Outcome-Based Regulation and Adverse Selection in Lung Transplantation

Andrew Schaefer

April 4, 2023

andrew.schaefer@rice.edu

RICE UNIVERSITY

Acknowledgments

David Mildebrath — Amazon

Taewoo Lee, Ph.D. — The University of Pittsburgh

Saumya Sinha, Ph.D. — The University of Minnesota – Twin Cities

**Ahmed Gaber**, M.D. — Houston Methodist Hospital

Outline

1. Government regulation in transplantation
2. Optimization model
3. Structural properties and analysis
4. Numerical experiments

Government Regulation in Transplantation
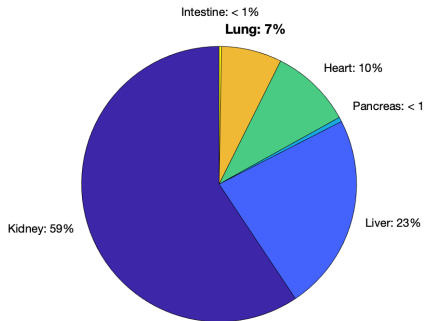
Organ Transplantation

- Organ transplantation is often the only treatment option for several end-stage diseases
- Growing transplant volume over the last few decades
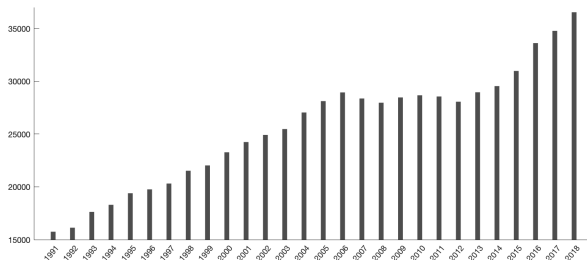
Number of solid-organ transplants in the US, 1991-2018

Lung Transplantation

- Lungs are the fourth-most transplanted organ in the US

- Constituted about 7% of all transplants in 2018

- Sparsely studied in the OR literature

- Afford some modeling advantages (more later)

Six largest single-organ transplants
in the US by volume in 2018

Transplant Regulations

Two sets of regulations in the past 20 years (CMS and OPTN)

|	CMS	OPTN
In Effect	2007–2019	2000–Present
Criteria	Frequentist	Bayesian (more stringent)
Lose Medicare/Medicaid?	Yes	Yes**

Medicare and Medcaid pay for a majority of transplants
at most programs in the United States

Because of the severity of CMS penalties (i.e. potential loss of
Medicare/Medicaid reimbursement), these regulations have been more widely
studied—we focus on them here

**Possible but less likely

Transplant Regulations

- We describe CMS regulations (OPTN are similar)
- In 2007, Centers for Medicaid & Medicare Services (CMS) noted large variability among outcomes across programs
- Introduced regulations to incentivize better post-transplant outcomes
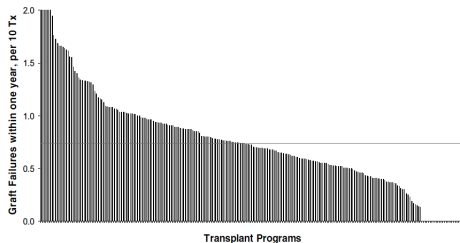
Not all transplant programs are created equal: Graft failure
per 10 transplants across different programs (Dickinson *et al.*, 2008)

CMS Conditions of Participation (CoPs)
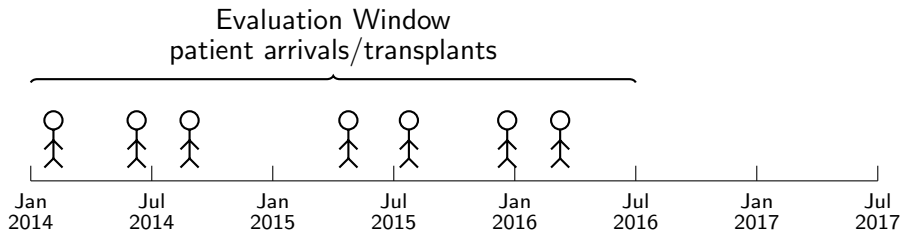
The CoPs evaluate transplant programs over 2.5-year evaluation windows
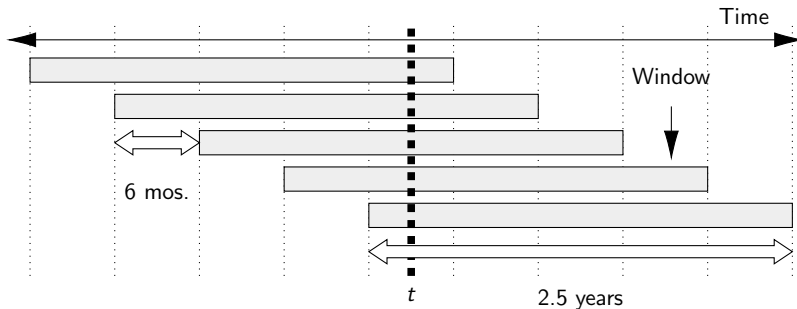
Five windows are 'active' at any time

In a single window, patients arrive and receive transplants

CMS Conditions of Participation (CoPs)

One year later, program reports to CMS

(i) all recipient-donor data, and

(ii) O = patient deaths/graft failures within one year of transplant
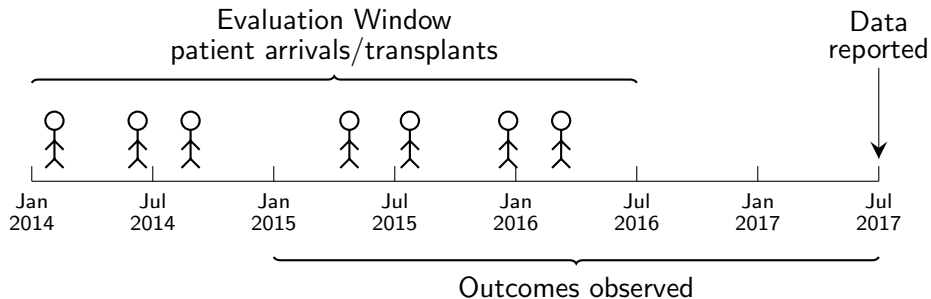
CMS Conditions of Participation (CoPs)

CMS computes survival function for each recipient; obtains expected number of deaths E
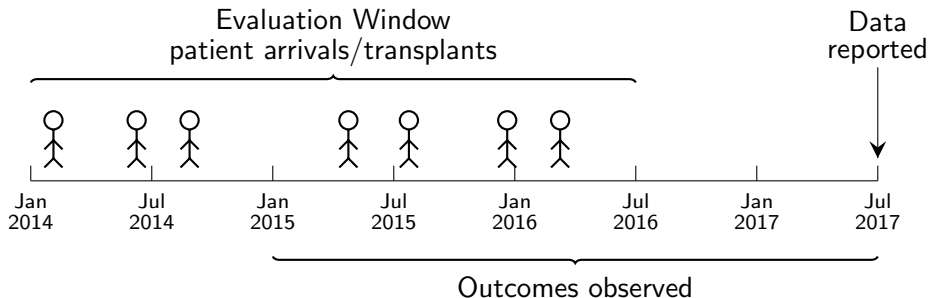
8/43

CMS Conditions of Participation (CoPs)

Program is flagged if O is
"much larger" than E

Program is flagged if O is "much larger" than E

In particular, if the following criteria are *violated*:
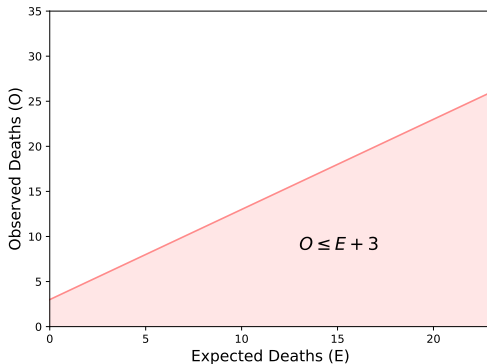
1. $O \leq E + 3$,

CMS Conditions of Participation (CoPs)

Program is flagged if O is "much larger" than E

In particular, if the following criteria are *violated*:

1. $O \leq E + 3$,
2. $O \leq 1.5E$, and

Program is flagged if O is "much larger" than E

In particular, if the following criteria are *violated*:

1. $O \leq E + 3$,
2. $O \leq 1.5E$, and
3. p-value $\leq$ threshold
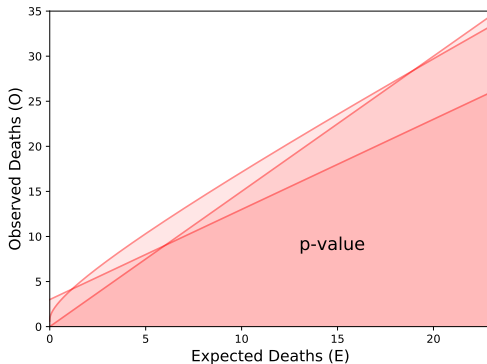
CMS Conditions of Participation (CoPs)

Program is flagged if O is "much larger" than E

In particular, if the following criteria are *violated*:
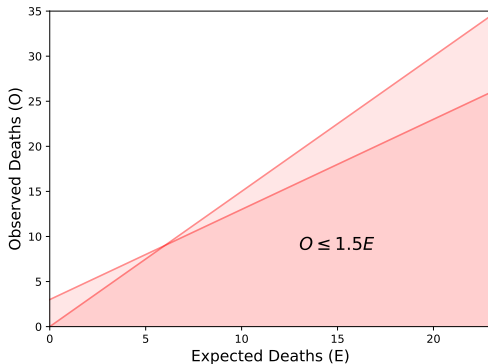
1. $O \leq E + 3$,
2. $O \leq 1.5E$, and
3. p-value $\leq$ threshold

OPTN uses a slightly different function to compare O/E
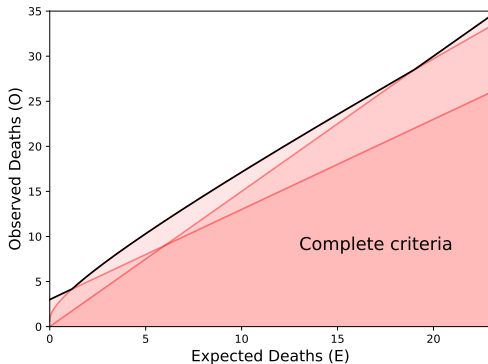
Penalties for Flagging

- Penalties associated with flagging can be severe
- Program must establish a remediation plan with CMS
- Negative publicity
- Temporary or permanent shutdown
- Estimated costs can run into tens of millions of dollars (USD)
- OPTN penalization perceived as less severe, but concerns remain [1] [2]

---

[1] Andreoni, American Journal of Transplantation, 2020, 20(8); 2026-2029

[2] Schold, Current Opinion in Organ Transplantation, 2020, 25(2); 158-162

- Widespread criticism from the medical community
- Risk-adjustment does not account for pre-transplant outcomes, co-morbidities, other mitigating factors (Weinhandl *et al.*, 2009)
- CoPs unable to identify truly underperforming programs; prevalence of false flagging (Massie and Segev, 2013)
- Programs forced to become risk-averse to both recipients and donors (Jay and Schold, 2017)

HEART FAILURE

Federal Inspectors Cite St. Luke's in Houston for Problems in a Heart Transplant

ProPublica, October 2018

STATNews, August 2016

DON'T MISS

Hospitals are throwing out organs and denying transplants to meet federal standards

By CASEY ROSS @caseymross / AUGUST 11, 2016

Deaths at OHSU heart transplant program spiked before program shut down

Updated Jul 12, 2019; Posted Jul 12, 2019

The Oregonian, July 2019

Empirical evidence suggests that when programs are flagged,
their volume decreases

Cause of this decrease is debated

- Programs reject risky patients to reduce chances of future penalization?

- Patients choosing to seek care at better programs?

- Temporary adjustment while programs "re-group"?

Do CMS and OPTN regulations create incentives for programs to reject patients?

Model a program which seeks to:

1. Maximize transplant volume

2. Control risk of penalization by OPTN/CMS

OR Literature
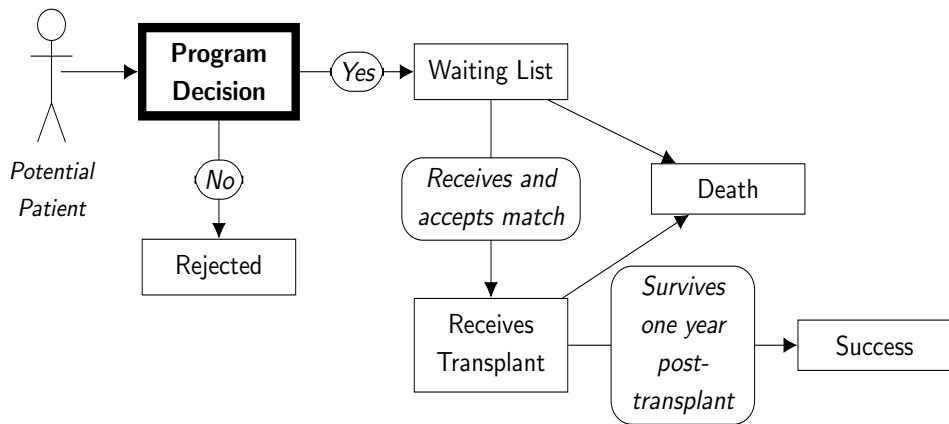
National/governmental perspective:

- Zenios *et al.*(2000)–kidney allocation
- Kong *et al.*(2010)–allocation region design
- Akan *et al.*(2012)–liver allocation
- . . .

Individual/patient perspective:

- Alagoz *et al.*(2004, 2007a, 2007b)–liver acceptance
- Sandıkçı *et al.*(2008, 2013)–patient perspective of waitlist
- . . .

Program/medical perspective:

This work (first from a program's perspective)

Clinical Literature

Editorial/opinion articles:

- Abecassis *et al.*(2009) - CoPs threat to innovation
- Schold and Axelrod (2014) - Bayesian approach to CoPs
- Hamilton (2013) - Impact of CoPs on patients

Simulation/data-driven:

- Massie and Segev (2013) - prevalence of false flagging
- Schold *et al.*(2013) - flagging and decline in volume
- Dolgin *et al.*(2016) - removal from liver transplant waitlist

Optimization: This work

Optimization Model

Idea Behind the Model

Each week, a batch of patients arrives

Each patient modeled by two numbers, c and e, where

$$1 - c = \mathbb{P}\left[\begin{array}{c|c} \text{survives} & \text{receives} \\ \text{one year} & \text{transplant} \end{array}\right] \rightarrow \text{according to program}$$

$$1 - e = \mathbb{P}\left[\begin{array}{c|c} \text{survives} & \text{receives} \\ \text{one year} & \text{transplant} \end{array}\right] \rightarrow \text{according to CMS}$$

Program model is a better predictor of patient survival

Assumptions

1. [Mild] Program probability $1 - c$ is the true survival probability
   Programs build higher fidelity prediction models with more robust data than CMS (Chan *et al.*, 2019)
2. [Medium] Survival probabilities independent of donor information
   Most factors associated with lung transplant failure depend on recipient, not donor (Diamond *et al.*, 2013)

3. [Medium] Patients receive transplants immediately after acceptance
   Median wait times for lung transplants less than six months – approx. 2
   years for kidneys (OPTN database).
   Short wait-time compared to length of evaluation window

4. [Mild] Patient arrivals are Poisson
   Common in the literature (e.g. Zenios (2000) and Shechter et al. (2005))
   and is also consistent with Houston Methodist data

Decision variables

Current patients: binary variables z_j for whom to accept from current pool of candidates

Future patients: $u_{it} \in [0, 1]$, fraction of type i patients accepted in week t

But what is a "type i" patient?

Classifying future patients

To control flagging risk, need to predict future patients

Define patient classes (indexed by i)

Class i has 3 associated values:

$$1 - c_i = \text{Survival probability (program)}$$
$$1 - e_i = \text{Survival probability (CMS)}$$
$$\lambda_i = \text{Mean \# patients to arrive each week}$$

Note that patients with different physiology could be grouped together because the regulations only care about patient-risk

Our Model

Z_{it} = number of type i patients to arrive in week t,
$Z_{it} \sim \text{Poisson}(\lambda_i)$

$A_{it}^j = 1$ iff patient j of type i accepted in week t,
$A_{it}^j \sim \text{Bernoulli}(u_{it})$ for all $j = 1, \ldots, Z_{it}$

Decision variable:
u_{it} = fraction of type i patients accepted in week t

Our Model

Y_{it} = number of type i patients accepted in week t,

$$Y_{it} = \sum_{j=1}^{Z_{it}} A_{it}^j \sim \text{Poisson}(\lambda_i u_{it})$$

$X_{it}^j = 1$ iff patient j dies within one year of transplant,
$X_{it}^j \sim \text{Bernoulli}(c_i)$ for all $j = 1, \ldots, Z_{it}$

Compute observed + expected deaths for window w:

$$O_w = \sum_{i,t} \sum_{j=1}^{Y_{it}} X_{it}^j \quad \text{and} \quad E_w = \sum_{i,t} e_i Y_{it}$$

Can easily compute mean and variance of O_w and E_w

Our Model

Flagging criteria: A program is flagged in window w if

1. **Absolute:** $O_w \leqslant E_w + 3$,
2. **Relative:** $O_w \leqslant 1.5 E_w$, and
3. **p-value:** p-value $\leqslant 0.05$

To limit risk of getting flagged, use chance constraints

$$\mathbb{P}[\min\{O_w - 1.5 E_w, O_w - E_w - 3, p\text{-value} - 0.05\} \leqslant 0\,] \geqslant 1 - \alpha_w,$$

where α_w is a pre-defined risk tolerance

Our Model

LHS of chance constraints famously hard to compute

Assume that O_w and E_w are normally distributed; valid for large transplant programs (e.g. Houston Methodist Hospital)

Then, each constraint[3] is normally distributed with parameters $(\mu_w^\ell, \sigma_w^\ell)$, $\ell = 1, 2, 3$

Conservative approximation to chance constraint given by

$$\min_{\ell=1,2,3} \left\{ \mu_w^\ell + \varphi_w \sigma_w^\ell \right\} \leqslant 0$$

where $\varphi_w = \Phi^{-1}(1 - \alpha_w)$, $\Phi = $ CDF of standard normal distribution

---

[3]after linearizing the p-value constraint

Our Model

To summarize,

- Model is solved once each week
- Decision variables: binary z_j and continuous u_{it}
- Objective function: total transplant volume
- Constraints: chance constraints (and others)

Main Inputs:

1. Data for patients under consideration
2. Patient class data for future patients
3. Data for the five "active" windows
   (Mean/variance of expected/observed deaths)
4. Risk tolerances for every window

Our Model

$$\max \quad \sum_j z_j + \sum_{i,t} \lambda_i u_{it} \qquad\qquad \mathbb{E}[\#\text{transplants}]$$

$$\text{s.t.} \quad \min_{\ell=1,2,3} \left\{ \mu_w^\ell + \varphi_w \sigma_w^\ell \right\} \leqslant 0 \ \forall \ w \qquad \text{Chance constraints}$$

$$\mu_w^\ell = (a_w^\ell)^T(u, \ z) \qquad\qquad \text{Mean (linear)}$$

$$(\sigma_w^\ell)^2 = (b_w^\ell)^T(u, \ z) \qquad\qquad \text{Variance (linear)}$$

$$L_w \leq \sum z_j + \sum \lambda_i u_{it} \leq U_w \ \forall \ w \quad \text{Operational constraints}$$

$$u_{it} \in [0,1] \ \forall \ (i,t) \qquad\qquad \text{Future patients}$$

$$z_j \in \{0,1\} \ \forall \ j \qquad\qquad \text{Current patients}$$

$$\max \quad \sum_j z_j + \sum_{i,t} \lambda_i u_{it} \qquad \mathbb{E}[\#\text{transplants}]$$

$$\text{s.t.} \quad \min_{\ell=1,2,3} \left\{ \mu_w^\ell + \varphi_w \sigma_w^\ell \right\} \leqslant 0 \ \forall \ w \qquad \text{Chance constraints}$$

$$\mu_w^\ell = (a_w^\ell)^T(u, \ z) \qquad \text{Mean (linear)}$$

$$(\sigma_w^\ell)^2 = (b_w^\ell)^T(u, \ z) \qquad \text{Variance (linear)}$$

$$L_w \leq \sum z_j + \sum \lambda_i u_{it} \leq U_w \ \forall \ w \qquad \text{Operational constraints}$$

$$u_{it} \in [0,1] \ \forall \ (i, t) \qquad \text{Future patients}$$

$$z_j \in \{0,1\} \ \forall \ j \qquad \text{Current patients}$$

Hard to solve!

Structural Properties and Analysis

Numerical solution exhibits steady-state behavior that captures long-term strategy – conveys program's inherent risk of getting flagged, ignores transient effects

Steady-State Behavior
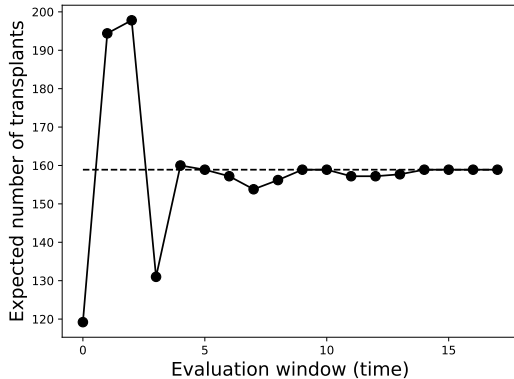
Can we characterize the steady-state?

Easier to analyze a single-window model
Justified because. . .

Theorem

For sufficiently long horizons T, the optimal solution converges to that given by solving a **single-window** model (under mild assumptions)

Proof: Compute upper and lower bounds on multi-window objective, show they converge to each other (squeeze theorem)

Single-Window Model

$$\max \quad \lambda^T u$$
$$\text{s.t.} \quad \min_{\ell=1,2,3} \left\{ a_\ell^T u + \varphi \sqrt{b_\ell^T u} \right\} \leqslant 0 \qquad (M_{sw})$$
$$u_{it} \in [0,1] \ \forall \ (i,t)$$

Can ignore t dependence, because ...

Proposition

(M_{sw}) has a stationary optimal solution (i.e., $u_{it_1} = u_{it_2} \ \forall \ t_1, t_2$)

Proof: Analysis of characteristic functions of O_w and E_w

Single-Window Model

$$
\max \quad \lambda^T u
$$
$$
\text{s.t.} \quad \min_{\ell=1,2,3} \left\{ a_\ell^T u + \varphi \sqrt{b_\ell^T u} \right\} \leqslant 0 \qquad (M_{sw})
$$
$$
u_i \in [0,1] \ \forall \ i
$$

Can ignore t dependence, because . . .

Proposition

(M_{sw}) has a stationary optimal solution (i.e., $u_{it_1} = u_{it_2} \ \forall \ t_1, t_2$)

Proof: Analysis of characteristic functions of O_w and E_w

Single-Window Model

(M_{sw}) is a **reverse convex program**, with the following special structure

---

Theorem (Hillestad and Jacobsen–1980)

Consider the problem

$$\max\{c^T x \mid Ax \leqslant b, \ g(x) \leqslant 0\} \qquad (P)$$

for some continuous, strictly concave function g. If (P) has an optimal solution, then it has an optimal solution that lies on an edge of the polyhedron $\{Ax \leqslant b\}$

Single-Window Model

$$\max \quad \lambda^T u$$

$$\text{s.t.} \quad \min_{\ell=1,2,3} \left\{ a_\ell^T u + \varphi \sqrt{b_\ell^T u} \right\} \leqslant 0 \quad \text{strictly concave if } \varphi > 0 \text{ (i.e., } \alpha < 1/2)$$

$$u_i \in [0,1] \; \forall \; i. \qquad\qquad \text{unit cube (polytope)}$$

By Hillestad and Jacobsen: there exists an optimal solution on the edge of the unit cube

Upshot: At optimality, at most one u_i is fractional. (All other $u_i \in \{0,1\}$)

Could it be optimal for a program to stop transplants?

Consider a large program (transplant volume ≥ 30)

For large programs, we can ignore the $E + 3$ and p-value constraints (Dickinson, 2006)

Convexify the non-convex constraint $a^T u + \varphi \sqrt{b^T u} \leqslant 0$

Theorem

Let $\mathcal{H}^- = \{i \mid c_i < 1.5e_i\}$. Then $u^* \equiv 0$ is optimal for the convex relaxation iff

$$\sum_{i \in \mathcal{H}^-} \lambda_i \frac{(c_i - 1.5e_i)^2}{(c_i - 1.5e_i)^2 + c_i(1 - c_i)} < \frac{\varphi^2}{130^2}$$

Proof: KKT conditions

How to interpret this condition?

Keeping Programs Open

Program closes if $\sum_{i \in \mathcal{H}^-} \dfrac{\lambda_i}{1 + \nu_i^2} < \dfrac{\varphi^2}{130^2}$

$\nu_i =$ Coeff. of variation of $X_i - 1.5e_i$; $\mathcal{H}^- = \{i \mid c_i < 1.5e_i\}$

Want the LHS to be as large as possible, that is...

$$
\begin{aligned}
|\mathcal{H}^-| \text{ large} &\Rightarrow \text{ Many classes with } c_i < 1.5e_i, \\
\lambda_i \text{ large} &\Rightarrow \text{ Many patients of type } i \in \mathcal{H}^-, \\
\nu_i \text{ small} &\Rightarrow c_i \text{ close to 0 or 1, certainty about outcomes}
\end{aligned}
$$

Numerical Experiments and Insights

Creating Patient Classes

Patient classes created from $n = 469$ patients added to the waitlist
at Houston Methodist hospital (HMH) between Jan 2014 and Dec 2018
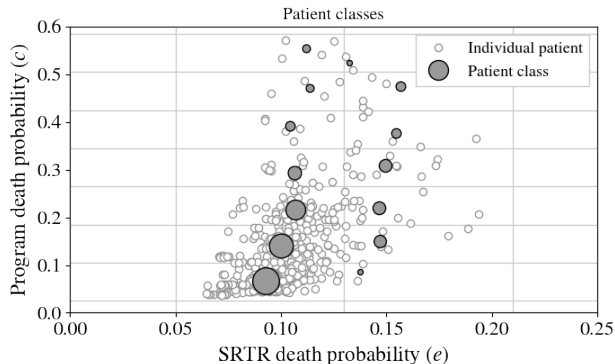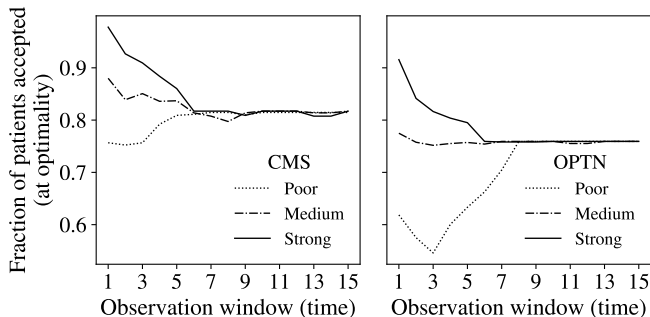

Generated 22 patient classes.
Final results are robust to clustering method

If a program is currently in a favorable position
how many more patients can receive transplant?

Initial positions from HMH data for 2016

Effect of initial position on convergence

Unfair Penalization of Medium-Sized Programs

How does a program's response depend on incoming patient volume?

Consider a program, with fixed risk tolerance (3%)
Vary the patient arrival rate, keeping patient 'mix' the same

Medium-sized programs accept fewer patients than larger programs

Effect of program size on transplant volume
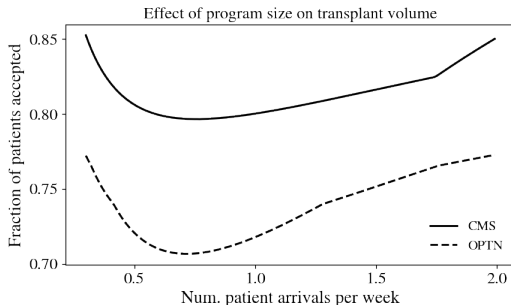
Response After Flagging

The previous result has further implications

Common hypothesis: if a program is flagged, its transplant volume declines as patients choose to seek care at better programs[4]

Then, flagged programs get hit *twice*:

- Patient volume declines due to patient choice, *and*
- Program forced to accept patients at lower rate to limit flagging risk

The second effect has not been studied before

---

[4]See, e.g., Howard and Kaplan, *Do report cards influence hospital choice? The case of kidney transplantation.* Inquiry (2006) 43:150–159

What if the Program Meets Expectations?

What if a program exactly meets regulatory expectation? Can they transplant 100% of their patients? No

Risk profile for programs meeting expectations
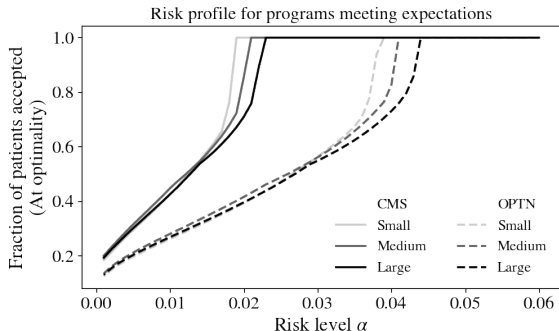
All programs satisfy $c_i = e_i$ for all patient classes i

As risk tolerance α decreases, the fraction of patients accepted at optimality drops below 100%

Key Insights

- CMS and OPTN regualtions *do* create incentives for programs to reject patients
- This incentive does not disappear even with adequate risk adjustment
- Medium-sized programs may be unfairly penalized under these regulations

Conclusion

- Developed first optimization model from a transplant program's perspective
- Presented first rigorous analysis of misaligned incentives under CMS/OPTN regulation
- Demonstrated previously unobserved problems with outcome-based regulation

Thank you!

CMS and OPTN Criteria: Frequentist vs. Bayesian

CMS used three (frequentist) criteria

Program flagged if all three hold

1. **Actionability**: $O/E > 1.5$ (or, later, 1.85)
2. **Importance**: $O - E > 3$
3. **Significance**: One-side p-value < 0.05

CMS and OPTN Criteria: Frequentist vs. Bayesian

OPTN uses Bayesian criteria (beginning in 2014)

Put a $\mathrm{Gamma}(2, 2)$ prior on hazard ratio HR.
Assume $O \sim \mathrm{Poisson}(\mathrm{HR} \times E)$.
Then posterior for HR is a Gamma distribution
with mean $(O + 2)/(E + 2)$ and variance $(O + 2)/(E + 2)^2$.

Program flagged if *either*

$\mathbb{P}[\mathrm{HR} > 1.2] \geq 75\%$ OR $\mathbb{P}[\mathrm{HR} > 2.5] \geq 10\%$

Parameters 1.2, 2.5, etc. chosen via simulation

Convexification of Reverse Convex Program

Recall constraint

$$\min_{\ell=1,2,3} \left\{ a_\ell^T u + \varphi \sqrt{b_\ell^T u} \right\} \leqslant 0$$

Can ignore two of the constraints ($E + 3$ and p-value) for large programs

In the non-convex constraint, replace each u_i with u_i^2.
That is, replace

$$a^T u + \varphi \sqrt{b^T u} \leqslant 0 \quad \text{with} \quad a^T u + \varphi \sqrt{u^T B u} \leqslant 0,$$

where $B = \text{diag}(b)$. New constraint is convex if $\varphi > 0$.
This gives a convex relaxation.

Convexification of Reverse Convex Program

Rewrite the convex constraint

$$a^T u + \varphi \sqrt{u^T B u} \leqslant 0 \quad \text{as} \quad a^T u + \varphi \|B^{1/2} u\|_2 \leqslant 0.$$

Possible because $B = \text{diag}(b)$, each $b_i = 130\lambda_i[(c_i - 1.5e_i)^2 + c_i(1 - c_i)] > 0$.

Therefore, we have a convex (second-order conic programming) relaxation for large programs:

$$
\begin{aligned}
\max \quad & \lambda^T u \\
\text{s.t.} \quad & a^T u + \varphi \|B^{1/2} u\|_2 \leqslant 0 \\
& u_i \in [0, 1] \ \forall \ i
\end{aligned}
$$