give very similar results if we considered instead manipulations that reduce the (expected) time till a patient is matched, or decreased the (expected) KDPI of the organ to which a patient is matched. Analysis of time till matching or the quality of the matching is slightly more complex as we need to take account of patients that are never matched. We can, for instance, give them a "virtual" organ with a KDPI of 101 at the time of their departure or at the end of the market if they remain till then.

We say that a mechanism for online organ matching is **strategyproof** if and only if for any matching market, no patient can improve their probability of being matched to an organ by declaring a false EPTS. We consider **responsive** mechanisms where there is at least one market where one agent can change the probability of receiving an organ by declaring a different EPTS. For example, both MIN and BOX are responsive. To return stable matchings other than by chance, a mechanism needs to be responsive. We contrast this with **irresponsive** mechanisms where in no market can any agent change their probability of receiving an organ by declaring a different EPTS. For example, the mechanism that matches organs uniformly at random amongst compatible patients is irresponsive. As a second example, the mechanism that matches an arriving organ to the oldest compatible patient on the waiting list is also irresponsive. Before the introduction of KDPI and EPTS in the US, matching was largely driven by time on waiting list. The following very simple result demonstrates that responsiveness rather unsurprisingly leads to the possibility of manipulation.

Theorem 4 *A mechanism for online organ matching is strategyproof if and only if it is irresponsive.*

Proof. Suppose a mechanism is responsive. Then there exists a market on which one patient declaring a new and different EPTS at some point changes the probability of this patient being matched.

Without loss of generality, we can suppose that it increases, otherwise we simply swap the new EPTS for the old. Hence, the mechanism is not strategy-proof.

Suppose a mechanism is irresponsive. Then in any market, declaring a different EPTS does not change the probability of being matched. Since the probability of being matched does not change, it cannot improve. Hence the mechanism is strategy-proof. □

Irresponsive mechanisms are undesirable since they will likely return matches in which EPTS and KDPI are not aligned. This simple result does not mean manipulations cannot be avoided. We can still identify some barriers to manipulation. Two possible barriers are secrecy and computational complexity. If we can keep the waiting list secret, then a risk averse patient may have an incentive to be sincere. Similarly, even if the waiting list is public, we may be able to construct a mechanism where it is computationally intractable to compute a successful manipulation. Similar computational barriers have been proposed in other areas of social choice [8] and fair division [22].

7 Fairness in Organ Allocation

The charter of the Organ and Tissue Authority of Australia requires that the mechanism employed be both fair and efficient, though it is not formally defined what this means. In this section we consider three notions of fairness: age, geography, and blood type. Due to space constraints we have moved the results about blood type to the appendix. In general, with fairness to blood type, we see that depending on definition, it may be necessary to restrict intra-blood type flows to ensure an equal allotment of organs.

7.1 Fairness to Age

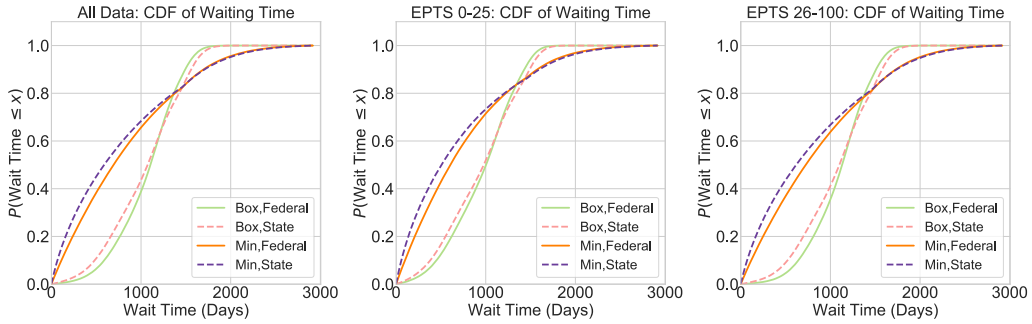


Figure 2: Cumulative distribution function (CDF) of waiting time for all patients and broken down to those with a low and high EPTS. Under the BOX mechanism we see a statistically significant movement to the right as EPTS increases, hence older patients must wait longer to receive their transplants.

We first consider the fairness of the two proposed matching mechanisms. We argue that, unlike the BOX mechanism, the MIN mechanism is procedurally fair to patients of different ages. Figure 2 we plot the cumulative distribution function (CDF) of waiting time for patients using the two mechanisms in a simulation as discussed in Section 8. When an organ arrives, it is matched nationally with a compatible patient using either the BOX or MIN mechanisms on a State or Federal allocation.

We see that the BOX mechanism is not procedurally fair. Those patients on the waiting list with an EPTS of 25 or less spend less time waiting than those with an EPTS of greater than 25. This is to be expected as the BOX mechanism preferentially favors patients with an EPTS of 25 or less. By comparison, the MIN mechanism is much more procedurally fair. The time on the waiting list is less dependent on EPTS (i.e., near the line y=x). Waiting time with the MIN mechanism is almost constant as is expected given that EPTS and KDPI are population percentiles.

7.2 Fairness to Geography

Matching at the national, rather than the state or hospital level increases the pool of potential donors. Especially for rarer blood types, matching that takes account of KDPI and EPTS will be better if we can use the larger national pool. We expect matches to be better nationally as we can perform all state level matches as well as those now possible nationally. There is a time penalty to matching nationally, however, it only takes about five hours to fly an organ across the country and this has no ill effect on the outcomes. There is thus little to be lost, and much to be gained, if the states and territories can be persuaded to match nationally.

In Figure 4, we plot the distribution of donors and patients waiting for transplant according to state and territory. We compare this with the distribution of the population within Australia. Donation largely tracks population, as might be expected. The major exception is New South Wales (NSW) and the Australian Capital Territory (ACT).[3] Donation rates in NSW/ACT are much lower than in the rest of Australia. To compound this issue, the waiting list in NSW/ACT is proportionally much longer than in the other states and territories. In part, this may reflect that doctors are more likely to list patients on the waiting list in New South Wales than in a state like Queensland (QLD) where the waiting list is smaller proportional to the population. The data may also reflect that patients gravitate towards the more sophisticated medical facilities available in a populous state like New South Wales.

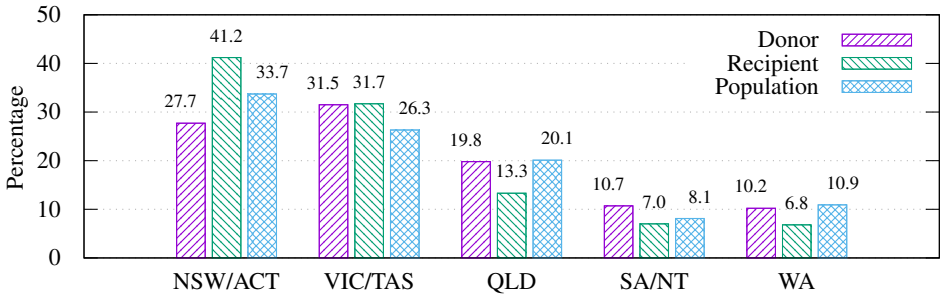


Figure 3: Distribution across states of donated organs, patients waiting transplant, and the wider population of Australia from 2010 to 2014.

There is concern, especially amongst the smaller states like South Australia (SA), Western Australia (WA) and the Northern Territory (NT), that donated organs will flow into New South Wales due to its comparatively much longer waiting list. In 2014, there were 614 kidneys transplanted nationally. If we had matched nationally, rather than at the state level, we could expect on aggregate that 21 out of the 63 organs donated in Western Australia would flow out of the state (exactly one third), 23 out of the 66 donated organs would flow out of South Australia and the North Territory (slightly over one third), 40 out of the 122 organs donated in Queensland would flow out of state (just under one third). Almost all of the inflow of organs would be to New South Wales. Only one of the organs flowing out of SA/WA/NT/QLD would be expected on aggregate to end up in Victoria (VIC). The other 83 organs flowing between states and territories would end up on aggregate being given to patients in New South Wales. This inflow of 83 organs into New South Wales represents 13.5% of the total number of deceased organs donated nationwide.

Matching nationally ensures that everyone gets the same chance of a match irrespective of geography. This is very far from the case currently. In 2014, for example, the waiting list in New South Wales contained 474 patients at the start of the year, and increased to 500 by the end. Only 152 of the patients on the waiting list received a deceased organ. By comparison, in South Australia, the waiting list began and ended the year with 64 patients on it. During the course of the year, patients

[3]There are no facilities for kidney transplant in ACT so all ACT patients are dealt with in NSW.

were added to and left the waiting list[4] but a total of 67 patients received a deceased organ. Patients waitlisted in South Australia thus have a much greater chance of receiving a transplant than in New South Wales. On average, patients waitlisted in South Australia wait about one year for transplant whilst those in New South Wales wait around three.

Matching nationally would result in greater fairness as waiting times across states would become more equal. On the other hand, matching nationally would result in a flux of organs out of the smaller states into New South Wales. The only way to prevent this is to harmonize donation rates within states and a campaign to this effect is underway.

8 Experimental Analysis

From an axiomatic perspective, online organ matching appears to be a challenging problem. In this section, we run some experiments using historical data to determine whether ideals including efficiency, stability, and fairness are achieved in practice when allocations are run at both a state and federal level. The Scientific Registry of Transplant Recipients publishes detailed models of expected donations and transplant survival times for the US Market, see `www.srtr.org` but no such detailed data exists for the Australian market. In what follows we generate patients and kidneys using some of the same techniques but tuned for the Australian market. We do not go into detail on post transplant survival as we are more concerned with the properties of the initial allocation. Subjecting this allocation to optimization under the expectation of future survival times is an interesting avenue for future work with some recent initial investigation [13].

Data Generation. Since we have only one historical data set, we first need to build a simulator that is as realistic as possible so we can run thousands of experiments with realistic data. This relies on two sources of data provided by the Organ and Tissue Authority of Australia: historical statistics published as the Public ANZDATA, `www.anzdata.org.au` and more detailed statistics provided to us as the Research ANZDATA which includes detailed information about the Australian market from 2010 – 2014. We cross validate this data with additional information from the Australian Bureau of Statistics; the Public and Research sets; and Wikipedia, we determined the probability distribution over the *Blood Type* and *State* of donors and patients.

We generate EPTS for the patients and KDPI for the kidneys according to the distributions contained in the Research ANZDATA for the years 2010–2014 with a small amount of random noise added for the least significant digit. We use a Poisson process to simulate arrival of patients to the market. Specifically, the number of patients arriving each day to the market is $P(x) = e^{-\mu} \cdot \mu^x / x!$. We use the long run average of 340 donors a year giving us a $\mu = 340/365$. Each patient that arrives donates up to 2 identical kidneys with the average being 1.72. The waiting list in Australia is held relatively constant at ≈ 1200 people. When kidneys are transplanted, new patients replace those that depart the market with a match. On average 150 patients are removed from the waiting list randomly throughout the year for a number of reasons including death and off list transplant. We model removal from the list again as Poisson process with $\mu = 150/365$; a new patient is generated to replace those removed. The EPTS of patients degrades over time for a number of reasons and ≈ 180 patients are removed from the list each year due to their EPTS becoming too high. We increase the EPTS of patients every 30 days between 1–2 points, uniformly at random, and remove patients whose EPTS increases past 100. Our simulator, along with one for the US kidney allocation market, is available at `www.preflib.org` [26].

Experimental Treatments and Evaluation. For our experiments we generate an initial list of patients and kidneys and then simulate the arrivals, departures, and donations by stepping through a simulated 8 years. We repeat this process 1000 times to gain confidence in the statistics we report here [12]. We use the first four years of data to burn in our simulator, so that the allocation has

[4] During 2014, one patient died waiting in South Australia, two received an organ from a living donor, 9 were taken off the list for medical and other reasons, and 79 new patients were added.

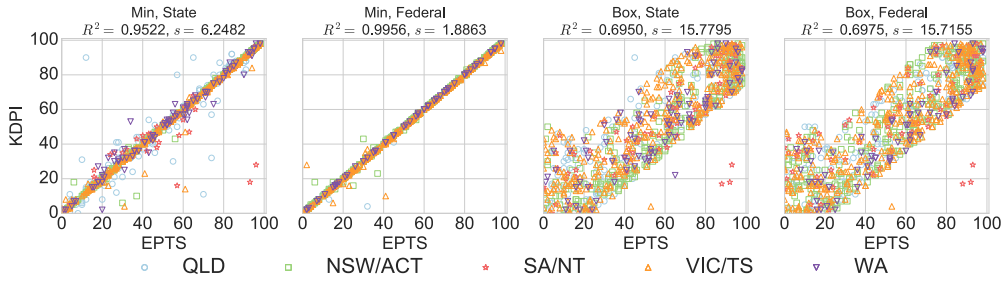


Figure 4: Visualization of the results of one run of our algorithm for both the MIN and BOX algorithms with both the State and Federal restrictions. MIN performs significantly better than BOX in terms of minimizing $|KDPI - EPTS|$ and the Federal allocation outperforms the State allocations in ensuring better outcomes.

reached a steady state, and report statistics based on the latter four years of data. The same list of kidneys and patients (their order of arrival) are used for all of the treatments (both mechanisms and both state and federal allocations). Hence, the difference in statistics are generated only by the particular matching strategy.

For all experiments we enforce basic blood type compatibility. That is, A can donate to A and AB, B to B and AB, AB to AB, and O can donate to anyone. We also differentiate between two treatments regarding the level the allocation takes place at. In a **Federal** allocation, a kidney can be transplanted to any state in Australia. We enforce that in the case of a tie, the kidney goes to the instate patient. In a **State** allocation all allocations are run within states only. Organs not able to be transplanted in state due to type restrictions are then matched to any patient outside the state where the kidney originated. Ties for all mechanisms are broken by state, time on list, and then randomly.

Ideally we would transplant every organ into a patient such that KDPI and EPTS are equivalent. This gives us a notion of optimality that is defined by the line $y = x$. We use two error metrics to judge how well our points match this idealized line. Note that since we have regression through the origin here we must be careful with computing our statistics [12]. Firstly, the **Coefficient of Determination** (R^2), if we treat KDPI as a dependent variable w.r.t. ETPS then we can use R^2 to get an indication of the proportion of the variance of KDPI that is predictable from EPTS. If we want to ensure that KDPI and EPTS are perfectly correlated ($x = y$) then we would get an $R^2 = 1$. We compute: $R^2 = 1 - \sum_i (EPTS_i - KDPI_i)^2 / \sum_i (KDPI_i - \overline{KDPI})^2$. Secondly, we use the **Standard Error of the Regression** (s), since s is expressed in the same units as the independent variable (KDPI) it gives us an intuitive measure of how how much KDPI varies, on average, from EPTS. Since we are forcing our regression through the origin we only have one degree of freedom in the model giving $s = \sqrt{\frac{1}{n-1} \sum_i (KDPI_i - EPTS_i)^2}$.

8.1 Efficiency, Waiting Time, and Participation

Figure 4 shows the results for one simulation, which is typical for all. We see that, no matter the state, the patient results are better for the MIN algorithm and Federal matching. MIN outperforms BOX by a statistically significant margin for both the R^2 measure and s for all 1000 instances that we tested. For the MIN algorithm, Federal significantly outperforms State, across all instances.

One concern with the Federal v. State treatments on the practical level is the flow of organs out of a state will exceed the flow of organs into a state. In effects, states do not want to lose organs to other states. When tracking organ flow we use the term inflow to mean organs that come into a particular state or type and outflow to mean organs that move to a different state or type. A flow between organ types means that, for instance, an O kidney was donated to an AB patient. Figure 8a (moved to the appendix) illustrates the flow of organs between states and we can see that the majority of organs are allocationd within the state with the exception of Western Australia. The cost in terms of efficiency

		Mean Number of Swaps		Mean Number of Patients		Mean Gain in Welfare	
MIN	Federal	3.90	(3.4)	3.59	(3.0)	0.1	(0.01)
	State	314.12	(53.9)	181.4	(22.6)	3.81	(0.85)
BOX	Federal	24,548.6	(2,233.1)	937.83	(30.87)	14.18	(0.34)
	State	23,555.6	(2,173.6)	934.62	(29.79)	14.18	(0.32)

Table 1: Mean number of Pareto improving swaps per patient, mean number of patients who could participate in a swap, and mean gain in welfare (standard deviation) for the four treatments. The MIN, Federal algorithm achieves the greatest efficiency and lowest frequency of opportunities for Pareto improving trades.

to the overall system, illustrated in Figure 4, is a much greater concern than organ flows.

Time on List. One of the main concerns for patients is the amount of time that they expect to spend on the waiting list. To investigate this we computed the mean waiting time in days by EPTS, the results for all states and for NSW/ACT are broken out in Figure 9a (moved to the appendix). We have omitted variance/error bars for this graph as they completely overlap. In general, the variance is constant for each of the mechanisms and treatments with the MIN mechanism having a lower variance of about ± 200 days while the BOX mechanism has a higher variance of about ± 600 days. We observe that the variance for the BOX algorithm is strictly higher than the variance for the MIN algorithm across all treatments.

Looking at Figure 9a we see that the MIN mechanism has a lower mean waiting time versus BOX except for patients with very high EPTS. This gap closes for patients with higher EPTS, likely due to the fewer organs that are donated with very high EPTS. It is interesting to note that the State v. Federal question is roughly negligible for mean time on list. Consequently, Federal allocations increase patient welfare without a significant impact on waiting time. The BOX has large equivalence classes between patients, illustrated in Figure 1. We conjectured that these large equivalence classes would give lower wait times as the mechanisms tie-break based on the time on the waiting list. However, this is not the behavior that we see in the data except for those high EPTS patients.

Efficiency. To investigate the efficiency of the various mechanisms in practice we investigate the number of Pareto swaps [11], i.e., between kidney/patient pairs such that the welfare of one of the pairs is increased while the other is not decreased. To do this we took, for each of the 1000 iterations, the complete set of transplants and checked to see for each kidney/patient if there exists another kidney/patient that arrives later in time that would enable a Pareto swap. Since one kidney/patient pair may be involved in multiple potential swaps, and a welfare efficient matching would only use at most one of these swaps *per patient*, we also compute the number of unique kidney/patient pairs that could be part of a Pareto improving swap. We also track the average increase in welfare over the set of potential swaps per instance. These values are reported in Table 1.

The metric we compute is a worst case one. We do not take into account the aging factor for the later swap. Hence, the number of swaps reported is an upper bound and would be lower, in reality, if we aged the patient participating in the swap. Additionally, kidneys need to be matched on the same day they are donated so these swaps are not even realistically possible. In the experiments for MIN, there were no improving swaps possible on the same day (though there were some for BOX). We report improving swaps that could have happened on *any day* to give an upper bound on the worst case that could happen if we could store the kidneys.

The lower the numbers across the board in Table 1 for the MIN algorithm shows that it is doing much better in terms of welfare efficiency than BOX. It is encouraging to see that the Federal allocations result in significantly fewer patients who would benefit from a Pareto swap. Additional evidence for the quality of the MIN algorithm can be seen in the CDF of $|KDPI - EPTS|$ over all 1000 runs seen in Figure 8b; the clear winner is MIN with the Federal treatment. There is a vanish-

ingly small probability of observing a value of $|KDPI - EPTS|$ larger than 10. Turning to the right side of 8b, we see that, in fact, most of the efficiency loss is due to AB organ transplants. As there are so few AB organs, it is hard to match them in way that maximizes welfare.

9 Related Work

Online problems have been studied primarily in online scheduling [2]. Mechanism design concerns have been extensively addressed in the online scheduling literature [30] though online allocation has received less attention. Online matching markets without money have been a common area of study in computer science. Gujar and Parkes [20] study an online matching market where there is no money, much like organ allocation, though only one side of their market is dynamic. Another online matching market was studied by Bosek *et al.* [10]. In their model the market incrementally increases in size and at each increase, a new maximum stable matching must be found. In our work we are not concerned with stability but fairness. Additionally, in online organ matching, unlike general online matching, all agents have a shared preference model and *both* sides of the market are dynamic. Finally, Aleksandrov *et al.* [3] consider the online allocation problem faced by foodbanks, a related online matching problem where only one side of the market is dynamic.

Kidney *exchanges* have been extensively studied in the economics [33, 32], medical [29], and computer science literatures [15]. These exchanges have also been studied in online settings [6] and in predictive settings using machine learning techniques [13]. Axiomatic and fairness concerns have received attention in the computer science literature in both theory and practice [16], as well as the online feature that such exchanges are repeated over time [6, 13]. Additional work in computer science has also focused on strategyproof mechanisms for kidney exchange at the patient and hospital level [4] and on merging kidney exchanges with other organs such as lungs [24] and livers [14]. However, as we have pointed out, the majority of kidney donations are performed from deceased patients and the online version of the deceased donor has historically received little attention.

A related line of work is determining what factors should be included in the scores such as KDPI and EPTS. Bertsimas *et al.* [9] study the problem of using data driven methods for finding national organ allocation policies. In general, they adopt a statistical approach that designs policies that work well on average based on historical data. By comparison, we have taken a more axiomatic approach. The Organ and Tissue Authority in Australia wish to use a new mechanism based solely on blood/tissue type and KDPI/EPTS. This prevents an approach like Bertsimas *et al.* [9] where we compute weights for terms going into KDPI/EPTS.

In proposing a new mechanism, we cannot overlook the challenging issues of managing the transition [1]. The US has already adopted a matching mechanism that takes account of KDPI and EPTS and others including the National Kidney Allocation Scheme [21] introduced in the UK in 2006 has as well. Interestingly, the UK scheme is a points system that uses a number of features including blood group points which, as we have proposed in considering fairness across blood type in Australia, end up allocating a proportion of O type kidneys to B type recipients. It may be difficult to persuade patients to buy into a new mechanism if they will be treated differently and we have shown that in many respects, the MIN mechanism treats most patients the same in terms of waiting time but results in more efficient outcomes, though there will be some individual winners and losers.

10 Conclusions

We have proposed the novel MIN mechanism for the online matching of deceased organs to donors and compared this both theoretically and empirically with the BOX under consideration in Australia. We argued that fairness needs to be considered on multiple levels: patient, region, and blood type to name just three. We investigated other axioms including stability and strategyproofness and showed that despite the dynamic nature of this two-sided market, a stable allocation always exists, and that

the MIN mechanism is likely to find an allocation close to this. We find that the MIN algorithm outperforms the proposed BOX mechanism across all studied areas. For the future, mitigating temptation for strategic behavior remains an important dimension of the policy-making in this area as does extending our studies to incorporate post-transplant success.

References

[1] M. Ahlert and H. Kliemt. Problems of priority change in kidney allocation and beyond. *The European Journal of Health Economics*, 14(3):383–390, Jun 2013.

[2] S. Albers. Online algorithms: A survey. *Mathematical Programming*, 97(1-2):3–26, 2003.

[3] M. Aleksandrov, H. Aziz, S. Gaspers, and T. Walsh. Online fair division: Analysing a food bank problem. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.

[4] I. Ashlagi, F. Fischer, I. A. Kash, and A. D. Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 91:284–296, 2015.

[5] Australia & New Zealand Dialysis & Transplant Registry. 38th annual ANZDATA report, 2015. Available from `http://www.anzdata.org.au/`.

[6] P. Awasthi and T. Sandholm. Online stochastic optimization in the large: Application to kidney exchange. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, 2009.

[7] H. Aziz, C. Cahan, C. Gretton, P. Kilby, N. Mattei, and T. Walsh. A study of proxies for Shapley allocations of transport costs. *Journal of Artificial Intelligence Research*, 56:573–611, 2016.

[8] J.J. Bartholdi and J.B. Orlin. Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8(4):341–354, 1991.

[9] D. Bertsimas, V. F. Farias, and N. Trichakis. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, 61(1):73–87, 2013.

[10] B. Bosek, D. Leniowski, P. Sankowski, and A. Zych. Online bipartite matching in offline time. In *Proceedings of the 55th Symposium on Foundations of Computer Science (FOCS)*, pages 384–393, 2014.

[11] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[12] P. R. Cohen. *Empirical Methods for Artificial Intelligence*. MIT Press, 1995.

[13] J. P. Dickerson and T. Sandholm. Futurematch: Combining human value judgments and machine learning to match in dynamic environments. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pages 622–628, 2015.

[14] J. P. Dickerson and T. Sandholm. Multi-organ exchange: The whole is greater than the sum of its parts. *Journal of Artificial Intelligence Research*, 2016.

[15] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Optimizing kidney exchange with transplant chains: Theory and reality. In *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 711–718, 2012.

[16] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Price of fairness in kidney exchange. In *Proceedings of the 13th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1013–1020, 2014.

[17] R. Freedman, J. Schaich Borg, W. Sinnott-Armstrong, J. P. Dickerson, and V. Conitzer. Adapting a kidney exchange algorithm to align with human values. In *Proceedings of the 1st AAAI/ACM Conference on AI, Ethics, and Society*, 2018.

[18] D. Gale and L.S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69:9–15, 1962.

[19] P. Glander, K. Budde, D. Schmidt, T.F. Fuller, M. Giessing, H-H. Neumayer, and L. Liefeldt. The 'blood group O problem' in kidney transplantation—time to change? *Nephrology Dialysis Transplantation*, 25(6):1998–2004, 2010.

[20] S. Gujar and D. Parkes. Dynamic matching with a fall-back option. In *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI)*, pages 263–268, 2010.

[21] R. Johnson, S.V. Fuggle, L. Mumford, J.A. Bradley, J.L. Forsythe, and C.J. Rudge. A new UK 2006 national kidney allocation scheme for deceased heart-beating donor kidneys. *Transplantation*, 89(4):387–394, 2010.

[22] T. Kalinowski, N. Narodytska, T. Walsh, and L. Xia. Strategic behavior when allocating indivisible goods sequentially. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI 2013)*. AAAI Press, 2013.

[23] S.R. Kaufman. Fairness and the tyranny of potential in kidney transplantation. *Current Anthropology*, 54(S7):S56–S66, 2013.

[24] S. Luo and P. Tang. Mechanism design and implementation for lung exchange. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 209–215, 2015.

[25] M. Maschler, E. Solan, and S. Zamir. *Game Theory*. Cambridge University Press, 2013.

[26] N. Mattei and T. Walsh. Preflib: A library for preferences, HTTP://WWW.PREFLIB.ORG. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*, 2013.

[27] N. Mattei, A. Saffidine, and T. Walsh. Mechanisms for online organ matching. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 2017.

[28] N. Mattei, A. Saffidine, and T. Walsh. Fairness in deceased organ matching. In *Proceedings of the 1st AAAI/ACM Conference on AI, Ethics, and Society*, 2018.

[29] R. A. Montgomery, S. E. Gentry, W. H. Marks, D. S. Warren, J. Hiller, J. Houp, A. A. Zachary, J. K. Melancon, W. R. Maley, H. Rabb, C. Simpkins, and D. L. Segev. Domino paired kidney donation: a strategy to make best use of live non-directed donation. *The Lancet*, 368(9533):419, 2006.

[30] R. Porter. Mechanism design for online real-time scheduling. In *Proceedings of the 5th ACM Conference on Electronic Commerce (ACM-EC)*, pages 61–70, 2004.

[31] A. E. Roth and M. A. O. Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, 1992.

[32] A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488, 2004.

[33] A. E. Roth, T. Sönmez, and M. U. Ünver. Pairwise kidney exchange. *Journal of Economic theory*, 125(2):151–188, 2005.

[34] T. Walsh. Challenges in resource and cost allocation. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pages 4073–4077, 2015.

Nicholas Mattei
IBM Research
New York, USA
Email: n.mattei@ibm.com

Abdallah Saffidine
UNSW Sydney
Sydney, Australia
Email: abdallah.saffidine@gmail.com

Toby Walsh
UNSW Sydney, Data61, and TU Berlin
Sydney, Australia
Email: tw@cse.unsw.edu.au

A Fairness

A.1 Fairness to Blood Type

Reducing the pool of candidate matches by considering the age of organs and patients creates fresh challenges especially for rarer blood types. How do we treat different blood types fairly? We could, for example, permit organs coming from donors of blood type O to be transplanted into patients of blood type B since blood type O are universal donors. But this would disdvantage patients of blood type O waiting for transplant. Patients of blood type B would be matched out of the larger pool of organs of blood type O and B, whilst patients of blood type O would be matched out of the smaller pool of organs of blood type O only. This phenomenon has been observed in the European transplant market [19]. On the other hand, there might be a greater demand for organs of blood type B relative to supply compared to blood type O. Transplanting a small number of organs from donors of blood type O into patients of blood type B could help to correct any imbalance.

Another concern when moving to a new mechanism based on age of organs and age of patients is that the quality of matching for rarer blood types will decrease as the pool size decreases since some matches will be are ruled out based on age considerations. Historical data demonstrates where challenges might arise in the Australian market. In Figure 6, we plot the distribution of donors and patients waiting for transplant according to blood type. We compare this with the distribution of blood types within the population. The exact blood type distribution in the population as a whole is not tracked, but we do have the distribution in the subset of the population donating blood.

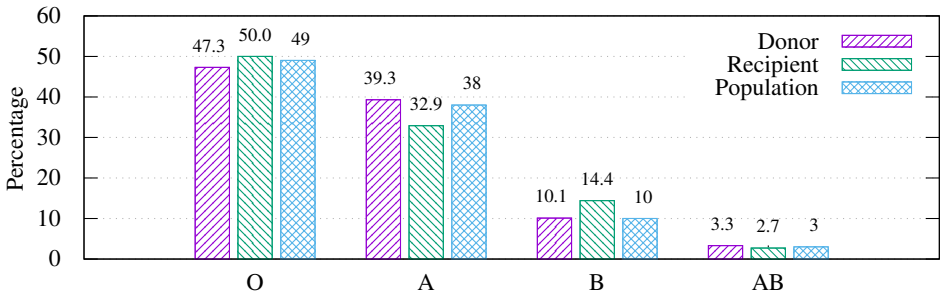


Figure 5: Distribution of blood types in donated organs, in patients waiting transplant, and in the wider population for Australia from 2010 to 2014.

Donation tracks population quite well. Those ethnic and other groups willing to donate their blood are perhaps also likely to be those willing to donate their organs. However, demand is somewhat different. In particular, there is a greater percentage of patients of blood type B waiting transplant than donation. 14.4% of the waiting list have blood type B yet only 10.1% of organs donated are blood type B. On the other hand, patients of blood type A are better off. 39.3% of donated organs are of blood type A, yet only 32.9% of the waiting list have this blood type. There is some medical evidence that people of blood type A are less prone to kidney disease.

To see how we might improve fairness across blood types, we set up a simple linear model. Let p_t be the fraction of the patients with blood type t. For instance, p_A and p_{AB} are the fraction with blood type A and AB respectively. Let o_t be the fraction of the organs with blood type t. For instance, o_A and o_{AB} are the fraction with blood type A and AB respectively. We suppose matching takes into account other factors like HLA type but consider here just the size of the pool from which possible matches are drawn.

We suppose that within our matching procedure a fraction x_{t_1,t_2} of the donated organs are of type t_1 and are considered for transplant to patients of blood type t_2 subject to HLA match and other factors like age. Thus $0 \leq x_{t_1,t_2} \leq o_{t_1}$. We insist on blood type compatibility for transplant. Hence, $x_{A,O} = x_{A,B} = x_{B,O} = x_{B,A} = x_{AB,O} = x_{AB,A} = x_{AB,B} = 0$. That is, we cannot transplant organs of type A to patients of type O, organs of type A to patients of type B, etc.

We have a conservation law for organs of each blood type. This requires:

$$\begin{aligned} x_{O,O} + x_{O,A} + x_{O,B} + x_{O,AB} &= o_O \\ x_{A,A} + x_{A,AB} &= o_A \\ x_{B,B} + x_{B,AB} &= o_B \\ x_{AB,AB} &= o_{AB}. \end{aligned}$$

We suppose that the mechanism is fairest when the fraction of organs available for a given blood type is as close as possible to the fraction of patients of this blood type. We introduce therefore some variables to measure this. Let z_t be the ratio of the fraction of organs available for blood type t and the fraction of patients of blood type t.

$$\begin{aligned} z_O &= \frac{x_{O,O}}{p_O} \\ z_A &= \frac{x_{O,A} + x_{A,A}}{p_A} \\ z_B &= \frac{x_{O,B} + x_{B,B}}{p_B} \\ z_{AB} &= \frac{x_{O,AB} + x_{A,AB} + x_{B,AB} + x_{AB,AB}}{p_{AB}}. \end{aligned}$$

To maximize fairness, we consider an egalitarian objective in which we maximize $z = min(z_0, z_A, z_B, z_{AB})$. This can be solved in polynomial time using linear programming. Let's consider how this model fares on the historical Australian data. We have the following input data.

	O	A	B	AB
o_t	0.473	0.393	0.101	0.033
p_t	0.500	0.329	0.144	0.027

Maximizing fairness gives the following organ fractions.

x_{t_1,t_2}	O	A	B	AB
O	0.446	0.000	0.027	0.000
A		0.393	0.000	0.000
B			0.101	0.000
AB				0.033

And the corresponding z fractions.

	O	A	B	AB
z_t	0.89	1.19	0.89	1.22

Because of the drowning effect of the minimum function, there are multiple solutions with the same maximal minimum z value. To minimize transfer between blood types, we chose the solution with the maximal number of decision variables x_{t_1,t_2} for $t_1 \neq t_2$ set to zero. In this case, the only organs matched across blood type are organ of blood type O which are matched with patients of blood type B.

There are a number of ways we could translate this into practice. The simplest would simply be to toss a coin when a new organ arrives to decide which groups of patients is used in its matching. In this case, when an organ arrives of blood type O, with probability $\frac{x_{O,O}}{o_O}$, we consider patients of blood type O (that is, with probability 0.943), and otherwise (with probability $1 - \frac{x_{O,O}}{o_O} = \frac{x_{O,B}}{o_O} = 0.057$) we consider patients of blood type B. In short, with a 5.7% chance, we transfer an organ of type O to a patient of type B, otherwise we match organs to patients of the same blood type.

Note that we have been unable to achieve complete fairness as the z values are not identical. We can equalize the treatment of patients of blood type O and blood type B (that is, we can equalize z_O and z_B). But these blood types are at a disadvantage compared to blood types A and AB (since z_A and z_{AB} are larger). We cannot use the relative excess of organs of blood type A and AB to help the relative excess of patients of blood type O and B. We simply need more organs of blood type O and B to give to patients of blood type O and B. Blood type A is also at a slight disadvantage compared to AB (since z_A is smaller than z_{AB}). We also cannot fix this problem by transferring organs between blood types. This illustrates a fundamental impossibility to be fair to the different blood types. As organs can only be transferred across blood type in one direction, there will be online organ matching problems, like the one in Australia, where we cannot treat patients of different blood type equivalently. Based on this analysis, we have advised the Organ and Tissue Authority to consider a publicity campaign to increase donation of organs from members of the public with blood types O and B.

A.2 Flow of Organs Between States, 2010–2014

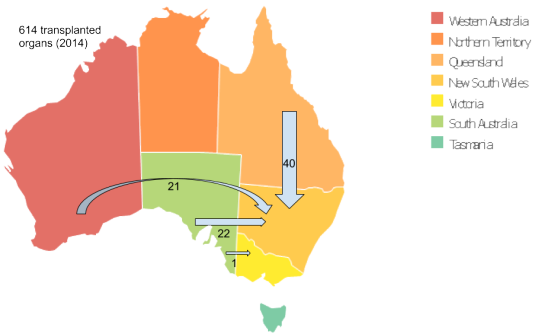


Figure 6: Aggregate flow of organs between states and territories in 2014 if matching had been nationally.

B Additional Experimental Results

In this section we provide an analysis of stability and manipulation in practice, along with additional graph and results related to the mechanisms.

B.1 Stability in Practice

		Number of blocking pairs	% patients affected	ΔKDPI	ΔKDPI%	ΔEPTS	ΔEPTS%
MIN	Federal	121 (77)	1.1% (0.6)	4.5 (3.0)	15.0% (8.7)	4.7 (3.1)	14.1% (6.9)
	State	378 (122)	6.0% (1.5)	12.0 (2.8)	27.5% (5.2)	12.4 (2.8)	22.8% (4.3)
BOX	Federal	23,312 (1,664)	65.4% (1.4)	11.7 (0.4)	25.9% (0.8)	11.6 (0.4)	24.2% (0.9)
	State	5,445 (398.1)	47.4% (1.4)	11.4 (0.4)	25.8% (0.9)	11.3 (0.4)	23.9% (1.0)

Table 2: Mean and standard deviation over 1000 trial of the total number of blocking pairs, percentage of total patients that have incentive to form a blocking pair, absolute change in KDPI for the patient, percentage change in KDPI for the patient, absolute change in EPTS for the kidney, and percentage change in EPTS for the kidney under the MIN and BOX mechanisms. The Federal, MIN mechanism provides a significantly ($p \leq 0.05$) more stable match with a smaller window for improvement than all other mechanisms. Under the BOX mechanism nearly 50% of patients have an incentive to form a blocking pair.

We now empirically check if the matching returned by the MIN or BOX mechanism is stable in practice. A matching is *stable* if for every kidney (resp. patient) there is no other patient (resp. kidney) to whom the kidney would prefer to be matched *and* that patient would prefer to be matched to this kidney. Informally, there exists no kidney/patient pair that would prefer to leave the market together rather than be matched. In this experiment we only consider strict stability, meaning that both the kidney and the patient must strictly prefer each other to their current match.

Formally, a matching μ is a list of pairs (p,k) indicating that patient p is matched to kidney k. We use the notation $\mu(k)$ to denote the patient matched to k and $\mu(p)$ to denote the kidney matched to p. A matching is stable if there exists no pair $(p,k) \notin \mu$ where k prefers p to its current match in μ, i.e., $p \succ_k \mu(p)$ where $\succ_k$ is the preference of k corresponding to the EPTS of p is strictly less than $\mu(p)$. The definition is symmetric for patients.

In order to measure the stability of the matching returned by the two algorithms we search for instances of blocking pairs that *could* be formed. For simplicity, we ignore patients who leave list without a match as we do not know their ultimate fate. Many of these patients become to sick to receive a transplant (or even die) so are unlikely to be a blocking pair. Others receive a satisfactory kidney elsewhere so are also unlikely to be a blocking pair. To find blocking pairs we track the EPTS of all patients until they would have aged out of the market, i.e., their EPTS exceeds 100. We ignore the possibility that they could have been randomly removed in the window after they received their match. Hence, our results are a worst case measure since there a patient may have left the market earlier than the time taken to age out, decreasing opportunities to form a blocking pair.

For each patient/kidney pair (p,k) that is matched at time t, we check from t until p would have aged out of the market if there arrived another kidney k' that was matched to p' where $k' \succ_p k$ and $p \succ_k p'$. Hence, (p,k') would have formed a blocking pair and the final match is not stable. We assume that p would have won all tie-breakers over p' but we assume that both the kidney and the patient must strictly prefer to be part of the blocking pair. We track the mean number of blocking pairs and the mean change (Δ) that the patients and kidneys would have by blocking. Results are given in Table 2.

In our experiments, the MIN mechanisms provide a significantly more stable matching for both the State and Federal treatments ($p \leq 0.05$) based on a t-test corrected for correlated samples [12]. The Federal, MIN matching has significantly fewer blocking pairs and significantly smaller room for improvement for both the patients and kidneys compared to all other treatments. In addition, under the Federal, MIN mechanism only $\approx 1\%$ of patients could form a blocking pair, implying a very stable match. Interestingly, the scope for improvement for the State, MIN and both treatments of the BOX mechanism are very similar. We note that the matches returned by the BOX mechanism

are not very stable as nearly half the patients have opportunities to form blocking pairs that would significantly increase the quality of their match.

B.2 Manipulation in Practice

We now consider strategic behavior on the part of the patients and doctors who, as in the case in Germany, may misreport their EPTS. We focus on two questions: (1) How does lowering or increasing the claimed EPTS impact on the quality of the matching? (2) How risky is misreporting your EPTS as it might increase the chance of a matching, it may lower the quality of this match?

In the Germany scandal in 2012, doctors increased the apparent EPTS of their patients by mixing blood in their urine samples. It is also possible for a lower EPTS to be reported than is sincere. For instance, the guidelines of the Organ Procurement and Transplantation Network in the US remind doctors of their duty to update reported EPTS scores when the status of patients change (e.g. they receive a liver transplant, or they start dialysis). Intuitively, since both the MIN and BOX mechanisms attempt to match low EPTS patients with low KDPI organs, we might expect that decreasing a patient's EPTS should be more often helpful than increasing it. Since misreporting the EPTS of a patient by a large amount may be easily detectable, we consider a *margin* parameter such that an insincere patient is allowed to claim any number between EPTS $-$ margin and EPTS $+$ margin.

In our experiments, we check for combinations of margin $\in [1,100]$ and actions increasing, decreasing, or both, computing the proportion of patients who could have improved their outcome by claiming a different EPTS. In each case, we focus on the list of patients who obtained a match in the truthful case. We report the mean proportion and standard deviation over 420 trials for margins in $\{1,3,10,30,100\}$ in Table 3.

	Margin	Increase	Decrease	Either
State	1	36.7±2.0	39.0±2.2	52.4±2.3
	3	49.3±2.2	58.8±2.3	70.0±2.2
	10	61.0±2.2	79.8±2.1	85.6±1.8
	30	63.6±2.1	90.2±1.5	93.1±1.2
	100	63.8±2.1	93.6±1.0	96.1±0.7
Federal	1	34.2±2.3	40.2±2.6	53.7±2.7
	3	51.8±2.3	64.2±2.6	74.8±2.4
	10	63.6±2.2	84.7±2.0	89.1±1.8
	30	65.7±2.2	93.6±1.3	95.5±1.1
	100	65.8±2.1	96.2±0.7	97.9±0.6

Table 3: Percentage of patients who could improve their assigned KDPI in MIN matching by adding, subtracting, or either adding or subtracting up to "margin" from their true EPTS.

The results suggest, as might be expected, that increasing the misreporting margin increases the proportion of patients who can improve the quality of their match. Decreasing the reported EPTS offers more opportunity for improvement to more agents than increasing it. On the other hand, since the "Either" column has higher proportions than the "Decrease" column, there are situations where a small decrease in the reported EPTS cannot help, but an increase can. Finally, we observe that strategic manipulation is easier in Federal, MIN than in State, MIN matching; this is not surprising as, in a larger market, there is a bigger trade-off possible between efficiency and manipulability.

Even if we discard the "margin = 30" and "margin = 100" cases as unrealistic or unfeasible and only look at "margin $\in \{1,3,10\}$", the percentage of patients who could get a strictly better organ by misreporting their EPTS is between 50% and 90%. To better understand incentives for manipulation, we also look into the percentage of patients running the risk of being allocated a lower-quality organ

by mis-reporting their EPTS. Figure 7 gives the data for Federal, MIN and omits the similar curves for the State matching for clarity. This graph shows that decreasing the reported EPTS is fairly risk-free while potential having a positive impact on the allocated KDPI. Conversely, increasing the reported EPTS may result in a match with a better organ, but not surprisingly it can also make things worse.

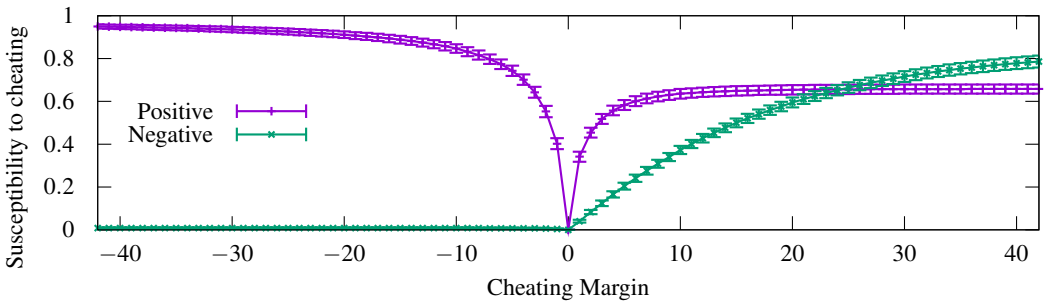


Figure 7: Proportion of patients whose allocated KDPI may change through their own manipulation as a function of the magnitude of EPTS misreporting, under Federal, MIN.

Although these experiments clearly show that the opportunity for strategic behaviour exists, the opportunities in practice will not be as great. For example, our experiments suppose complete information. In practice, doctors will not know the EPTS of all future patients yet to join the waiting list, or the KDPI of all future organs.

B.3 Results on Participation

We examined if patients can strategically delay entrance into the market to improve their outcome. To do this we took, for each of the 1000 iterations, the complete set of transplants made during that iteration and checked if a patient could receive a kidney for which they had higher utility by declaring a later arrival time. We took each patient p and checked if, in a future time step, a compatible kidney arrives of better quality for p. If so, we check if p (1) could have arrived before p' who received the better kidney and (2) p would have had priority greater than or equal to that of p' where priority is determined by $|KDPI - EPTS|$ for the MIN algorithm and by placement in the same equivalence class for the BOX algorithm. If so than we say p could have delayed for an improvement in utility. This is a worst case measure, we assume that p has full knowledge of all the kidneys that arrive in the future and p wins tie breakers against p'.

Our results for participation are shown in Table 4, standard deviations over 1000 runs in parenthesis. The MIN algorithm is much less susceptible to manipulation from arriving later. This is due to the large equivalence classes that are crated by the BOX mechanism, which offer more opportunities for optimistic manipulation. Additionally, the magnitude of the improvement in KDPI of the transplanted kidney is enormous for the BOX mechanism, indicating that arriving later can significantly increase patient utility. Interestingly, we see an increase in the number of patients that could manipulate in the MIN, State treatment. This is likely due to the smaller pools of patients and donors which leads to an increased disparity between EPTS and KDPI of patient and donor, illustrated in Figure 4, leading to an increase in the delay opportunity.

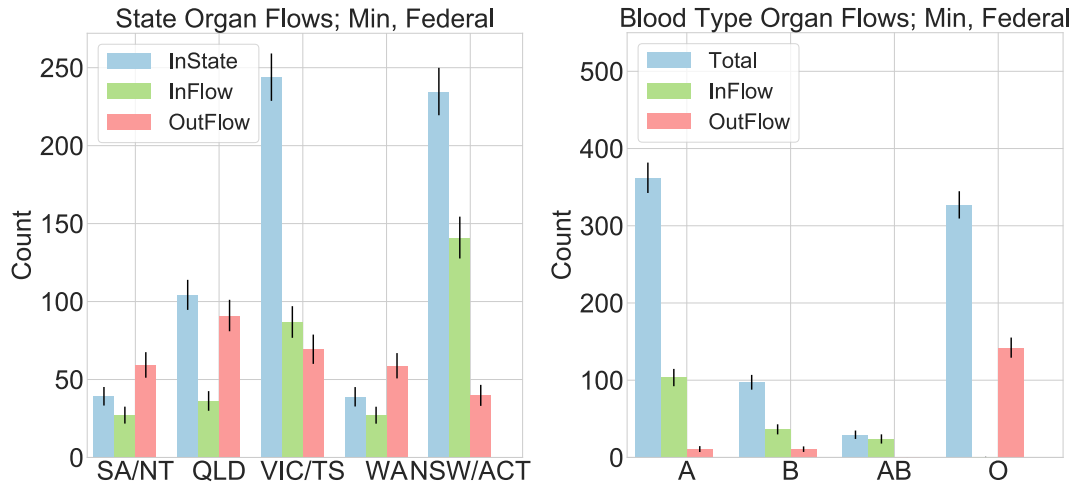
B.4 Waiting Time and Organ Flow Figures

In Figure 9b (moved to the appendix) we slice the mean waiting time data along the blood type axis. This reveals another interesting property of online organ matching: different blood types get

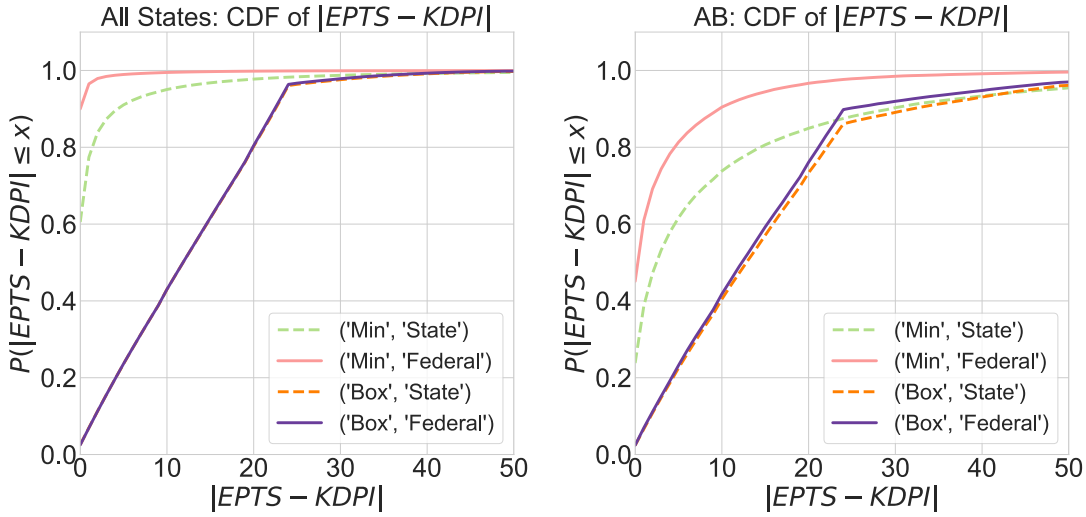
		Mean Num. Patients Delay Improves	Mean Num. of Improvements/Patient	Mean Max Improvement in KDPI
MIN	Federal	138.2 (20.5)	2.7 (0.34)	3.0 (1.25)
	State	269.3 (24.1)	2.5 (0.31)	9.2 (2.45)
BOX	Federal	960.5 (29.9)	169.5 (7.1)	50.9 (0.93)
	State	928.8 (31.6)	40.6 (2.1)	49.6 (1.00)

Table 4: The mean number of patients who can delay and improve their received organ, the mean number of improvements per patient, and the mean of the maximum improvement a patient could see for all 1000 iterations (standard deviation) for the four algorithmic treatments. The MIN, Federal algorithm often achieves allocations where KDPI=EPTS which affords few opportunities for strategic participation.

treated very differently. Those patients with type O blood, the most common amongst the types, track very closely to the waiting times see in Figure 9a while those with type AB blood have strictly lower waiting times. This is due to the small number of AB patients that require transplants and their ability to accept kidneys of any type.

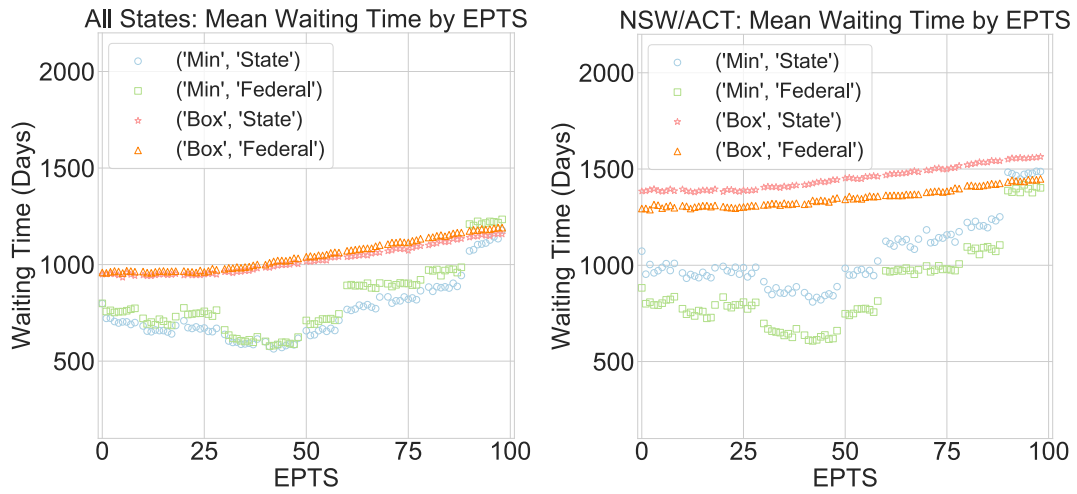


(a) The majority of organs are transplanted within the state of donation.

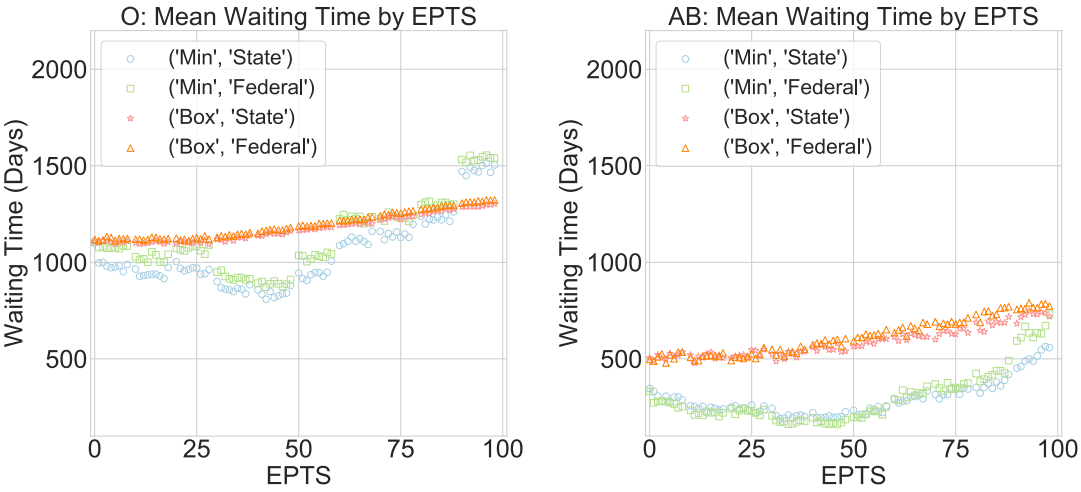


(b) With the MIN, Federal treatment, 99% of agents receive an organ where $|KDPI - EPTS| < 5$. Agents with a higher $|KDPI - EPTS|$ are almost all waiting for a rare AB organ.

Figure 8: In (a) we show the average over 1000 iterations of the flow of organs between states and blood types; error bars represent one standard deviation. In (b) we show the Cumulative distribution function (CDF) of $|KDPI - EPTS|$. The vertical axis shows the percentage of patients who have a $|KDPI - EPTS|$ below the value on the horizontal axis.



(a) Mean time on list for all states (left) and for NSW/ACT (right).



(b) Mean time on list for type O (left) and type AB (right).

Figure 9: In (a) we see that MIN gives lower average waiting times and that the difference between State and Federal allocations is negligible. In (b) Blood type O is the most common and thus the waiting times track closely to the overall mean while the rarest blood type AB does not.