Optimization of Health and Health Delivery: A Technology Overview

Sanjay Mehrotra
Industrial Engineering and Management Sciences
mehrotra@iems.northwestern.edu
847 491 3155

Co-Authors on various papers: Ashley Davis, Kibaek Kim, Mark Daskin, Jane Holl, Daniela Ladner, John Friedewald, Anton Skaro, Michael Abecassis, Reuben Thomas, Olga Lyandres, Richard P. Van Duyne, Joseph T. Walsh, Matthew R. Glucksberg, Carlos J. Paredes, Eleftherios T. Papoutsakis, and Vassily Hatzimanikatis

Methodology work funded by: National Science Foundation under grants NSF CMMI-1100868, NSF-CMMI-0928936, NSF-CMMI-0727532, Office of Naval Research grants ONR-N00014-09-10518, Department of Energy grant DOE-SP0011568, NIH R01 GM065475, NIH-1R36HS021078-01

Images used in the presentation are taken from http://www.google.com/imghp

Outline

- ❖ Understanding Optimization
- ❖ Univariate Decisions (non-dichotomous) with Data Uncertainty
  - Example: Nurse Shift Planning
- ❖ Robust Predictive Modeling – Multivariate Decisions; Few Data Samples
  - Example: Glucose/Pyridine Concentration Predictive Modeling
- ❖ Designing Networks
  - Example: Addressing Geographic Inequity in Kidney Allocation
- ❖ Reverse Engineering an Underlying Mechanism – Longitudinal Data
  - Example: Gene Regulatory Networks
- ❖ Multi-Objective Decision Making
  - Example: National Diabetes Budget Allocation for Prevention Programs

Optimization Problem Structure

Objective

Constraint

Available Resource
Structural Requirements
Ambiguity in Data

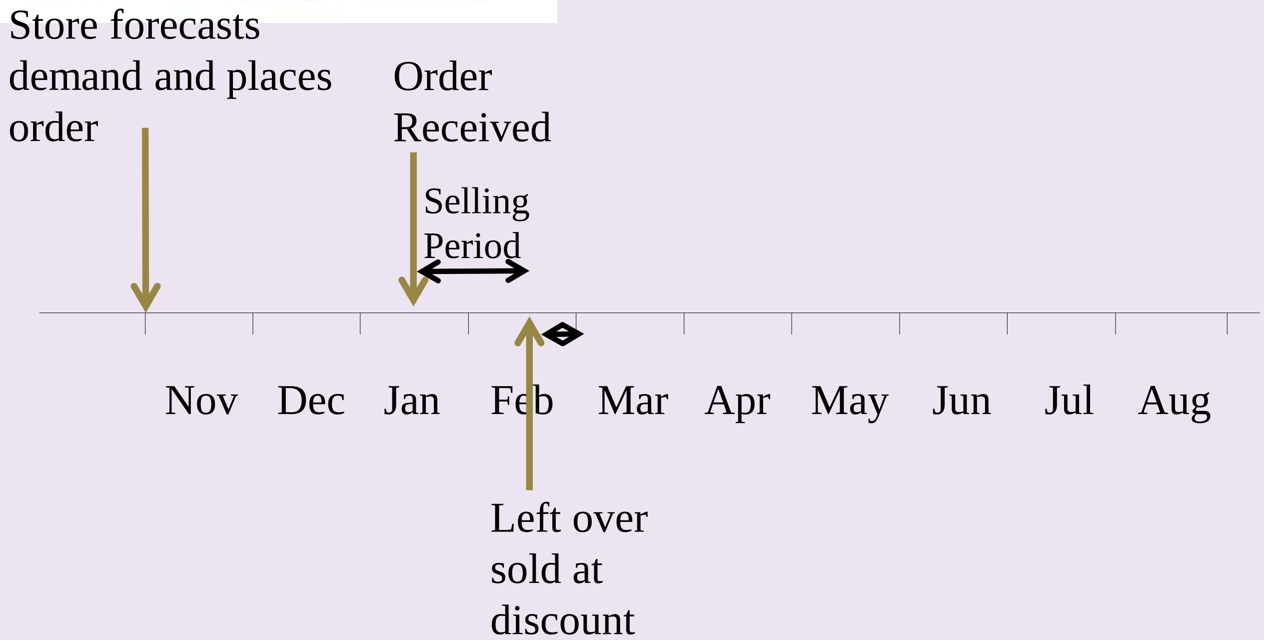
Decisions

How Much? -- Quantity
When? – Timing/Policy
Which? – Selection
Where? -- Locations
How? – Mechanism Design
Who? -- Scheduling

How Many Valentine Gift Boxes to Order?

Financial Input:

- Unit Sales Price p = $18
- Unit Purchasing Price c = $7/box
- Discounted Sales Price (Salvage Value) = $5

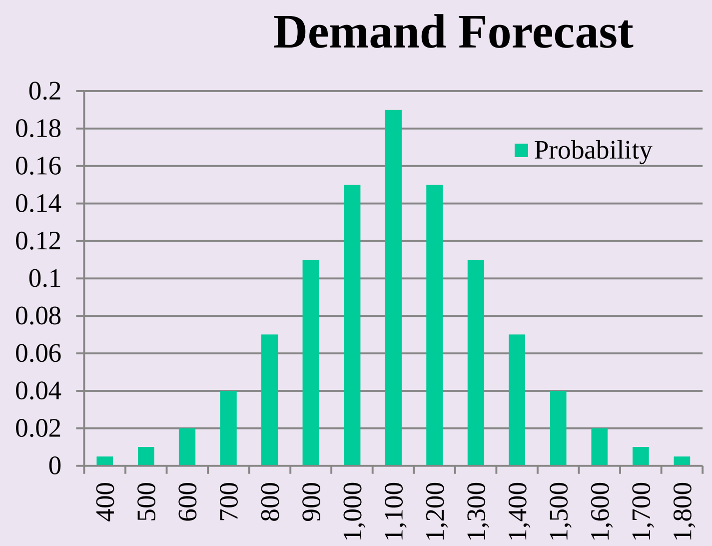


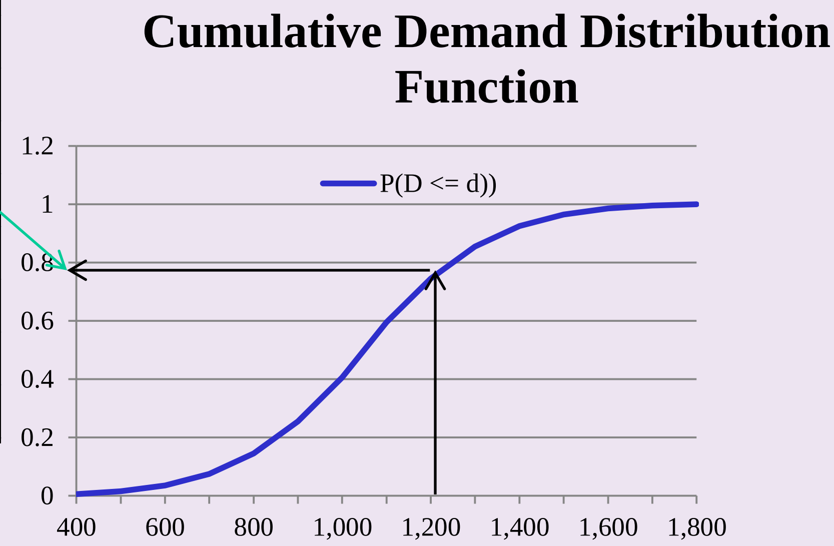
Elements of the Management Decision Problem

- Demand uncertainty
- An integer decision
- A reward function
- Decision before demand is realized
- Need a systematic method for finding the “best decision”

Demand Distribution

Demand d_i	Probability p_i	Cumulative Probability of demand being this size or less Prob(D $\leq$ d)
400	.005	0.005
500	.01	0.015
600	.02	0.035
700	.04	0.075
800	.07	0.145
900	.11	0.255
1,000	.15	0.405
1,100	.19	0.595
1,200	.15	0.745
1,300	.11	0.855
1,400	.07	0.925
1,500	.04	0.965
1,600	.02	0.985
1,700	.01	0.995
1,800	.005	1
Expected Demand = 1,100		





Profit as a Function of Demand

Profit = Sales Revenue + Salvage Value – Purchase Price

- Scenario 1:

Assume Order Quantity $q = 700$, Demand $d=500$

$$Profit = 18 * 500 + 5 * (700-500) - 7 * 700$$
$$= \$5,100$$

- Scenario 2:

$q = 700, d=1000$

$$Profit = 18 * 700 + 5 * (0) - 7 * 700$$
$$= \$7,700$$

For a given order quantity (q) depending on the demand we have different values profits!

Writing Profits Using Symbols

Profit = Sales Revenue + Salvage Value – Purchase Price

We can only sell minimum of demand and order quantity

q-*d* is the salvage quantity

$$\text{Profit}(q,d) = \underbrace{p \min(d,q)}_{\text{Sales Revenue}} + \underbrace{s \max(0,q-d)}_{\text{Salvage Revenue}} - \underbrace{c\, q}_{\text{Purchasing Cost}}$$

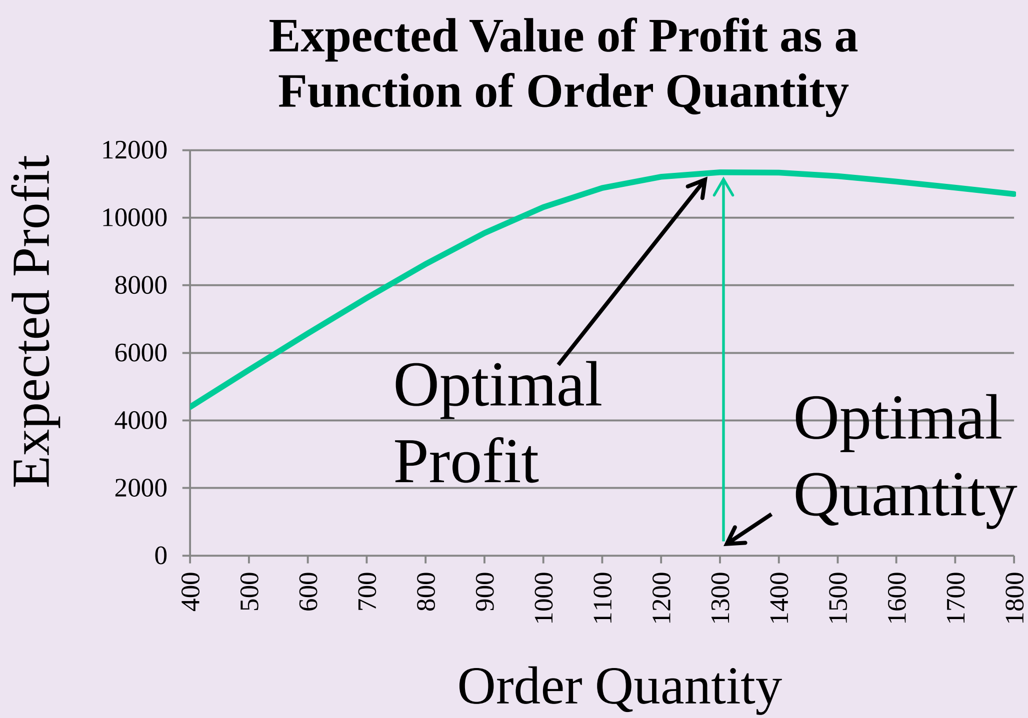
Expected Profit for a Given Order Quantity

Order Quantity *q = 700*			*q = 700*
Demand d_i	**Probability p_i**	**Profit**	**Probability * Profit**
400	.005	3800	19
500	.01	5100	51
600	.02	6400	128
700	.04	7700	308
800	.07	7700	539
900	.11	7700	847
1,000	.15	7700	1155
1,100	.19	7700	1463
1,200	.15	7700	1155
1,300	.11	7700	847
1,400	.07	7700	539
1,500	.04	7700	308
1,600	.02	7700	154
1,700	.01	7700	77
1,800	.005	7700	38.5
Expected Value of Profit			7628.5

- For each order quantity profit is a random variable.
- We calculate expected value of profit function, which is random for a given order quantity.
- We want know the “best” order quantity, i.e. one that maximizes expected profit!

Finding the Optimal Quantity

Order Quantity	Expected Profit
400	$ 4,400
500	$ 5,494
600	$ 6,574
700	$ 7,628
800	$ 8,631
900	$ 9,542
1,000	$ 10,311
1,100	$ 10,884
1,200	$ 11,211
1,300	**$ 11,342**
1,400	$ 11,331
1,500	$ 11,228
1,600	$ 11,074
1,700	$ 10,894
1,800	$ 10,700



- At optimum: Expected cost of lost sales due to under stocking = Expected cost of overstocking

Shift Staffing Levels

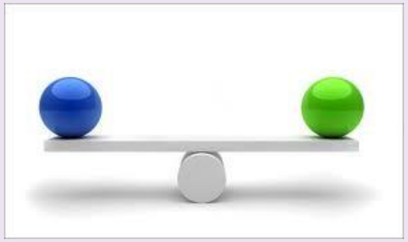
Situation: Need to staff a shift cost-effectively while not compromising patient safety

Common Practice: Staff with a mix of permanent and temporary (agency or float) nurses

Question: How many permanent nurses to use?

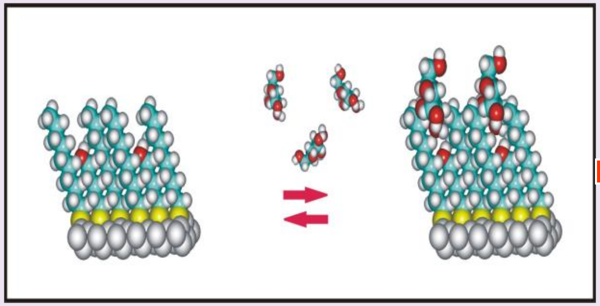
An Approach:
"Cost" := the per shift salary of a permanent RN
"Stock-Out Price" := daily per shift salary of a temporary RN
"Salvage Value" := benefit of having an extra permanent RN

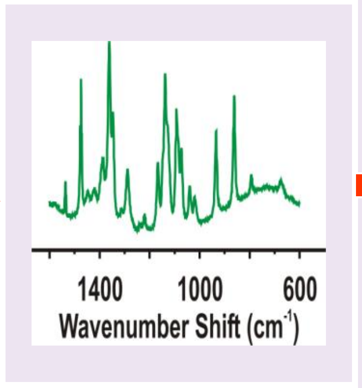
Total Cost = Regular staffing cost + under staffing costs + overstaffing costs



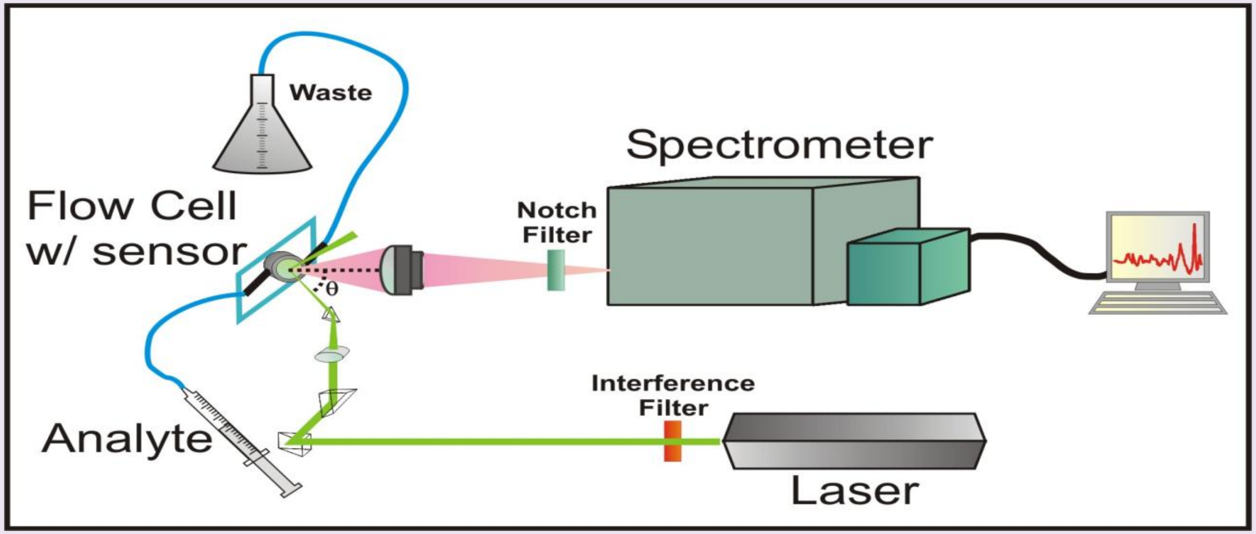
"Distribution-Robust Newsvendor Models for Shift Nurse Demand Estimation", Ashley Davis, Sanjay Mehrotra, Mark Daskin and Jane Holl (under review Asia Pacific Journal of Operations Research)

Predictive Modeling with Limited Noisy Experimental Data





Prediction



"Prediction range estimation from noisy Raman spectra with robust optimization", Olga Lyandres, Richard P. Van Duyne, Joseph T. Walsh, Matthew R. Glucksberg and Sanjay Mehrotra, *Analyst*, 135, 2111-2118, 2010

Predictive Modeling with Experimental Data

Question: How to predict true concentrations using limited calibration data?

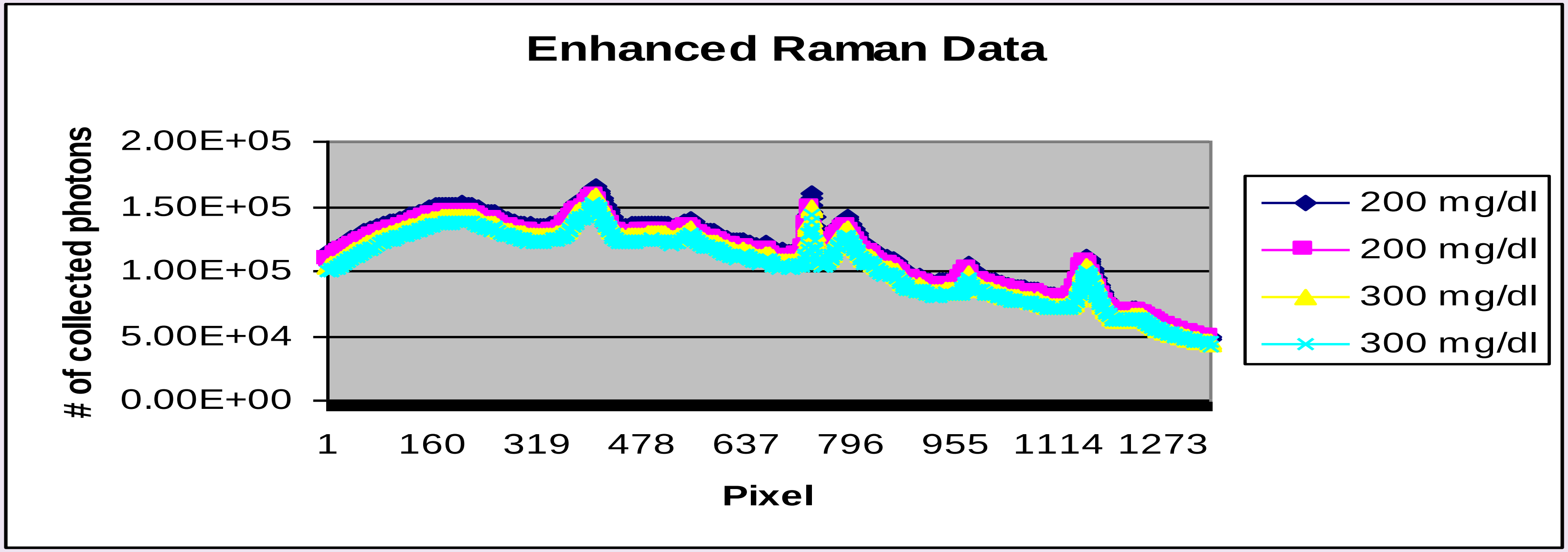
Problem: Calibration data is noisy and system is over determined.

Current Practice: Using Partial Least-squares from Statistics which is meant to filter noise and give prediction.

Alternative Approach: Build Robust Least-squares based optimization model

“Prediction range estimation from noisy Raman spectra with robust optimization”, Olga Lyandres, Richard P. Van Duyne, Joseph T. Walsh, Matthew R. Glucksberg and Sanjay Mehrotra, *Analyst*, 135, 2111-2118, 2010

Predictive Modeling with Experimental Data



Calibration

$[c]_g$=10, 20, 40, 60, 100, 150, 250, 350, 450 mg/dL

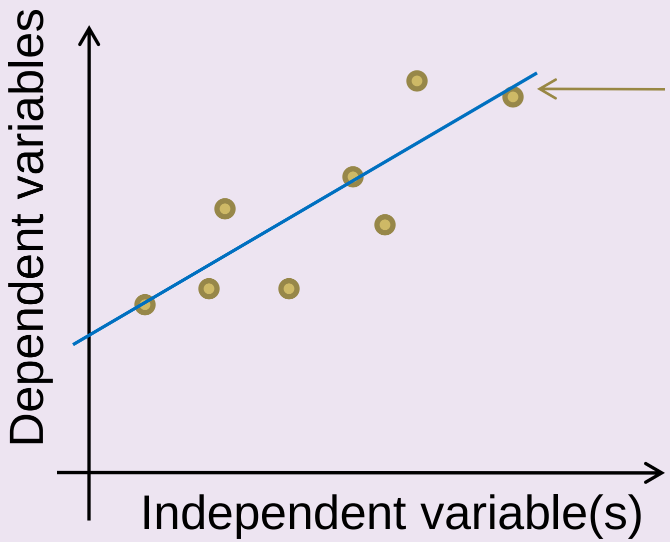
Validation

$[c]_g$=15, 50, 80,120, 200, 300 mg/dL

10 spectra at each concentration, baseline corrected, normalized

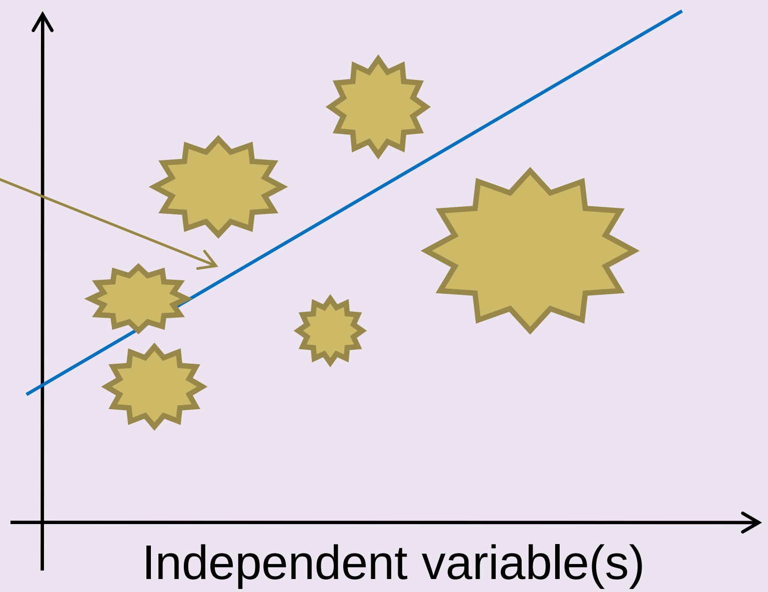
"Prediction range estimation from noisy Raman spectra with robust optimization", Olga Lyandres, Richard P. Van Duyne, Joseph T. Walsh, Matthew R. Glucksberg and Sanjay Mehrotra, *Analyst*, 135, 2111-2118, 2010

Robust Least-square Framework



Need to find a Linear Predictive Estimator

$$y^u := \sum_{j=1}^{n} \beta_j y_j$$



Ordinary Least-Square

$$\min_{\beta_1,\dots,\beta_n} \left\| \mathbf{x}^u - \sum_{j=1}^{n} \beta_j \mathbf{x}_j^0 \right\|^2$$

Robust Least-Square/Optimization

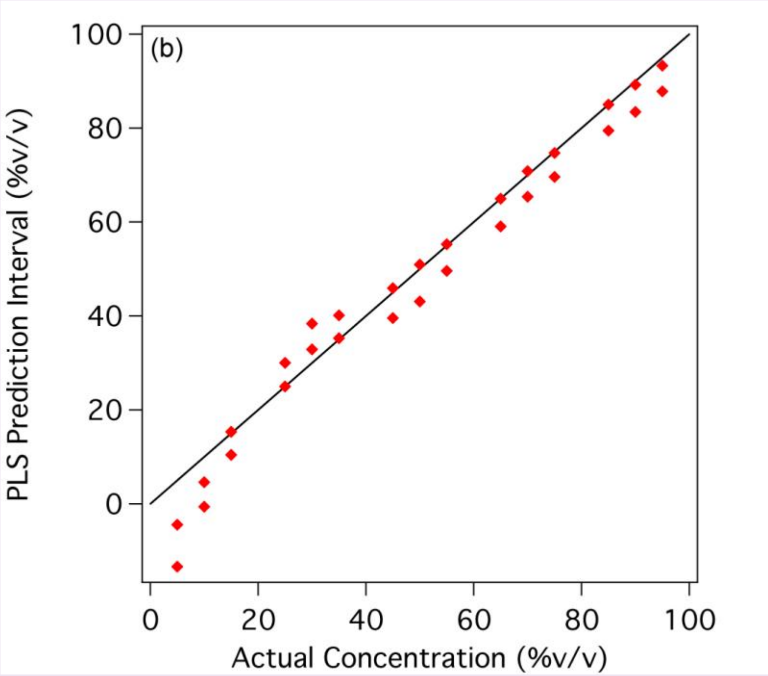
$$y_{lo}^* = \min \sum_{j=1}^{n} \beta_j y_j$$

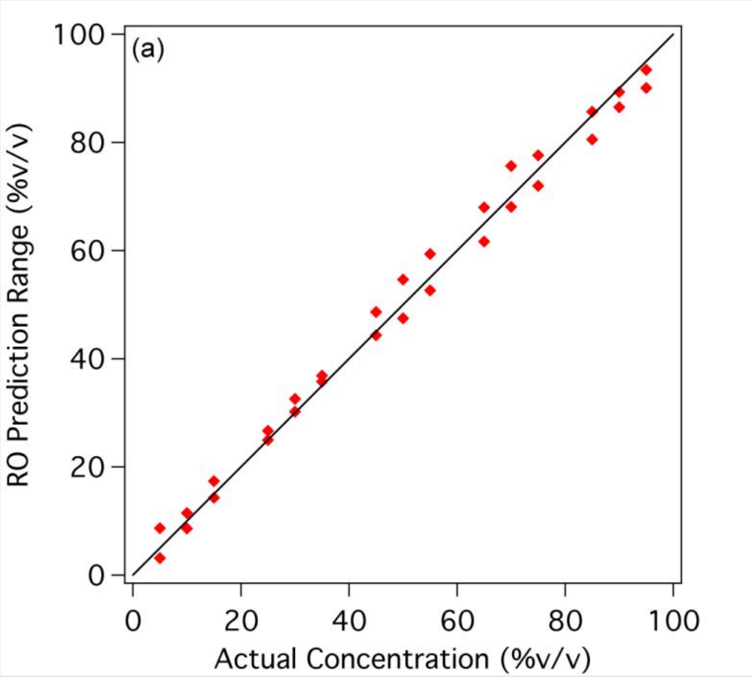
$$\text{subject to } \left\| \mathbf{x}^u - \sum_{j=1}^{n} \beta_j \mathbf{x}_j \right\|^2 \le z \qquad \mathbf{x}_j \in U_j, \; j = 1, \dots, n$$

$$\sum_{j=1}^{n} \beta_j = 1, \;\; \beta_j \ge 0$$

Uncertainty Set

Results: A pyridine System





True value is out of prediction interval in 12/15 times.

True value is out of prediction interval in 4/15 times.

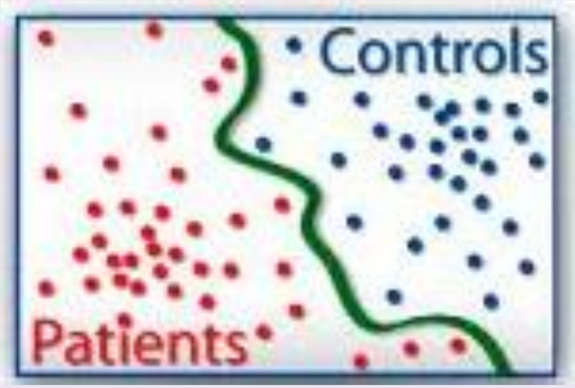
"Prediction range estimation from noisy Raman spectra with robust optimization", Olga Lyandres, Richard P. Van Duyne, Joseph T. Walsh, Matthew R. Glucksberg and Sanjay Mehrotra, *Analyst*, 135, 2111-2118, 2010

Results: Comparison with all Popular PLS Methods

Actual concentration	Robust optimization (RO) prediction range	Partial least squares (PLS) prediction – 99% prediction interval			
		Bootstrap method	Faber 96 method	Serneels method	Phatak method
Pyridine concentrations (% v/v)					
5	3.14 – 8.66/3.97	-11.81 – -6.03/-8.92*	-11.63 – -6.21/-8.92*	-13.38 – -4.46/-8.92*	-11.63 – -6.21/-8.92*
10	8.6 – 17.37/8.87	-0.52 – 4.48/1.98*	-0.49 – 4.45/1.98*	-0.62 – 4.58/1.98*	-0.49 – 4.45/1.98*
15	14.3 – 19.85/17.66#	10.41 – 15.32/12.87	10.42 – 15.31/12.87	10.41 – 15.32/12.87	10.42 – 15.31/12.87
25	24.96 – 26.7/25.87	25.08 – 29.90/27.49*	25.09 – 29.90/27.49*	24.97 – 30.01/27.49	25.09 – 29.90/27.49*
30	30.19 – 32.57/31.47*	33.18 – 38.08/35.63*	33.21 – 38.05/35.63*	32.89 – 38.36/35.63*	33.21 – 38.05/35.63*
35	35.75 – 36.88/36.16*	35.27 – 40.09/37.68*	35.27 – 40.08/37.68*	35.24 – 40.11/37.68*	35.27 – 40.08/37.68*
45	44.32 – 48.64/44.98	40.15 – 45.30/42.72	40.27 – 45.18/42.72	39.55 – 45.90/42.72	40.27 – 45.18/42.72
50	47.47 – 54.64/50.24	44.41 – 49.60/47.01*	44.41 – 49.60/47.01*	43.07 – 50.94/47.01	44.41 – 49.60/47.01*
55	52.62 – 59.35/54.85	49.91 – 54.97/52.44*	50.01 – 54.87/52.44*	49.60 – 55.28/52.44	50.01 – 54.87/52.44*
65	61.65 – 67.98/65.47	59.48 – 64.55/62.02*	59.58 – 64.45/62.02*	59.07 – 64.96/62.02*	59.58 – 64.45/62.02*
70	68.08 – 75.62/71.93	65.43 – 70.80/68.11	65.62 – 70.60/68.11	65.40 – 70.83/68.11	65.62 – 70.60/68.11
75	71.96 – 77.61/75.43	69.64 – 74.66/72.15*	69.75 – 74.54/72.15*	69.59 – 74.70/72.15*	69.75 – 74.54/72.15*
85	80.52 – 85.66/83.25	79.58 – 84.90/82.24*	79.82 – 84.66/82.24*	79.44 – 85.03/82.24	79.82 – 84.66/82.24*
90	86.50 – 89.31/87.2*	83.79 – 88.85/86.32*	83.89 – 88.74/86.32*	83.43 – 89.21/86.32*	83.89 – 88.74/86.32*
95	90.08 – 93.41/91.16*	88.08 – 93.06/90.57*	88.13 – 93.00/90.57*	87.82 – 93.31/90.57*	88.13 – 93.00/90.57*
Mean Range	**4.38**	**5.1**	**4.93**	**5.84**	**4.93**
Relative error	**0.81**	**2.04**	**2.12**	**2.95**	**2.12**
RMSEP	**1.7**	**5.1**	**5.1**	**5.1**	**5.1**

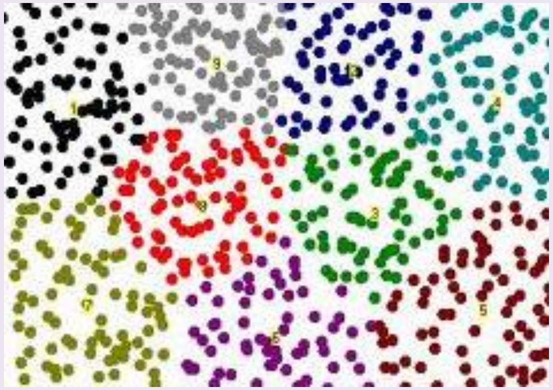
* indicates when actual value is not included in prediction range or interval
# coefficients initialized to values determined by least squares solution, for all other samples coefficients were initialized to 0.1

Other Examples



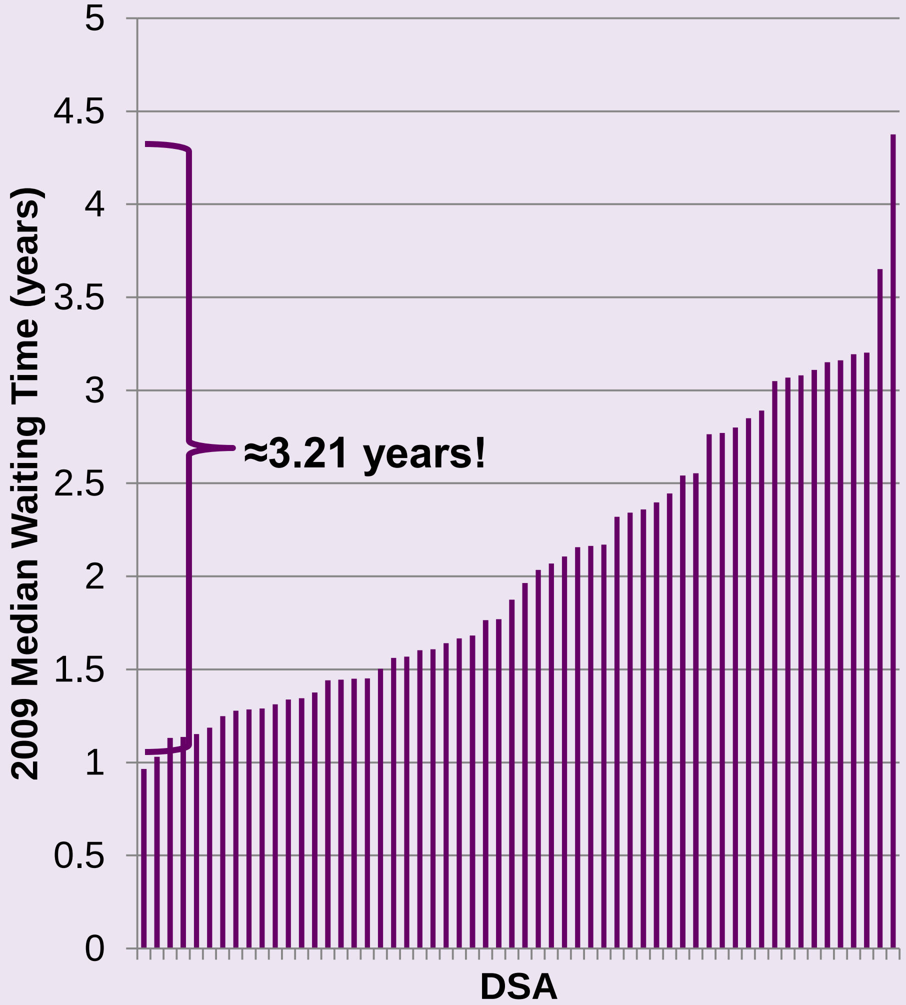
Classification

Recognition

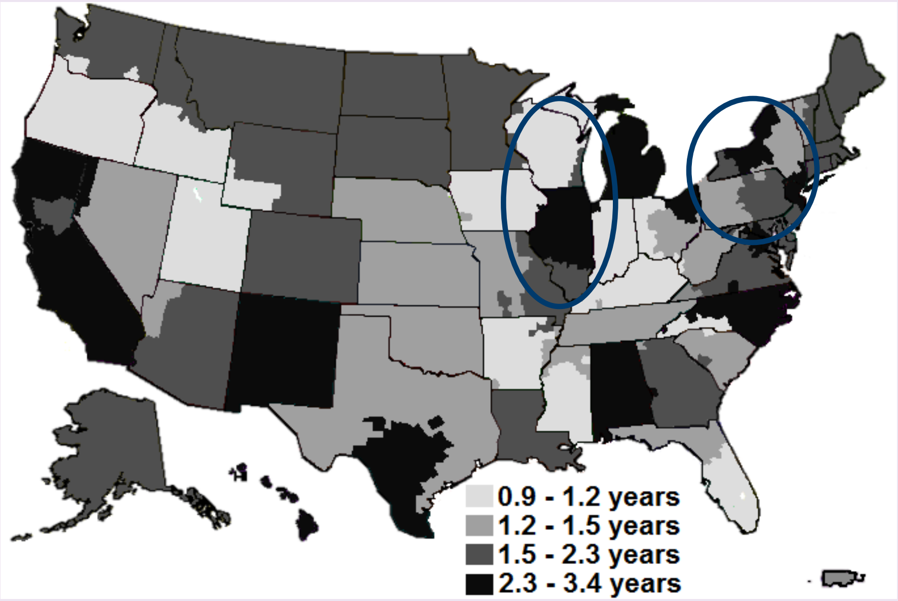


Clustering

Geographic Disparity in Kidney Allocation



2000-2009 Median Waiting Time Variability

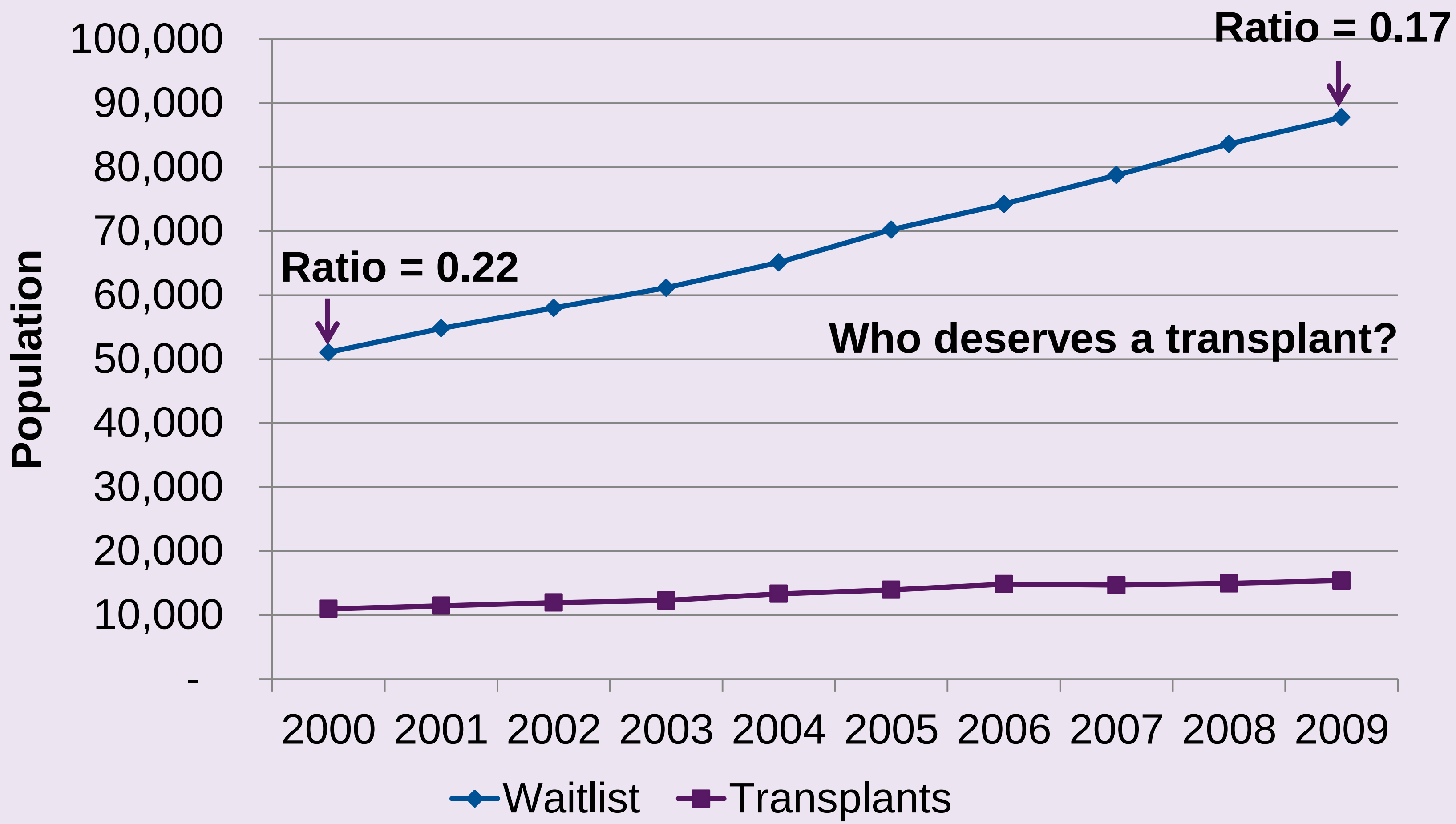


If you live in…

IL: 2.7 years vs **WI**: 1.4 years
NY: 3.0 years vs **PA**: 1.6 years

In Collaboration with:
Ashley E Davis, Mark S Daskin, Daniela P Ladner, John J Friedewald, Anton I Skaro, and Michael M Abecassis,

Geographic Disparity in Kidney Allocation

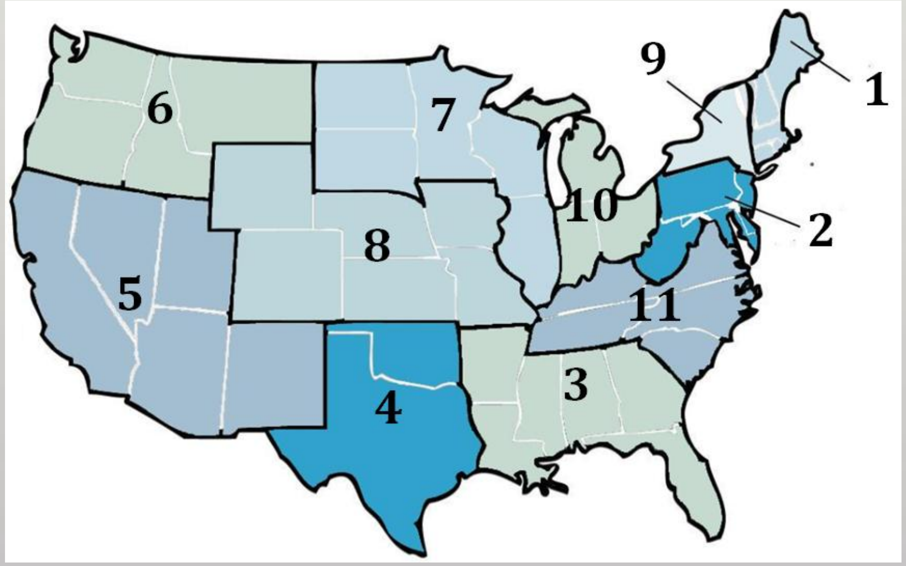


Ratio: Transplanted patients relative to waitlisted patients that year

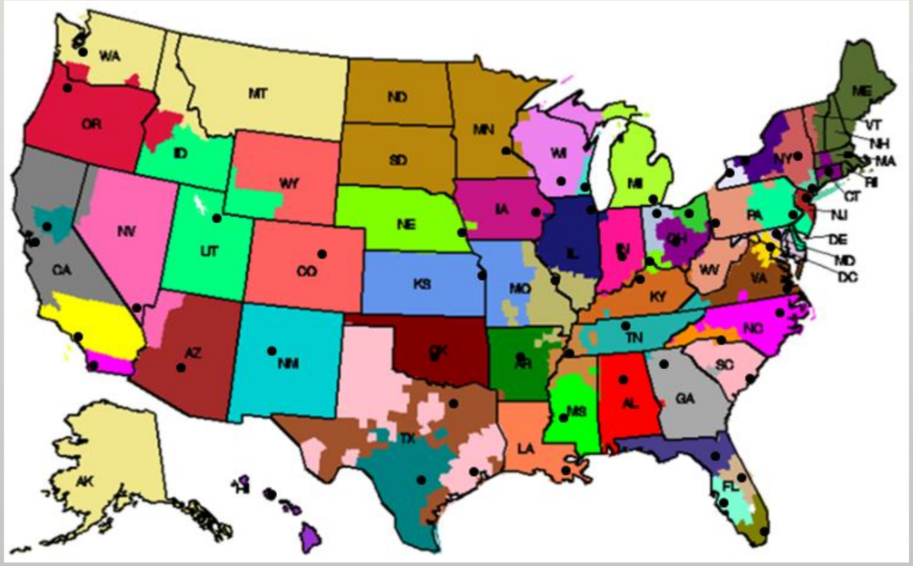
United Network for Organ Sharing (UNOS)

- Created by the National Organ Transplant Act in 1984
  - Facilitate all Organ Donation and Transplantation in the US

UNOS Regions [11]

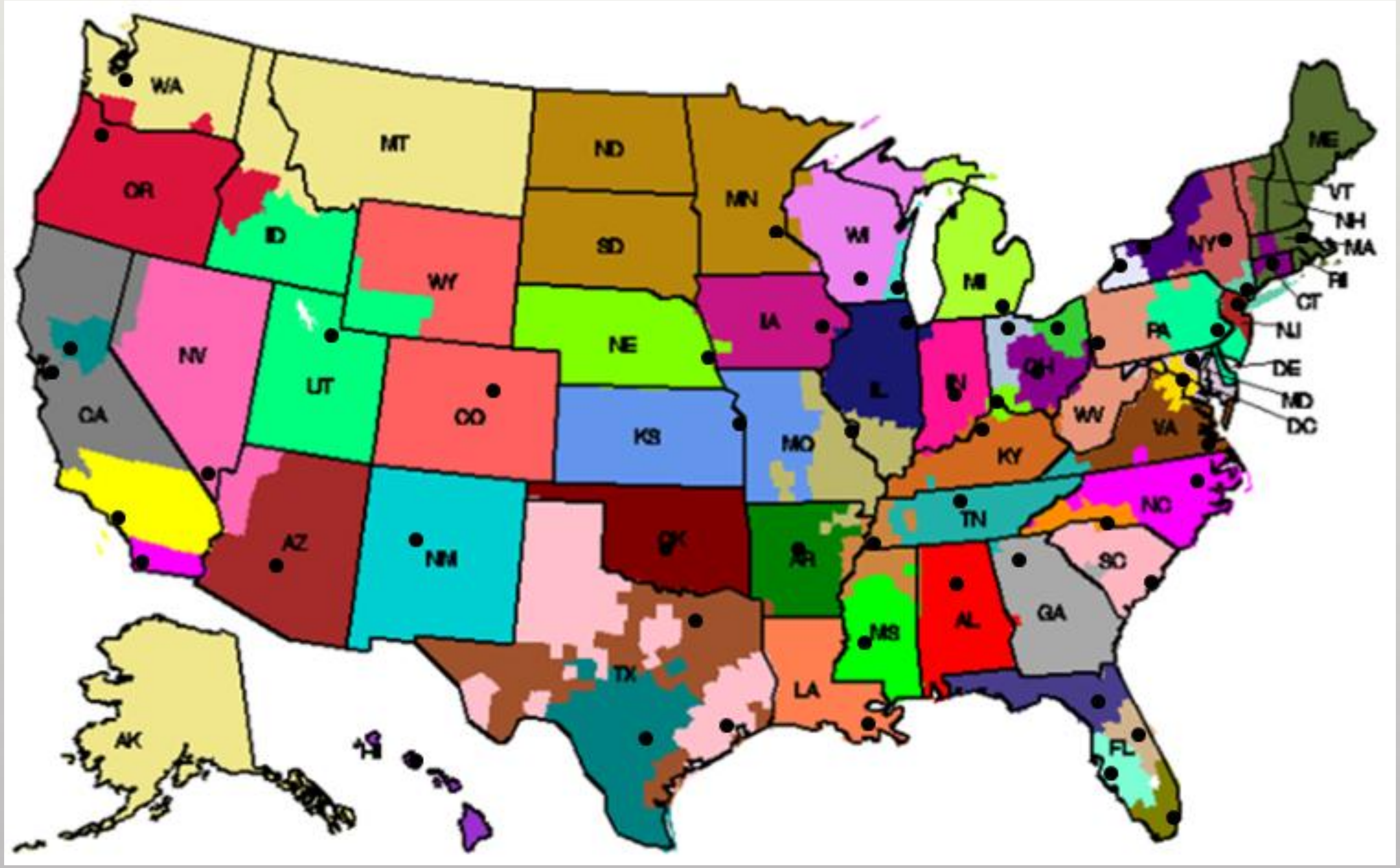


UNOS Donor Service Areas (DSA) [58]



Current Geographic Kidney Allocation: Local – Regional – National

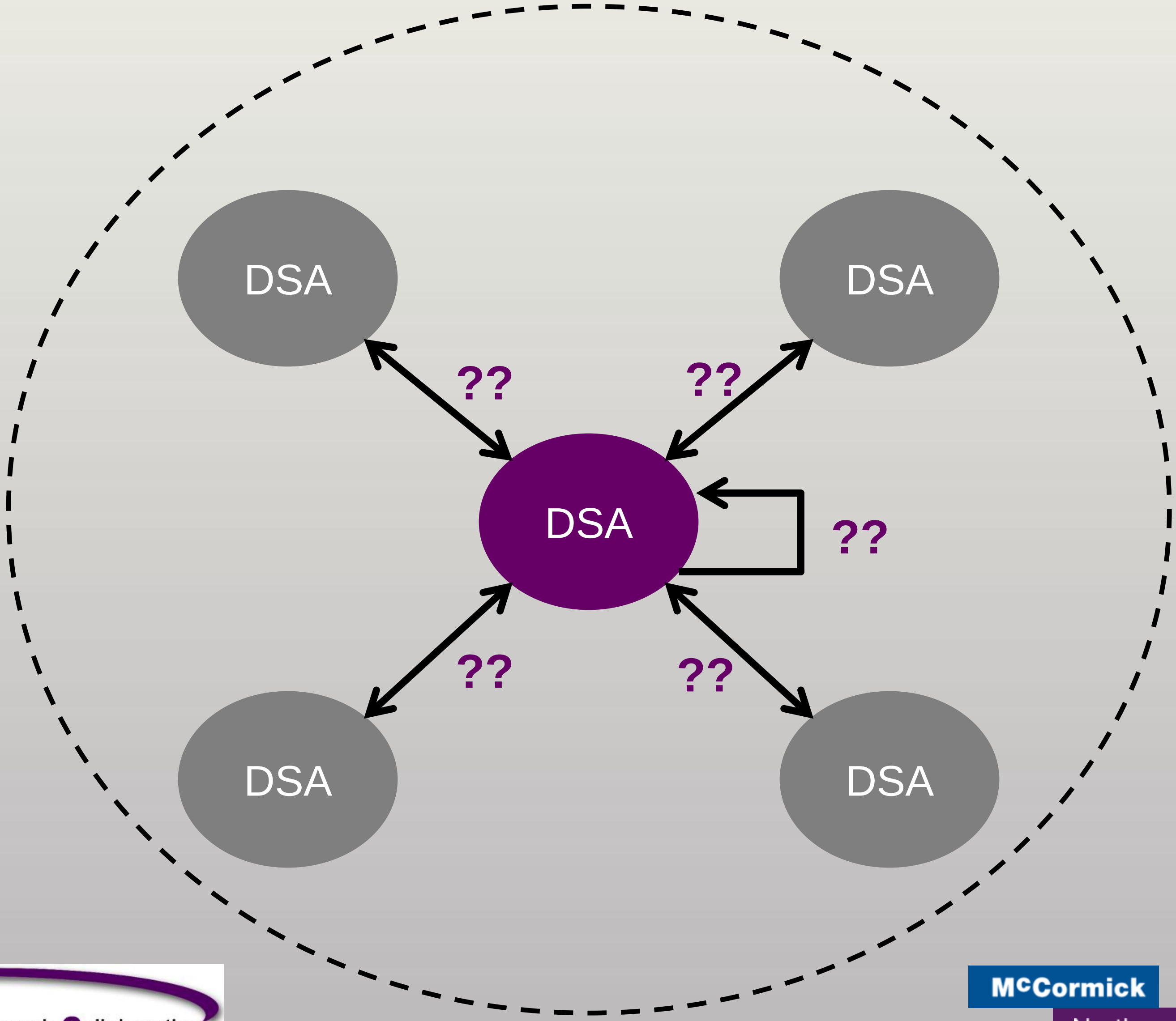
"National"



Actual 2009 DSA "Good Kidney" Sharing

Local: 76% **Regional: 8%** **National: 16%**

Proposed KSHARE Sharing Strategy



KSHARE Optimization Model

Assumptions

1. All patients treated the same
2. All kidneys accepted as optimal results dictate

Objectives

- Minimize DSA Transplant Rate Variability
  - $rate_{DSA} = \frac{Kidney\ Transplants\ in\ DSA}{Waitlist\ population\ in\ DSA}$
  - Equitable by the Institute of Medicine
- Maintain current local allocation
- 10 year phase-in with minimal:
  - DSA Sharing Partnerships
  - Changes to Yearly Sharing Strategy

KSHARE Inputs

- **Sets**
  - *I* : Set of all DSAs in the Continental US
  - *DSA(i)*: Set of Feasible Sharing DSAs for DSA *i*, including *i*
- **Parameters**
  - *w(i)*: DSA *i* Waitlist on Jan 1, 2000
  - *g(i,t)*: DSA *i* Waitlist Registrations – Non-Transplant Removals in Year 200*t*
  - *maxTR, minTR*: Limits on transplant rate range to be attained by 2009
  - *l(i)*: Yearly percentage of locally allocated kidneys in DSA *i*
  - *s(i,t)*: DSA *i* kidney procurement in year 200*t*
  - *M* >>0, T = 10 years, $\varphi = 10^{-8}$ (scaling parameter)
- **Variables**
  - ***A(i,j,t)*: Kidneys allocated from DSA *i* to DSA *j* in year 200*t***
  - *WL(i,t)*: DSA *i* Waitlist size on January 1st , 200*t*
  - *TX(i,t)*: DSA *i* total kidney transplants in year 200*t*
  - *SP(i,j)*: equals 1 if DSA *i* ever shares kidneys with DSA *j*
  - *FS(i,j,t)*: Percent of DSA *i* procurement allocated to DSA *j* in year 200*t*
  - *maxFS(i,j), minFS(i,j)*: Max/Min annually percent of DSA *i* kidneys allocated to DSA j

KSHARE Formulation

$$\min \sum_{i \in I} \left(\sum_{j \in DSA\ i} \varphi * SP(i,j) + maxFS(i,j) - minFS(i,j) \right)$$

Minimize Sharing Partnerships and Variation in Yearly Sharing Strategy

subject to:

$$WL(i,0) = w(i), \quad \forall\, i \in I$$

Initialize Waitlist

$$WL(i,t+1) = WL(i,t) + g(i,t) - TX(i,t), \quad \forall\, i \in I, t = 0 \ldots T-1$$

Update Waitlist Annually

$$TX(i,t) = \sum_{j \in DSA\ i} A(j,i,t), \quad \forall\, i \in I, t = 0 \ldots T$$

Calculate Total Yearly Transplants

$$l(i) * s(i,t) = A(i,i,t), \quad \forall\, i \in I, t = 0 \ldots T$$

Meet local allocation levels

$$\sum_{j \in DSA(i)} A(i,j,t) = s(i,t), \quad \forall\, i \in I, t = 0 \ldots T$$

Allocate all procured kidneys

$$minTR * WL(i,T) \le TX(i,T) \le maxTR * WL(i,T), \quad \forall\, i \in I$$

Transplant rate in acceptable range

$$\sum_{t=0 \ldots T} A(i,j,t) \le M * SP(i,j), \quad \forall\, i \in I, j \in DSA(i)$$

Establish Sharing Partnerships

$$FS(i,j,t) * s(i,t) = A(i,j,t), \quad \forall\, i \in I, j \in DSA(i), t = 0 \ldots T$$

Calculate Yearly Sharing Strategy

$$minFS(i,j) \le FS(i,j,t) \le maxFS(i,j), \quad \forall\, i \in I, j \in DSA(i), t = 0 \ldots T$$

Range in yearly sharing strategies

KSHARE Formulation

$$\text{minimize: } \sum_{\text{All DSAs i}} \left(\sum_{\text{All DSA j in FeasibleDSAs(i)}} (\text{sharePair}(i,j) + \text{maxSharedLim}(i,j) - \text{minSharedLim}(i,j)) \right)$$

$$\text{waitlistSize}(i,0) = \text{initialWaitlist}(i)$$

$$\text{waitlistSize}(i,t+1) = \text{waitlistSize}(i,t) + \text{growthInWaitlist}(i,t) - \text{transplants}(i,t)$$

$$\text{transplants}(i,t) = \sum_{\text{All DSAs j in FeasibleDSAs(i)}} \text{allocation}(j,i,t)$$

$$\text{localAllocation}(i) * \text{kidneysProcured}(i,t) \leq \text{sharedAllocation}(i,i,t)$$

$$\sum_{\text{All DSAs j in FeasibleDSAs(i)}} \text{allocation}(i,j,t) = \text{kidneysProcured}(i,t)$$

$$\text{minRateLim} * \text{waitlistSize}(i,T) \leq \text{transplants}(i,T) \leq \text{maxRateLim} * \text{waitlistSize}(i,T)$$

$$\sum_{\text{All Years t}} \text{allocation}(i,j,t) \leq M * \text{sharePair}(i,j)$$

$$\text{fracShared}(i,j,t) * \text{kidneysProcured}(i,t) = \text{allocation}(i,j,t)$$

$$\text{minSharedLim}(i,j) \leq \text{fracShared}(i,j,t) \leq \text{maxSharedLim}(i,j)$$

Effect of Sharing Radius

DSA Transplant Rate Statistic	Actual 2009	Feasible Sharing Radius, miles							
		370	450	500	600	900	1,200	1,500	2,700
Min Rate (%)	3.0	4.3	4.5	4.6	5.0	5.8	6.3	6.5	6.6
Max Rate (%)	30.0	30.0	15.0	12.5	12.5	12.5	12.5	12.4	12.4
Max Rate/Min Rate	10	7.0	3.3	2.7	2.5	2.2	2.0	1.9	1.9
Range in Rates (%)	27.0	25.7	10.5	7.9	7.5	6.7	6.2	5.9	5.9

Only small reductions in:
Range in Rates: 1.6%
Max/Min Ratio: 0.6

Lessons Learned

- Global sharing is not required to fix the inequity problem

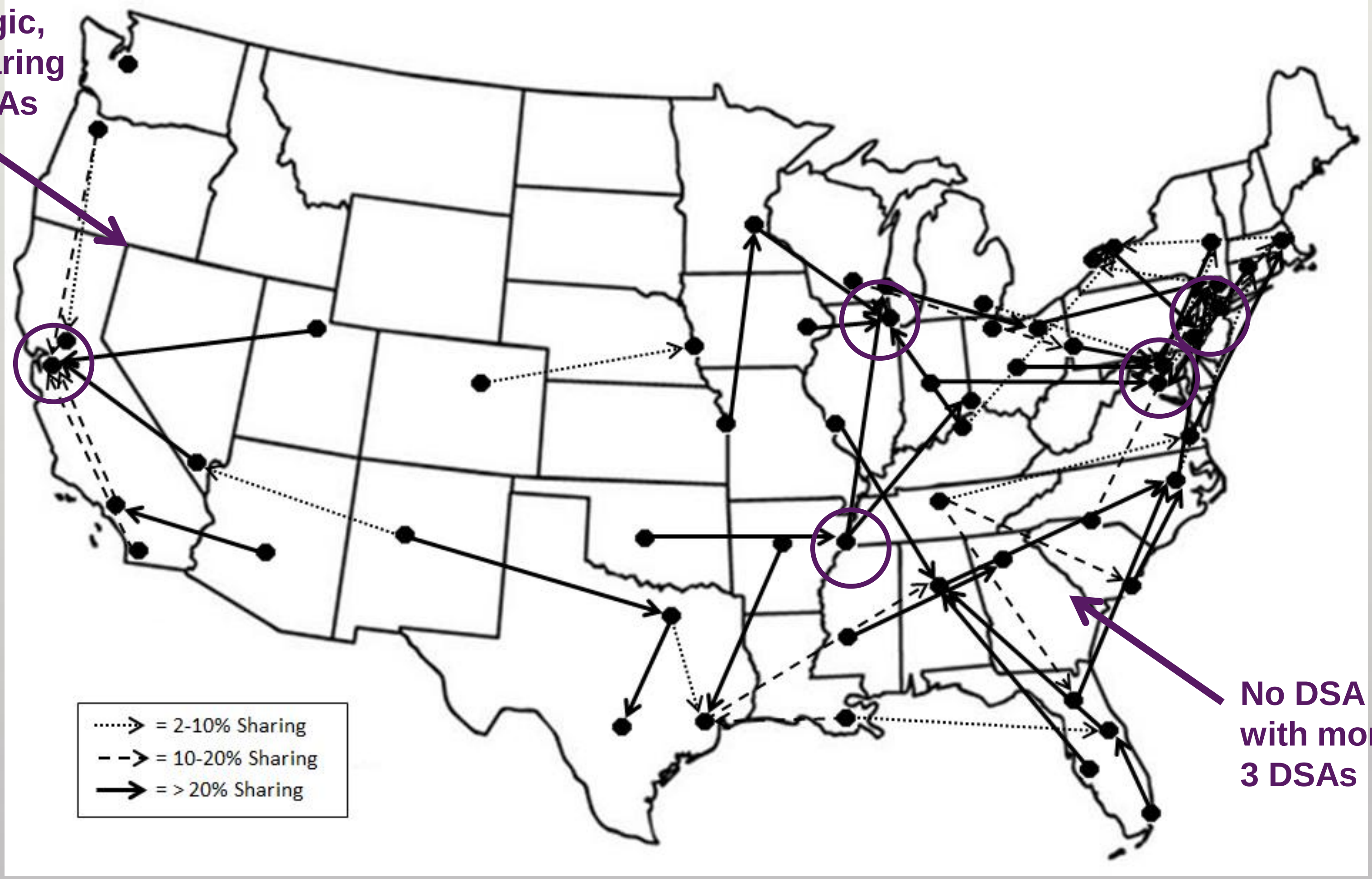
Comparison of 2000-2009 Allocation

Year	Actual Allocation				600 mile Allocation			
	Min Rate (%)	Max Rate (%)	Range (%)	Max Rate/ Min Rate	Min Rate (%)	Max Rate (%)	Range (%)	Max Rate/ Min Rate
2000	5.1	54.5	49.4	10.6	6.9	54.5	47.6	7.9
2001	5.1	54.6	49.5	10.8	5.9	38.3	32.4	6.5
2002	5.5	45.0	39.5	8.2	5.5	38.3	32.8	6.9
2003	4.7	44.1	39.4	9.4	4.7	39.5	34.8	8.4
2004	4.0	60.3	56.3	15	5.0	31.8	26.8	6.4
2005	3.9	45.8	42.0	11.9	4.4	25.0	20.6	5.7
2006	4.3	49.6	45.3	11.5	5.3	25.9	20.7	4.9
2007	4.4	37.7	33.3	8.6	4.4	23.4	19.0	5.3
2008	4.0	29.5	25.5	7.3	4.8	23.0	18.2	4.8
2009	3.0	29.9	27.0	10	5.0	12.5	7.5	2.5

Yearly KSHARE Sharing

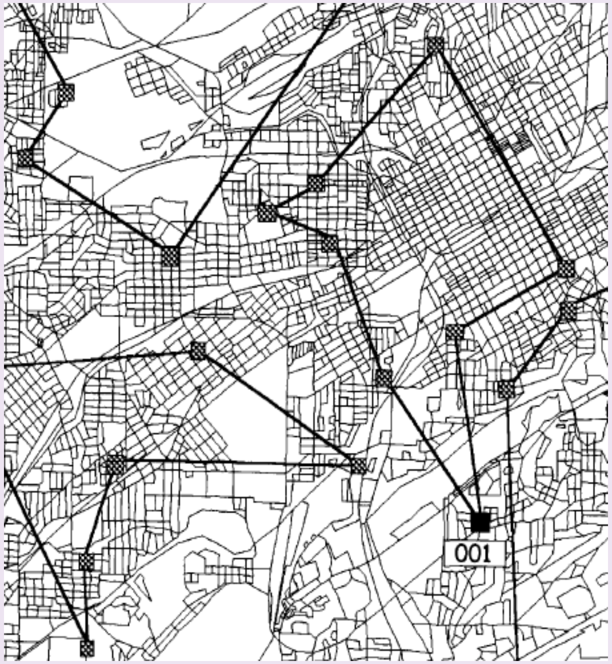
Local Allocation Increases by 3%!

More strategic, focused sharing between DSAs

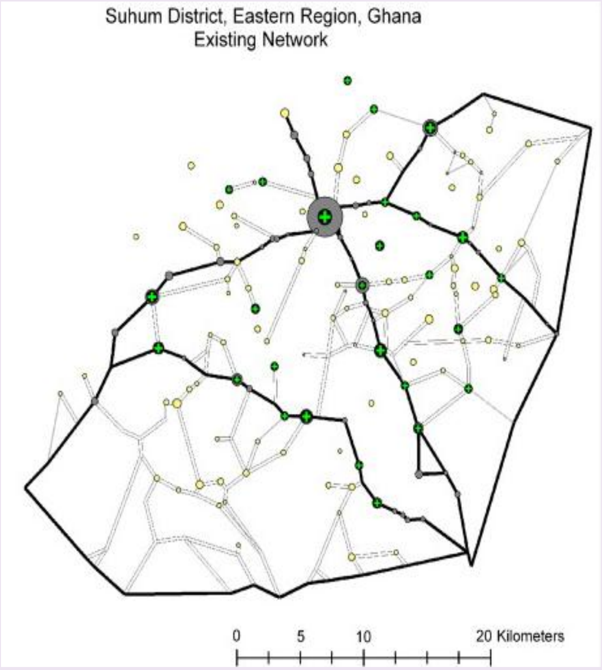


No DSA shares with more than 3 DSAs

Other Examples



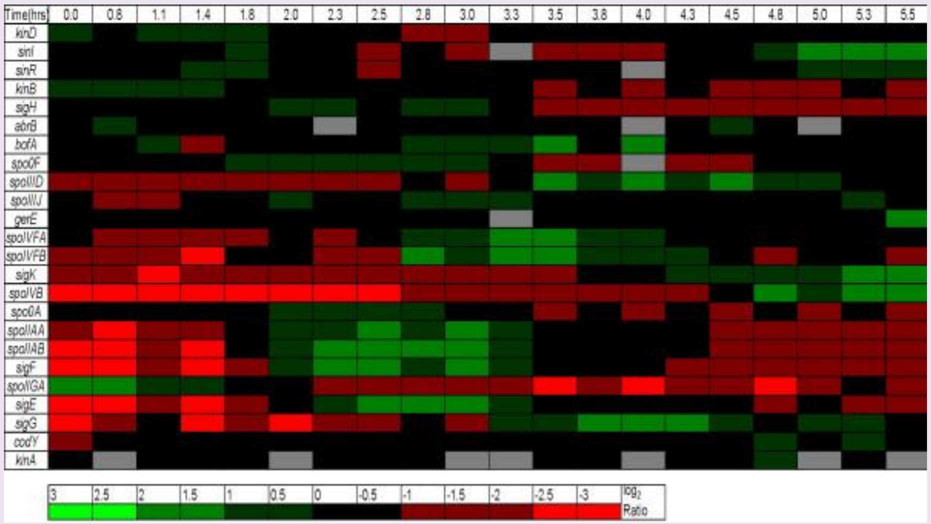
Routing of Home Health Services [1]

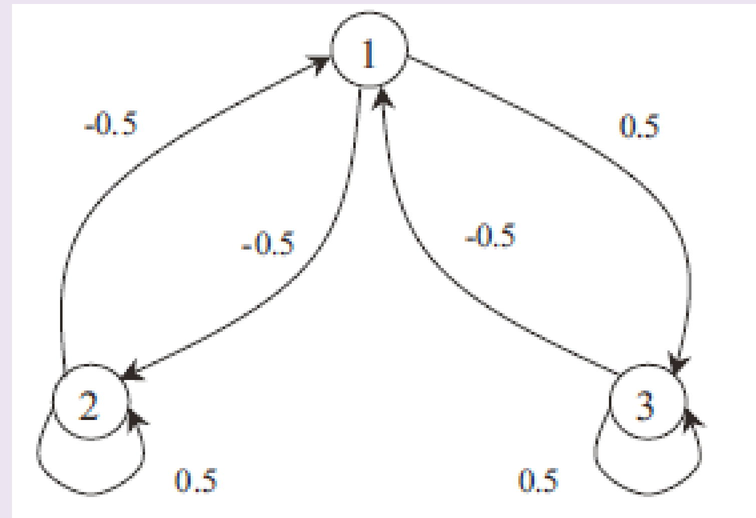


Design of Communities and/or Rural Infrastructure [2]

[1] An Integrated Spatial DSS for Scheduling and Routing Home-Health-Care Nurses, SV. Begur, D.M. Miller, and JR Weaver, Interfaces, 1997, 27(4). [2] "Improving accessibility to rural health services: The maximal covering network improvement problem," by Lisa Murawski, Richard L. Church, Socio-Economic Planning Sciences, 43(2), 2009

Reverse Engineering of Gene Regulatory Networks

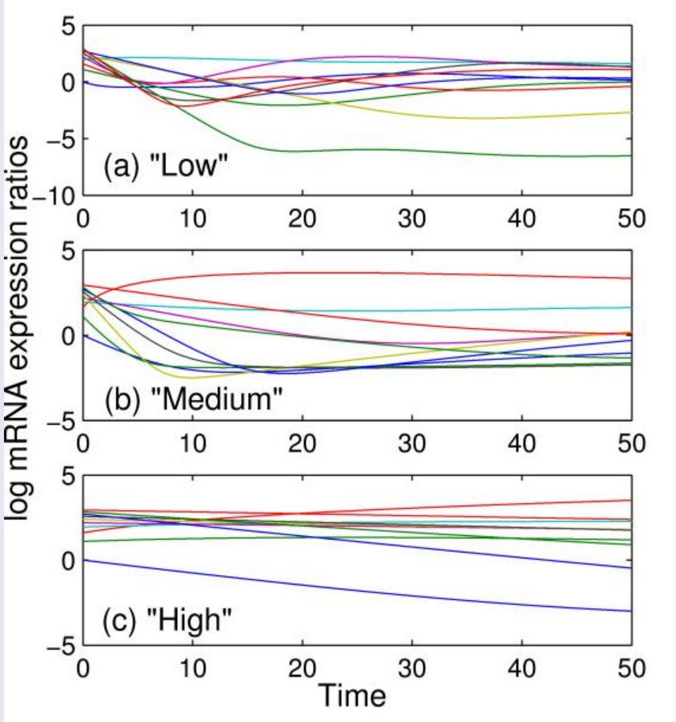




Time-course variation of subset of important genes in the sporulation cascade of *B. Anthracis*. The time-course (in hours) variation of the logarithm of expression ratios in color-coded format (green indicates up-regulation, red indicates down-regulation, grey indicates missing data and the intensity of the color indicates the level of regulation) of 24 important genes in the sporulation cascade of *B. anthracis* .

"A Model-based optimization framework for the inference of regulatory interactions using time-course DNA microarray expression data", Reuben Thomas, Carlos J. Paredes, Sanjay Mehrotra, Eleftherios T. Papoutsakis, and Vassily Hatzimanikatis, BMC Bioinformatics, 8:228, 2007.

Expressions from different Regulatory Networks



Time profiles of log mRNA expression ratios for three representative synthetic networks. Logarithm of mRNA expression ratios as a function of time for three networks. The network in (a), "Low" results in a relatively lower degree of similarity between the different gene-expression patterns in the system, the network in (b), "Medium" in a medium degree of similarity, while the network in (c), "High“ results in a relatively high degree of similarity. The units of time are arbitrary but are consistent with the units of the parameters of the system.

Mathematical Formulation for the Network Design

$$\min \sum_{j=1}^{N_t} \left(\log\left(\frac{d\tilde{m}_i(t_j)}{dt} + \beta_i \tilde{m}_i(t_j) \right) - \log(\alpha_i) - \sum_{g=1}^{n} \varepsilon_{ig} \log(\tilde{p}_g(t_j)) \right)^2 + \tau_i^2 \, \|\bar{\varepsilon}_i\|^2 \tag{9}$$

subject to

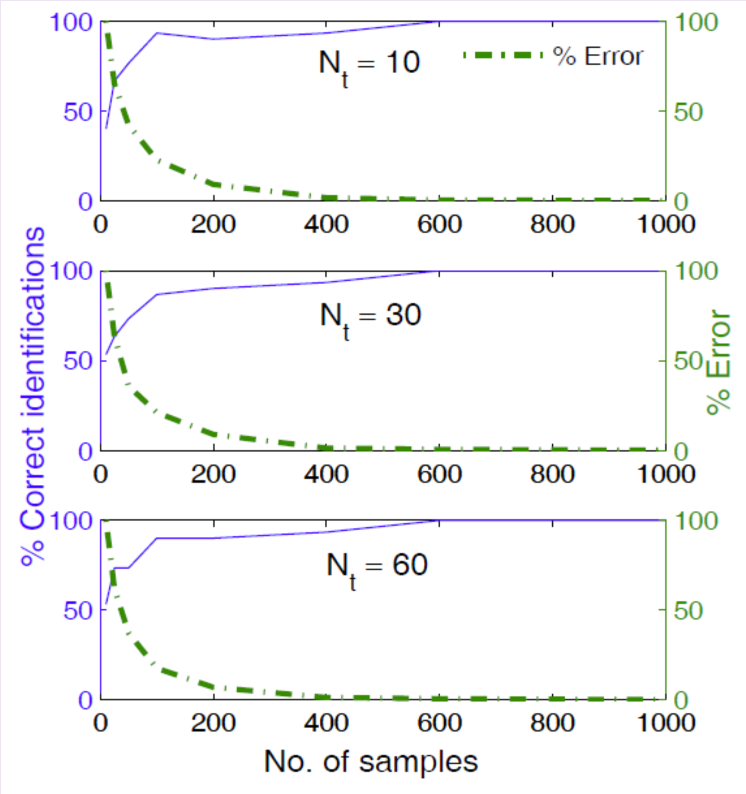
$$-DY_{ij} \leq \varepsilon_{ij} \leq DY_{ij},\ j = 1,2,\ldots,n \tag{10}$$

$$\sum_{j=1}^{N} Y_{ij} \leq k \tag{11}$$

$$Y_{ij} \in \{0,1\}, \mathrm{i}, j = 1,2,\ldots,n \tag{12}$$

$$\log(\alpha_i) \geq -A \tag{13}$$

Mathematical Formulation for the Network Design



Variation of correct identifications and identification errors with experimental samples and discretizations for "Low" network. Variation of the percentage of correctly identified interactions among 30 known interactions and the error as a percentage of the error obtained with the smallest number of samples. The variations are with respect to the number of experimental samples chosen and the number of discretizations, ***Nt***. The "experimental" data are obtained by simulation using the "Low" synthetic network (see Figure 1).

Multi-Expert Multi-Objective Decision Making

Situation: Centers for Disease Control and Prevention allocates diabetes budgets to different states for improving diabetic outcomes every year.

Question: How to allocate limited budget to the different states? Which Risk/Outcome criterion to use?

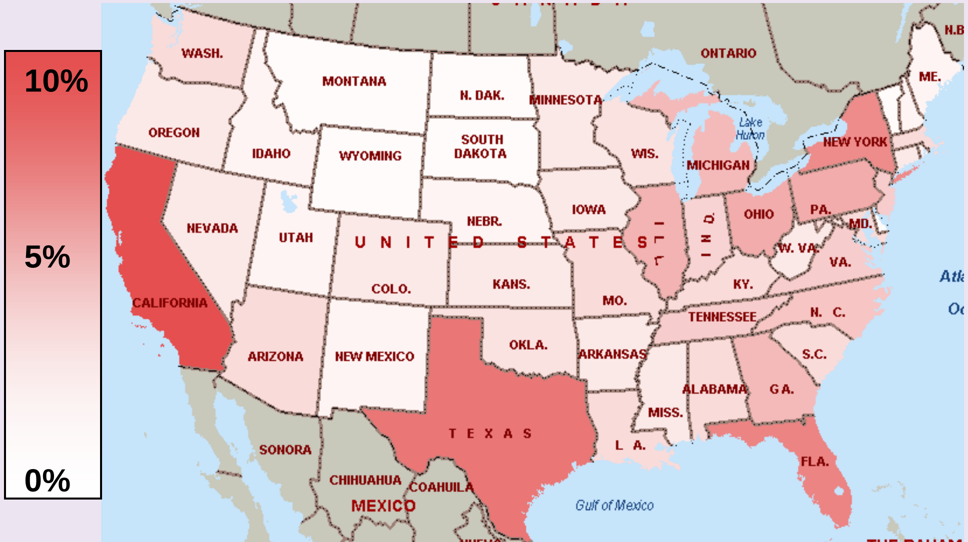
Problem: Stakeholders have different opinions about how to allocate money to states!

Current Practice: Don't know, but budgets are not correlated with patient outcomes

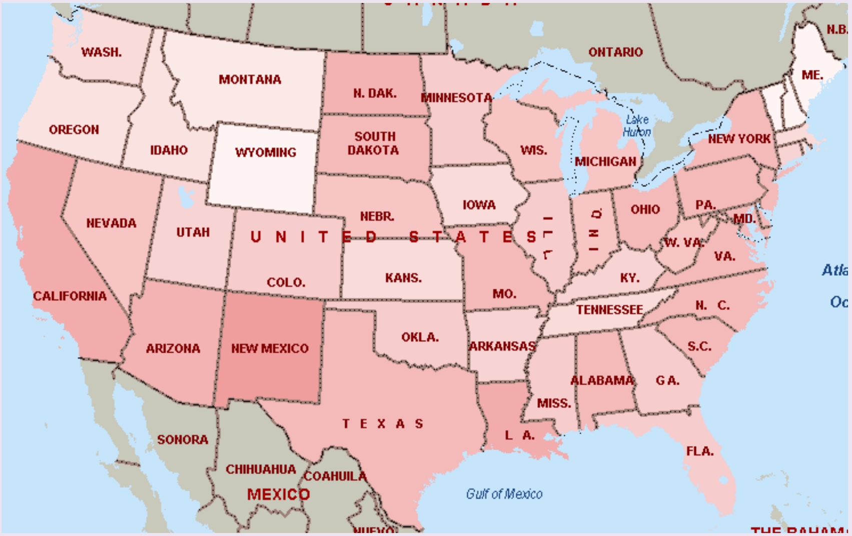
“Outcome Based State Budget Allocation for Diabetes Prevision Programs Using Multi-Criteria Optimization with Robust Weights,” Sanjay Mehrotra, and Kibaek Kim, *Health Care Management Sciences,* 2011

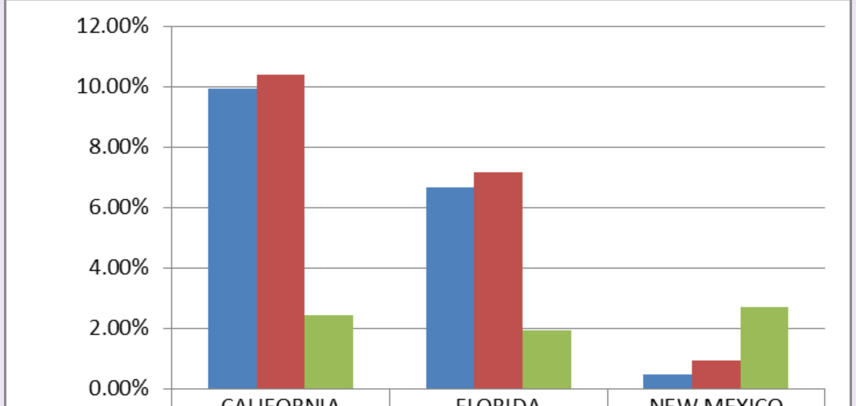
Diabetes Prevalence & Comorbidities Vary

Percentage of Diabetic Population



End Stage Renal Disease





	CALIFORNIA	FLORIDA	NEW MEXICO
Diabetes	9.92%	6.65%	0.46%
Mortality	10.38%	7.16%	0.94%
ESRD	2.43%	1.93%	2.72%

Decision Criteria from: Behavioral Risk Factor Surveillance Survey (BRFSS) Data

24 Possible Criteria

Decision Criteria	Description
1-1.	the average number of times for checking feet sores and irritations by a health professional in the past 12 month
1-2.	the average number of times for checking feet sores and irritations by themselves in the past 12 month
1-3.	the number of diabetic patients who have ever had feet sores or irritations for more than four weeks
2-1.	the number of diabetic patients who have not had an eye exam more than a month
2-2.	the number of diabetic patients having eyes affected by diabetes or having retinopathy
3-1.	the number of diabetic patients who have not had a flu shot during the past 12 months
3-2.	the number of diabetic patients who have not had a pneumonia shot during the past 12 months
4-1.	the average number of times for having a health professional checked for Hemoglobin A1c level
4-2.	the average number of times for personal checking blood glucose level
5-1.	the gap between the maximum and the minimum of the prevalence among races/ethnicities
6-1.	the number of diabetic patients who have ever diagnosed with heart attack
6-2.	the number of diabetic patients who have ever diagnosed with angina or coronary heart disease
6-3.	the number of diabetic patients who have ever diagnosed with stroke
6-4.	the number of diabetic patients who have not checked blood cholesterol more than one year
6-5.	the number of diabetic patients who have ever diagnosed with high blood cholesterol
6-6.	the number of diabetic patients who have ever diagnosed with high blood pressure
6-7.	the number of diabetic patients who are currently smoking
6-8.	the number of obese diabetic patients
6-9.	the number of diabetic patients who did not participate any physical activities or exercise during the past month
7-1.	the average number of times for seeing health professionals for diabetes in the past 12 months
7-2.	the number of diabetic patients who have never taken classes in managing diabetes
7-3.	the number of people who have been diagnosed as diabetes
8-1.	the crude rate of adults initiating treatment for diabetes-related ESRD
9-1.	the number of deaths per 1,000 population

Mapping Criteria to National Diabetes Objectives

#	Decision Criteria Group (# of criteria)	Data Source	NDO [21]	ICD-9
1	Limb Amputation (3)	BRFSS	1	250.6
2	Blindness (2)	BRFSS	2	250.5
3	Influenza and Pneumonia (2)	BRFSS	3	-
4	Glucose Control (2)	BRFSS	4	250.1 - 250.3
5	Disparity (1)	BRFSS	5	-
6	Cardiac (9)	BRFSS	6	250.7
7	Diabetic Prevalence (3)	BRFSS	-	250.8 - 250.9
8	Renal Failure (1)	NDSS	-	250.4
9	Mortality	WONDER (1)	-	-

Retrospective Principal Component Analysis (PCA) of the CDC Budget

09

Criteria	Principal Components 1	2	3	Communality
Budget	0.30	-0.28	**-0.66**	0.60
1-1	-0.25	**0.95**	0.10	0.97
1-2	-0.23	**0.95**	0.06	0.95
1-3	**0.95**	0.01	-0.11	0.92
2-1	**0.96**	-0.27	-0.07	1.00
2-2	**0.95**	0.01	-0.10	0.91
3-1	**0.96**	-0.25	-0.06	0.98
3-2	**0.97**	-0.20	-0.06	0.99
4-1	-0.21	**0.95**	-0.04	0.95
4-2	-0.26	**0.95**	0.12	0.98
5-1	**0.75**	-0.41	0.01	0.73
6-1	**0.91**	-0.34	-0.08	0.96
6-2	**0.95**	-0.28	-0.09	0.99
6-3	**0.93**	-0.31	-0.06	0.97
6-4	**0.95**	-0.24	0.02	0.96
6-5	**0.95**	-0.28	-0.09	0.99
6-6	**0.95**	-0.29	-0.07	1.00
6-7	**0.87**	-0.41	-0.03	0.93
6-8	**0.93**	-0.34	-0.07	0.99
6-9	**0.91**	-0.36	-0.05	0.96
7-1	-0.22	**0.95**	0.05	0.96
7-2	**0.96**	0.01	-0.10	0.92
7-3	**0.96**	-0.28	-0.07	1.00
8-1	0.01	-0.04	**0.88**	0.77
9-1	**0.92**	-0.30	-0.12	0.95

Principal component values: the bi-variate correlations between the actual data and the corresponding component.

The communality values: represent how well all the principal components can explain the variation in the observed data

- *All data is explained by 2 components, except Budget and ESRD which needs 3rd.*
- *ESRD and Budget have "negative correlation"*
- *Budget can not be explained by the risk criteria!*

Diabetic Relative Risk Measure and Excess Risk Function

State	Limp Amputation Risk	Glycemic Risk	Health Disparity	Mortality	Budget	Excess Risk of Mortality
Alabama	0.007135	0.005188	0.033331	0.031085	0.012626	0.018459
Alaska	0.093590	0.103378	0.003743	0.002332	0.017833	0.000000
Arizona	0.007471	0.006835	0.021334	0.025801	0.016535	0.009265
Arkansas	0.013907	0.011237	0.014212	0.017949	0.012481	0.005469
California	0.001529	0.001240	0.111161	0.157799	0.136228	0.021571
Colorado	0.014551	0.013602	0.020933	0.014483	0.013602	0.000882
Connecticut	0.015555	0.009757	0.023900	0.016730	0.010192	0.006538
Delaware	0.045371	0.051390	0.005189	0.004086	0.025840	0.000000
District of Columbia	0.078343	0.043969	0.005211	0.003936	0.043969	0.000000
Florida	0.002460	0.001507	0.079745	0.110648	0.093641	0.017006

Relative risk across states for each risk factor

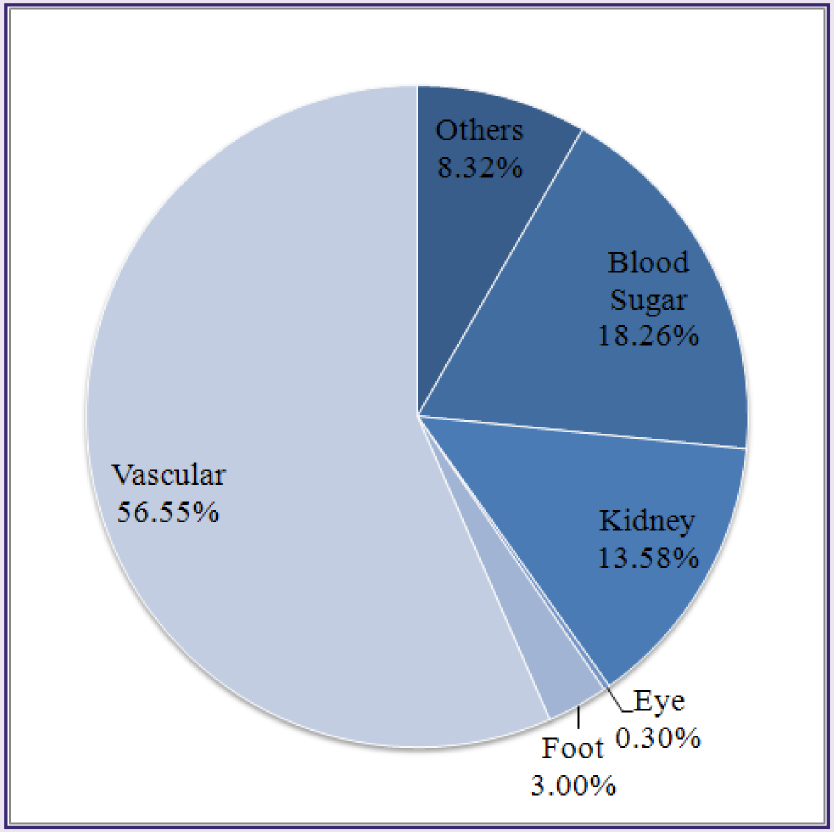
Excess Risk/Utility Function:

$$f_j(x) = \sum_{i=1}^{n} \max(z_{ij} - x_i, 0), \ j = 1, \ldots, m$$

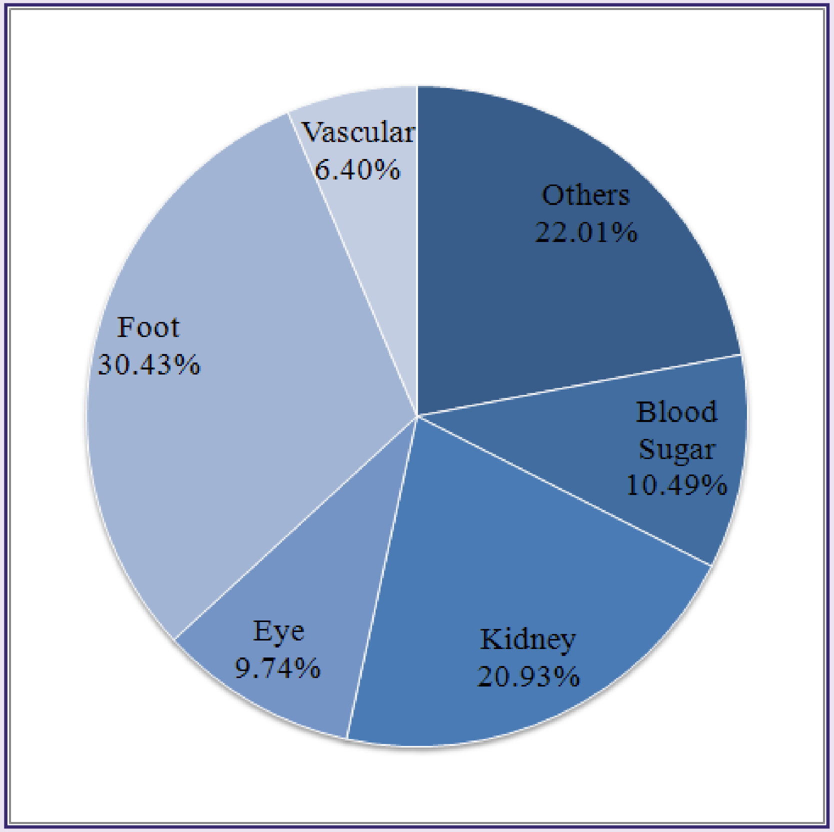
Relative risk of criterion i in state j

Allocated budget in state j

Possible Weights for Classical Weighted Sum Multi Criteria Optimization Approaches



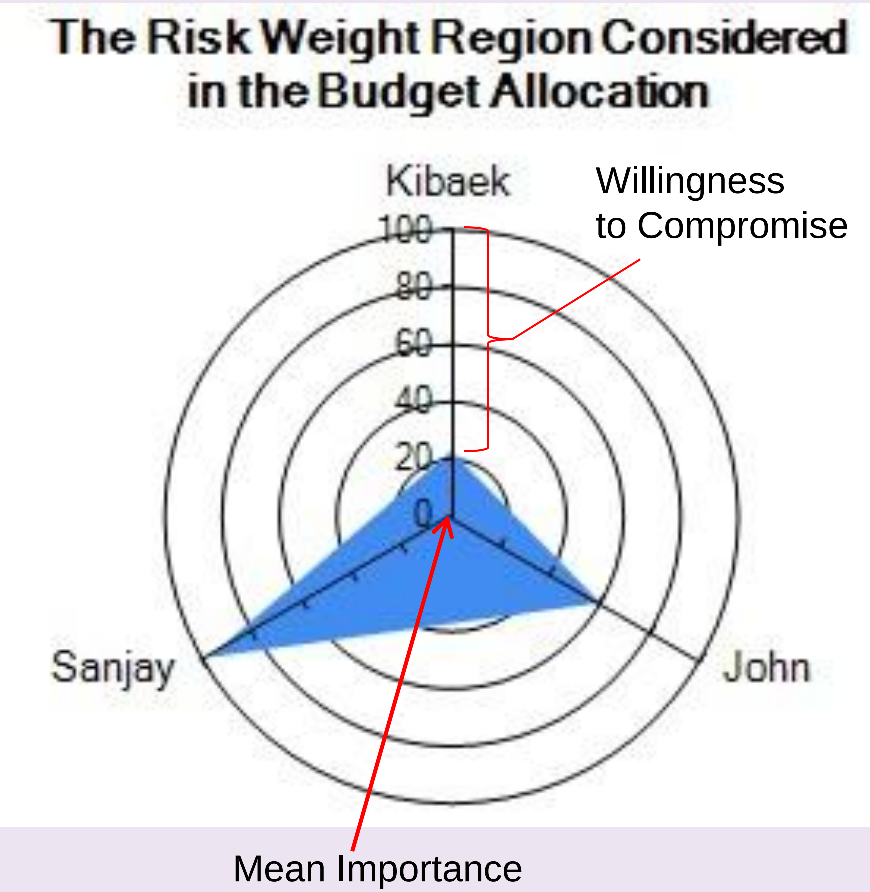
Mortality based Relative Importance



Hospital Discharge based Relative Importance

Interactive Multi-Objective Decision Making

http://optimize.iems.northwestern.edu/Diabetes



Risk Factor Weight Importance

User Profile: Sanjay Mehrotra

Willingness to Compromise: 2 %

Risk Factor	Importance	
	Absolute	Relative
Limb Amputation Risk	10 %	2.13 %
Blindness Risk	90 %	19.15 %
Influenza and Pneumonia Risk	20 %	4.26 %
Glycemia Risk	40 %	8.51 %
Disparity	50 %	10.64 %
Cardiac Risk	90 %	19.15 %
Prevalence Risk	70 %	14.89 %
Renal Failure	90 %	19.15 %
Mortality	10 %	2.12 %

Back to the Diabetes Case Study

Define Disparity for each criteria *j* as:

$$f_j(\mathbf{x}) = \sum_{i=1}^{n} (z_{ij} - x_i)_+, \quad j = 1, \ldots, m,$$
$$\text{where } (\cdot)_+ = \max\{\cdot, 0\}.$$

$$\min_{\mathbf{x} \in \mathcal{S}} \max_{\mathbf{w} \in \mathcal{P}} \sum_{j=1}^{m} w_j f_j(\mathbf{x}).$$

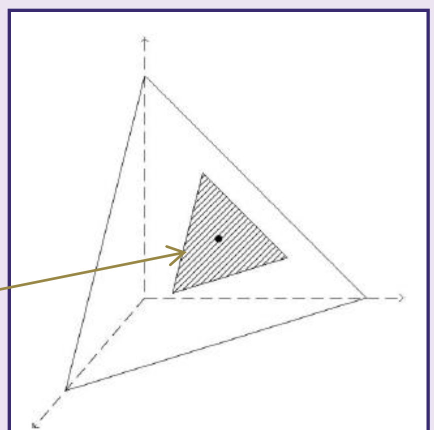
z_{ij} is budget share for state i if criteria j is the only criteria used for a proportional budget allocation according to this criteria.

x_i is the decision variable (model recommended budget)

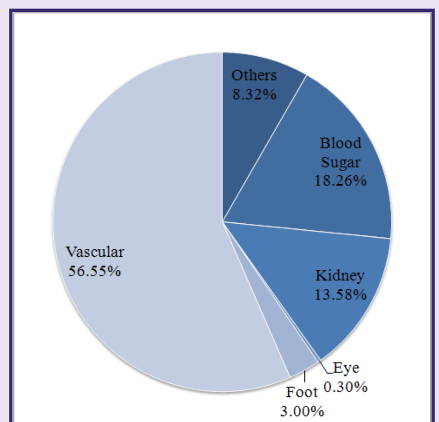
$$\mathcal{S} = \{\mathbf{x} \mid \textstyle\sum_{i=1}^{n} x_i = 1, x_i \geq 0, i = 1, \ldots, n\}.$$

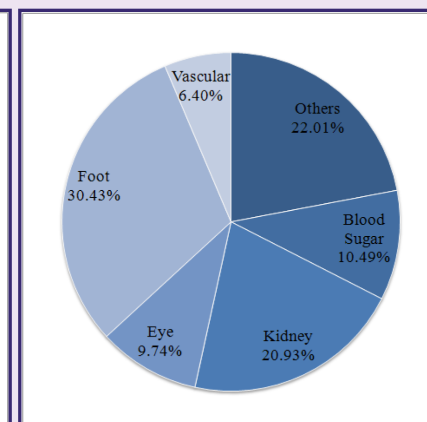
$\mathcal{P}$

Perturbation to each extreme around a center



Mortality and Discharge Ctrs





What did we learn?

State	CDC Budget		Mortality	Magnitude of Perturbation				Budget Change	
	Dollar	Prop.	Center	0.1	0.2	0.5	0.9	Relative	Absolute
AK	$424,661	2.32%	0.24%	0.25%	0.32%	1.78%	1.78%	7.4	1.54%
AL	$291,564	1.59%	3.25%	3.14%	3.14%	1.26%	1.11%	2.9	2.14%
AR	$464,177	2.54%	1.80%	1.80%	1.79%	1.42%	1.34%	1.3	0.46%
AZ	$250,017	1.37%	3.10%	3.10%	3.10%	1.61%	1.61%	1.9	1.49%
CA	$1,043,922	5.71%	16.14%	16.14%	15.92%	13.71%	15.16%	1.2	2.43%
CO	$507,359	2.77%	1.44%	1.44%	1.44%	1.44%	1.44%	1.0	0.00%
CT	$252,782	1.38%	1.45%	1.47%	1.47%	1.47%	1.20%	1.2	0.28%
DC	$261,917	1.43%	0.31%	0.31%	0.40%	4.68%	4.68%	15.3	4.37%
DE	$386,912	2.12%	0.52%	0.54%	0.55%	3.41%	2.58%	6.6	2.89%
FL	$694,394	3.80%	10.40%	10.40%	10.20%	9.36%	10.20%	1.1	1.04%
GA	$369,150	2.02%	6.44%	6.43%	6.43%	4.31%	3.92%	1.6	2.51%
HI	$328,887	1.80%	0.56%	0.57%	0.69%	3.14%	3.14%	5.6	2.58%
IA	$229,862	1.26%	1.31%	1.32%	1.37%	1.74%	1.74%	1.3	0.43%
ID	$330,291	1.81%	0.79%	0.80%	0.85%	1.92%	1.92%	2.4	1.13%

- ❖ Weight center and regions matter – though solutions are “stable” when perturbations are “reasonable”
- ❖ Some states may be significantly underfunded

Lessons Learned: Discharge Versus Mortality Center

State	Mortality Center	Hospital Discharge Center	Magnitude of Perturbation				Budget Change	
			0.1	0.2	0.5	0.9	Rel.	Abs.
AK	0.24%	0.38%	1.78%	1.78%	1.78%	1.78%	4.7	1.40%
AL	3.25%	2.52%	1.26%	1.14%	1.06%	1.11%	2.4	1.46%
AR	1.80%	1.68%	1.51%	1.46%	1.36%	1.34%	1.3	0.34%
AZ	3.10%	2.08%	1.61%	1.61%	1.61%	1.48%	1.4	0.61%
CA	16.14%	17.54%	14.29%	14.55%	15.11%	15.16%	1.2	3.24%
CO	1.44%	1.45%	1.44%	1.44%	1.44%	1.44%	1.0	0.01%
CT	1.45%	1.53%	1.47%	1.47%	1.28%	1.20%	1.3	0.33%
DC	0.31%	0.52%	4.68%	4.68%	4.68%	4.68%	9.0	4.15%
DE	0.52%	0.59%	2.58%	3.83%	2.58%	2.58%	6.5	3.24%
FL	10.40%	10.20%	9.76%	9.76%	9.98%	10.20%	1.0	0.44%
GA	6.44%	4.28%	3.92%	3.92%	3.92%	3.92%	1.1	0.36%
HI	0.56%	0.97%	3.14%	3.14%	3.14%	3.14%	3.2	2.16%
IA	1.31%	1.66%	1.66%	1.66%	1.66%	1.66%	1.0	0.00%
ID	0.79%	1.07%	1.92%	1.92%	1.92%	1.92%	1.8	0.84%

- Initial recommendations are different but they approach each other as weight regions are enlarged
- Rest of the conclusions are similar!

Policy Implications: Reverse Engineering

- Find Relative importance (weights) from a recommended budget.

$$\text{(WSO)} \quad \min_{\mathbf{x}\in\mathcal{S}} \sum_{j=1}^{m} w_j f_j(\mathbf{x})$$

$$\begin{aligned}
\max_{\lambda,\pi,w} \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} z_{ij}\lambda_{ij} - \pi \qquad \text{(InvLP)} \\
s.t. \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} z_{ij}\lambda_{ij} - \pi - \sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}w_j = 0, \\
& \sum_{j=1}^{m} \lambda_{ij} - \pi \leq 0, \quad i = 1,\dots,n, \\
& \lambda_{ij} \leq w_j, \quad i = 1,\dots,n,\ j = 1,\dots,m, \\
& \sum_{j=1}^{m} w_j \leq 1, \\
& \lambda_{ij} = 0, \quad (i,j) \in \mathcal{K}, \\
& \lambda_{ij} \geq 0, \quad (i,j) \notin \mathcal{K}, \\
& \pi \geq 0 \\
& w_j \geq 0, \quad j = 1,\dots,m.
\end{aligned}$$

$$\mathcal{K} = \{(i,j) \mid \hat{r}_{ij} + \hat{x}_i > z_{ij},\ i = 1,\dots,n,\ j = 1,\dots,m\}$$

$$\textit{where } \hat{r}_{ij} = \max\{z_{ij} - \hat{x}_i, 0\} \textit{ for all } i,j.$$

- Recommended weights emphasize regular self foot exam (1-3) and eye exam (2-2), blood glucose self checkup (4-2) and patient education (7-1) to reduce geographical disparity.
- In short, proper physician follow-up and education (7-1, 7-2) will help reduce disparity.

Conclusion

- Highly versatile technology for data analysis and decision making
- It is well developed with continued development
- Numerous use examples in areas other than health
- Whenever your problems can be framed with an “objective” to be minimized/maximized, you have an optimization problem!!!

Thank you