Keynote Address: **Healthcare Analytics: Leveraging Predictive and Prescriptive Methods to Prevent and Treat Diseases**

Brian Denton
University of Michigan
July 13, 2023

IFORS2023
The 23rd Conference of the International Federation of Operational Research Societies
July 10-14 · SANTIAGO, CHILE

IFORS International Federation of Operational Research Societies

ICHIO Instituto Chileno de Investigación Operativa

alio asociación latino iberoamericana de investigación operativa

Universidad de Chile

Pontificia Universidad Católica de Chile

ISCI Instituto Sistemas Complejos de Ingeniería

Brian Denton is the Stephen M. Pollock Professor of Industrial and Operations Engineering and the Chair of the Department of Industrial and Operations Engineering at the University of Michigan. His research interests are in data analytics and data-driven optimization under uncertainty with applications to medicine, public health, and healthcare delivery. He is a Professor in the Department of Urology (by courtesy) at Michigan Medicine and a member of the Institute for Healthcare Policy and Innovation and the Cancer Center at the University of Michigan. His research has been funded by the National Science Foundation, the Agency for Healthcare Research and Quality, the National Institutes of Health, the U.S. Department of Veterans Affairs, and industry research contracts. He is past President of the Institute for Operations Research and the Management Sciences (INFORMS), and he is an elected Fellow of INFORMS.

COLLEGE OF ENGINEERING
INDUSTRIAL & OPERATIONS ENGINEERING
UNIVERSITY OF MICHIGAN

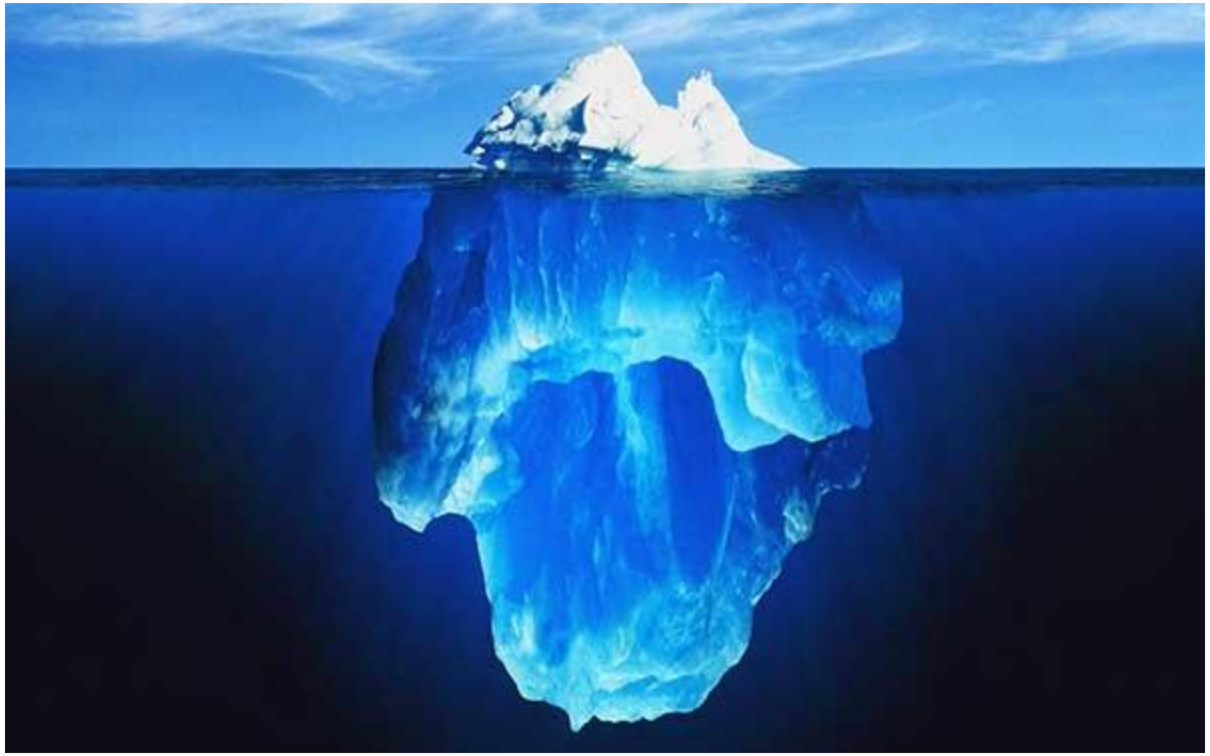
Healthcare Analytics: Predictive and Prescriptive Methods to Prevent and Treat Diseases

July 13, 2023

Brian Denton
Stephen M. Pollock Collegiate Professor
Department of Industrial and Operations Engineering
University of Michigan

(Download slides)

Healthcare Data



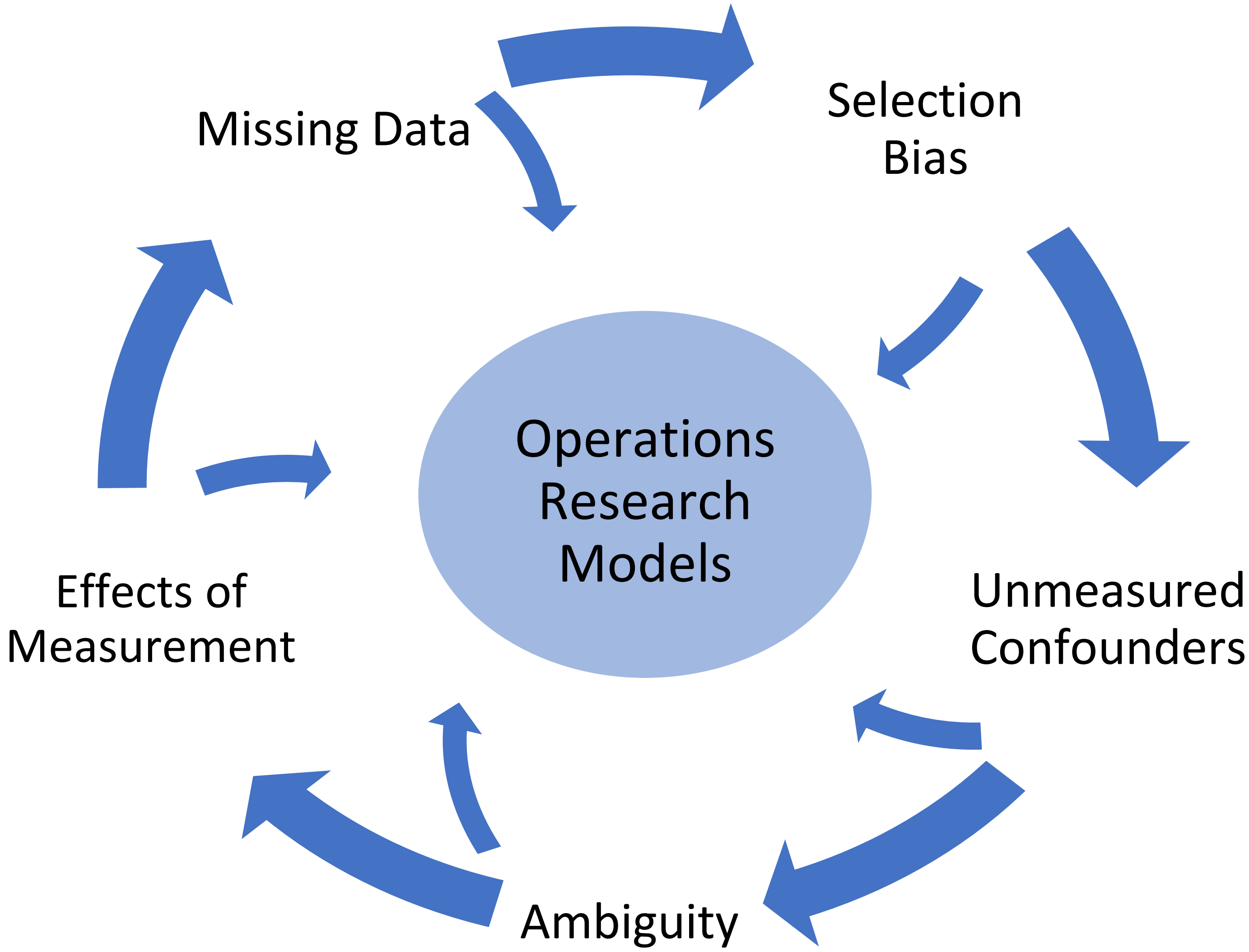
~~**Randomized Controlled Trials**~~

Observational Data: Patient data collected through observations of the natural healthcare delivery process during routine medical care.

Observational Data

- *Demographics:* age, sex, race, ethnicity, geography,…
- *Encounters:* blood pressure, weight, symptoms,…
- *Labs:* cholesterol, blood sugar, creatinine,…
- *Procedures:* biopsy, endoscopy, imaging,…
- *Insurance claims:* health services, prescription refills,…

A Whirl-Wind of Problems (Opportunities?)



Three Examples of OR & Analytics in Medicine

1. Prevention

2. Diagnosis

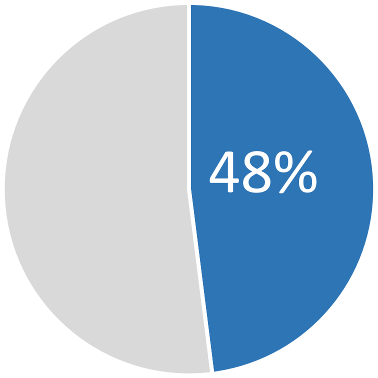
3. Treatment

1. Prevention

Setting: *Prevention of cardiovascular disease*

OR Challenge: *sequential decisions under uncertainty with sparse data*

1 in 3 deaths are due to cardiovascular disease (CVD)



$407 Billion

Percentage of people at risk of CVD in the U.S.

Annual cost of CVD in the U.S.

Heart Disease and Stroke Statistics - 2021 Update. American Heart Association

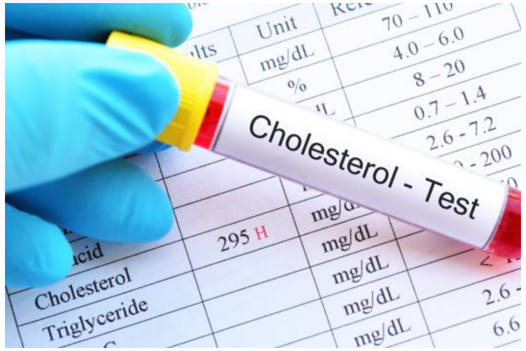
Cholesterol monitoring recommendations vary from 3 months to 6 years between testing

American College of Cardiology (ACC) Policy *

Age:

- Under 75: 4 to 6 years
- Over 75: 1 to 2 years

On treatment: 3 to 12 months



* Grundy, S. M. et al. (2018). 2018 AHA/ACC Guideline on the Management of Blood Cholesterol. American College of Cardiology 139 (25):e1082–e1143.

Should CVD risk factors be used to recommend cholesterol screening?

- Cholesterol
- Blood pressure
- Age
- Race
- Sex
- Smoking habits
- Treatment

Risk Calculator → 10-year risk of a heart attack or stroke

<u>Physician's Decision:</u> when to recommend the patient return for cholesterol screening

Finite horizon Markov Decision Process to maximize expected societal rewards

Decision epochs: t

40-year decision horizon with quarterly decision epochs

States: s_t

Static and dynamic risk factors, health outcomes

Actions: a_t

When the patient is advised to have next cholesterol test

Transition probabilities: $p_t(s_{t+1}|s_t, a_t)$

Cholesterol, blood pressure, risk of CVD

Rewards: $r_t(s_t, a_t)$

Expected societal benefits and costs

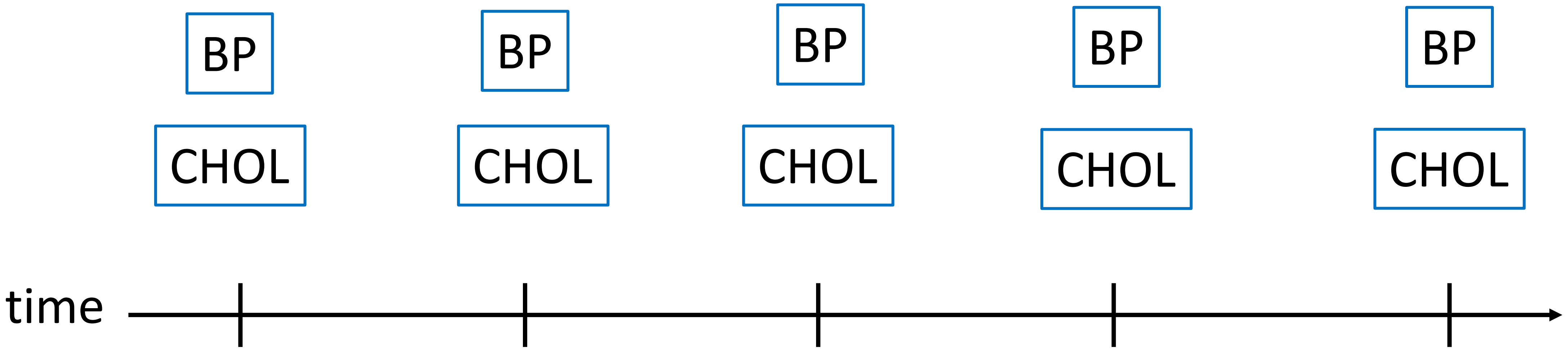
(Bellman's Equations)

Optimal Policy

State: s_t → **Action:** a_t

(Chol, BP, Age…) next Chol test

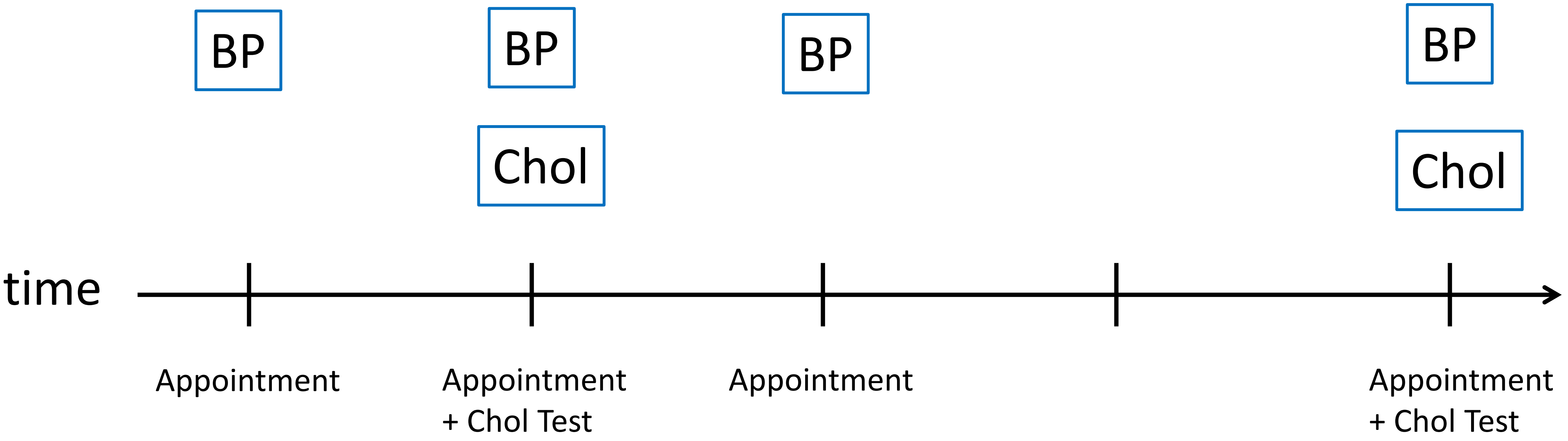
For complete data, transition probabilities are based on state transition frequency



S_{ij} ≔ Number of observations from state i to state j in one epoch.

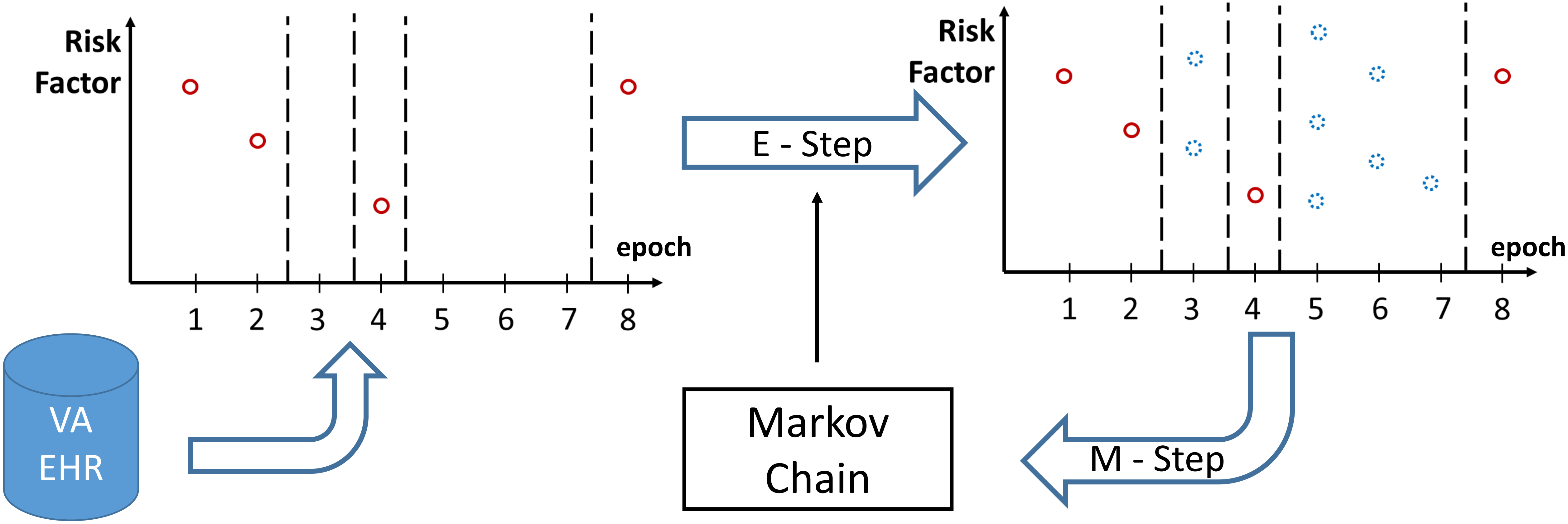
$P_{ij} = \frac{S_{ij}}{\sum_k S_{ik}}$ Fraction of S_{ij} over all observations of transitions from state i.

In reality, observational data are sparse



BP = Blood Pressure; Chol = Cholesterol

E-M Algorithm estimates transition probabilities for unequally spaced data



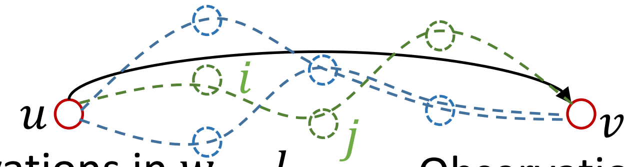
Dempster AP, Laird NM, Rubin DB (1978). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1-22.

Iterative estimation of transition probabilities using EM Algorithm

$P_{ijl,uvw}$ = Probability that a transition between i and j occurs in l epochs given the observations.

O_{uvw} ≔ Number of observations from state u to state v in $\boldsymbol{w}$ **epochs.**

E - Step



Observations in $w - l$

Observations in w

$S_{ij}(k) \coloneqq S_{ij}$ in iteration k

$$P_{ij}(k) = \frac{S_{ij}(k)}{\sum_l S_{il}(k)}$$

M - Step

Yeh, H.W., et al (2010). Estimating transition probabilities for ignorable intermittent missing data in a discrete-time Markov chain. *Communications in Statistics: Simulation and Computation* 39(2):433–448.

Finite horizon Markov Decision Process to maximize expected societal rewards

Decision epochs: t

40-year decision horizon with quarterly decision epochs

States: s_t

Static and dynamic risk factors, health outcomes

Actions: a_t

When the patient is advised to have next cholesterol test

Transition probabilities: $p_t(s_{t+1}|s_t, a_t)$

Cholesterol, blood pressure, risk of CVD

Rewards: $r_t(s_t, a_t)$

Expected societal benefits and costs

(Bellman's Equations)

Optimal Policy

State: s_t $\Longrightarrow$ **Action:** a_t

(Chol, BP, Age…) next Chol test

Optimal recommendations depend on the patient's age, race, sex, and CVD risk

Overall patient groups: → Recommendations depends on the patient's age and CVD risk

White men

Younger patients → More appointments compared to ACC Guideline

Black men

More appointments compared to white men.

White women

Fewer appointments compared to white men.

Otero-Leon, D, Lavieri, M., Denton, B., Sussman, J., Hayward, R. "Monitoring policy in the context of preventive treatment of cardiovascular disease." *Health Care Management Science* 26, no. 1 (2023): 93-116.

2. Diagnosis

Setting: *Imaging to detect metastatic cancer in patients diagnosed with prostate cancer*

OR Challenge: *machine learning, selection bias, class imbalance*

Imaging modalities to detect metastases in newly diagnosed prostate cancer patients

Bone Scan (BS)

- Detect bone metastasis

Computed Tomography (CT)

- Detects lymph node metastasis

Harms of not imaging

- Metastatic cancer may go undetected
- Missed diagnoses subject patients to unnecessary treatments (e.g., radical prostatectomy)
- Appropriate treatment (e.g., chemotherapy) is delayed

Harms of imaging

- Potentially harmful radiation exposure
- Incidental findings that require painful and risky follow-up procedures (e.g., bone biopsy)
- Blocks access to imaging resources for other patients and unnecessarily increases healthcare costs

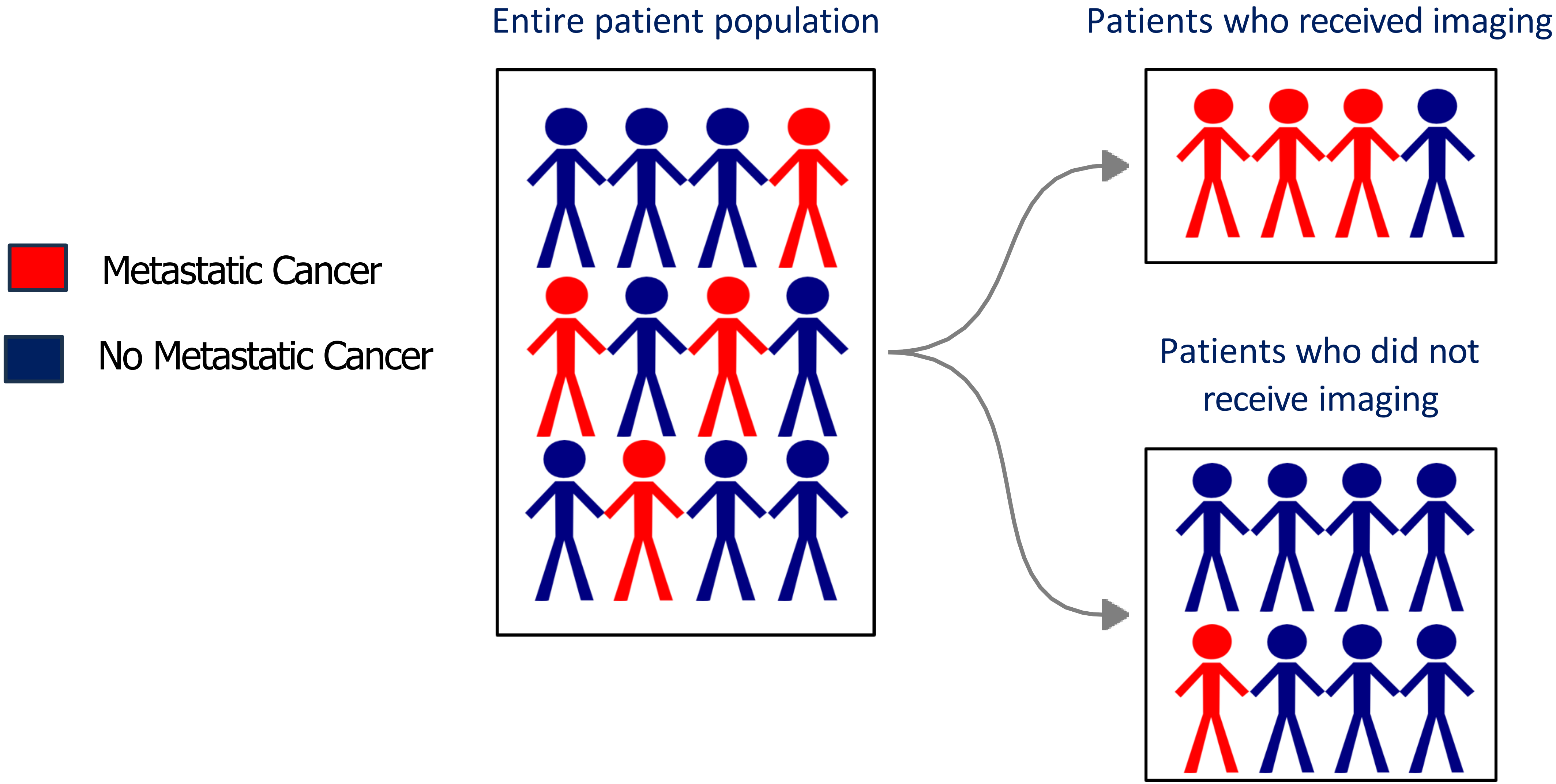
How can risk factors for metastatic cancer help decide which patients to recommend for imaging?

- Age
- Race and ethnicity
- Biomarkers
- Pathology
- Clinical tumor stage (e.g., T1a/b/c, T2a/b/c, T3/4)

MUSIC
Michigan Urological Surgery Improvement Collaborative

<u>Physician’s Decision:</u> when to recommend imaging for patients newly diagnosed with prostate cancer

Selection bias



Effects of selection bias

Guidelines (G)	Uncorrected Sensitivity	Uncorrected Specificity	Bias-corrected Sensitivity	Bias-corrected Specificity
Bone scan				
EAU	97.9	33.4	84.5	75.7
AUA	97.9	43.5	81.2	82.0
CT scan				
EAU	98.4	36.5	89.9	74.4
AUA	96.8	49.2	87.2	82.5

EAU: European Association of Urology; AUA: American Urology Association

Begg, C. B., Greenes, R. A. "Assessment of diagnostic tests when disease verification is subject to selection bias," *Biometrics*, 39:207, 1983.

Effects of verification bias

Guidelines (G)	Uncorrected		Bias-corrected	
	Sensitivity	Specificity	Sensitivity	Specificity
Bone scan				
EAU	97.9	33.4	84.5	75.7
AUA	97.9	43.5	81.2	82.0
CT scan				
EAU	98.4	36.5	89.9	74.4
AUA	96.8	49.2	87.2	82.5

Begg, C. B., Greenes, R. A. "Assessment of diagnostic tests when disease verification is subject to selection bias," *Biometrics*, 39:207, 1983.

Correcting for selection bias

Estimate sensitivity and specificity based on the entire population:

$$\Pr(Disease\ Present|G+)P(G+) + P(Disease\ Present|G-)P(G-)$$

$$P(G+|Disease\ Present) = \frac{P(Disease\ Present|G+)P(G+)}{P(Disease\ Present)}$$

$$\Pr(Disease\ \mathbf{not}\ Present|G+)P(G+) + P(Disease\ \mathbf{not}\ Present|G-)P(G-)$$

$$P(G-|Disease\ \mathbf{not}\ Present) = \frac{P(Disease\ \mathbf{not}\ Present|G-)P(G-)}{P(Disease\ \mathbf{not}\ Present)}$$

Main Assumptions: Data missing at random; Factors considered by the guideline are the only factors that influence imaging decisions.

Begg, C. B., Greenes, R. A. "Assessment of diagnostic tests when disease verification is subject to selection bias," *Biometrics*, 39:207, 1983.

Guideline optimization – which patients should be imaged?

- Two important challenges:
  - Learning from <u>unlabeled data</u>
    - Not all patients receive imaging at diagnosis
  - Learning from <u>imbalanced data</u>
    - A minority of patients have metastatic cancer
- To address these challenges, we combined <u>semi-supervised</u> and <u>cost-sensitive learning</u>

Cost-sensitive Laplacian Kernel Logistic Regression

Higher cost on missing metastatic cancers

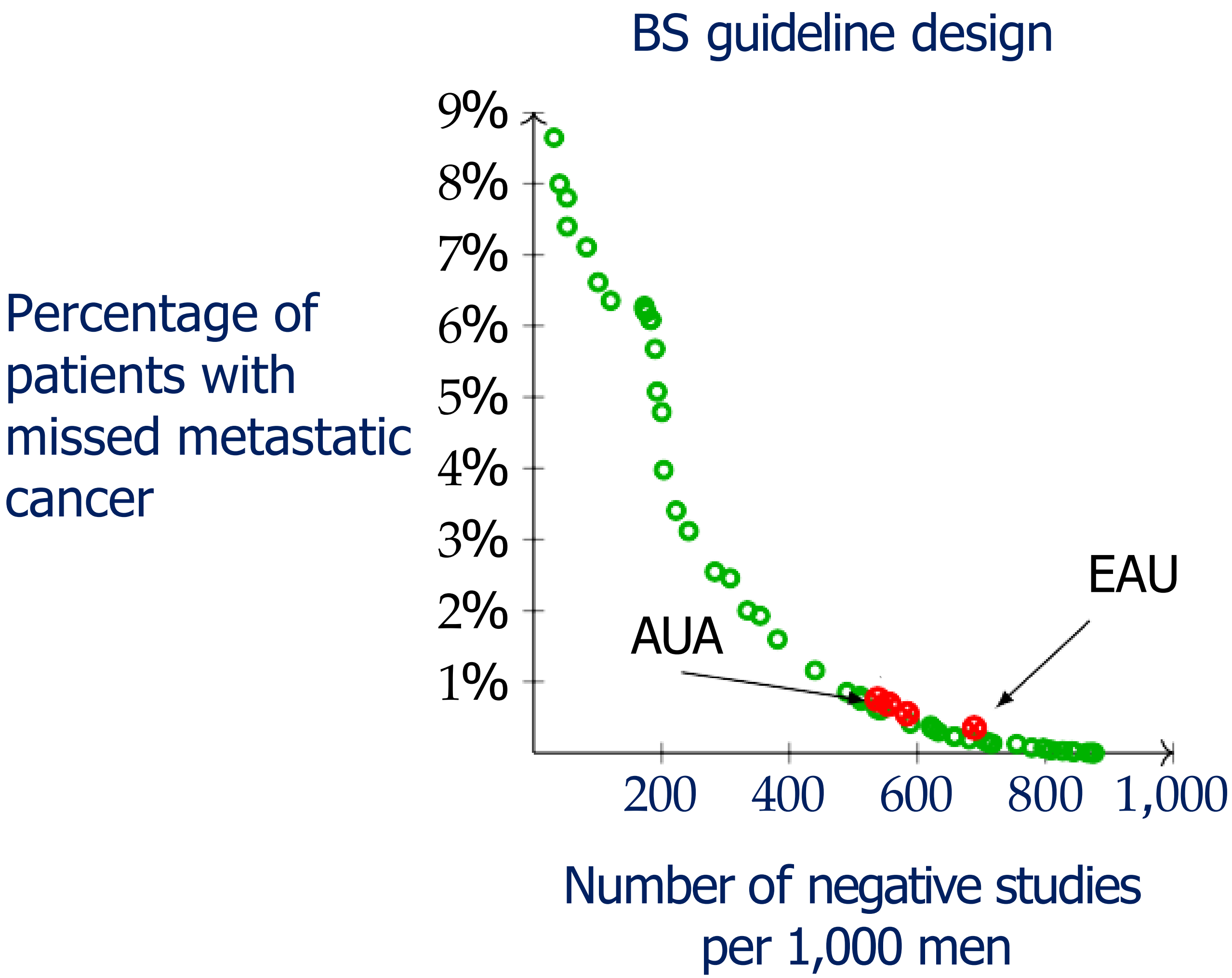
$$f^* = \underset{f \in \mathcal{H}}{\operatorname{argmin}} \frac{1}{l} \sum_{i=1}^{l} \boxed{\delta \mathbb{1}_{\{y_i=1\}} \log\left(1 + e^{-f(\mathbf{x}_i)}\right)} + (1-\delta)\mathbb{1}_{\{y_i=-1\}} \log\left(1 + e^{f(\mathbf{x}_i)}\right)$$

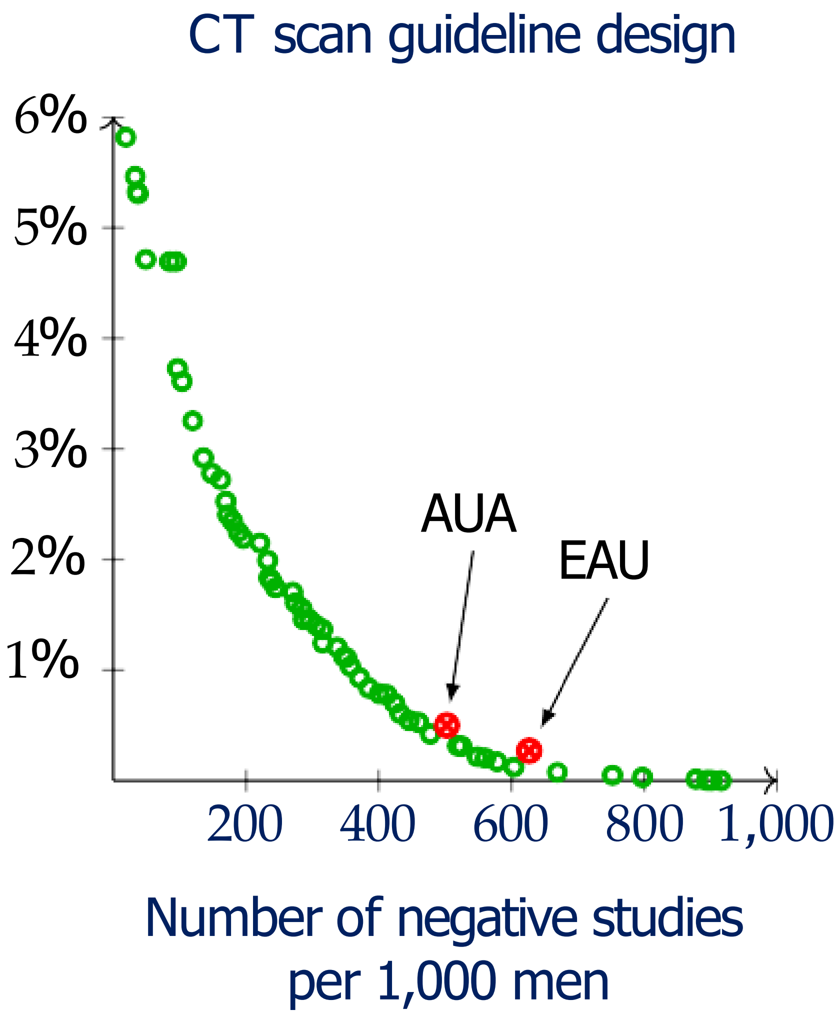
$$+ \boxed{\gamma_{\mathcal{H}} \|f\|_{\mathcal{H}}^2} + \boxed{\gamma_{\mathcal{M}} \mathbf{f}^{\mathbf{T}} \mathbf{L} \mathbf{f}}$$

Avoid overfitting

Extract information from unimaged patients

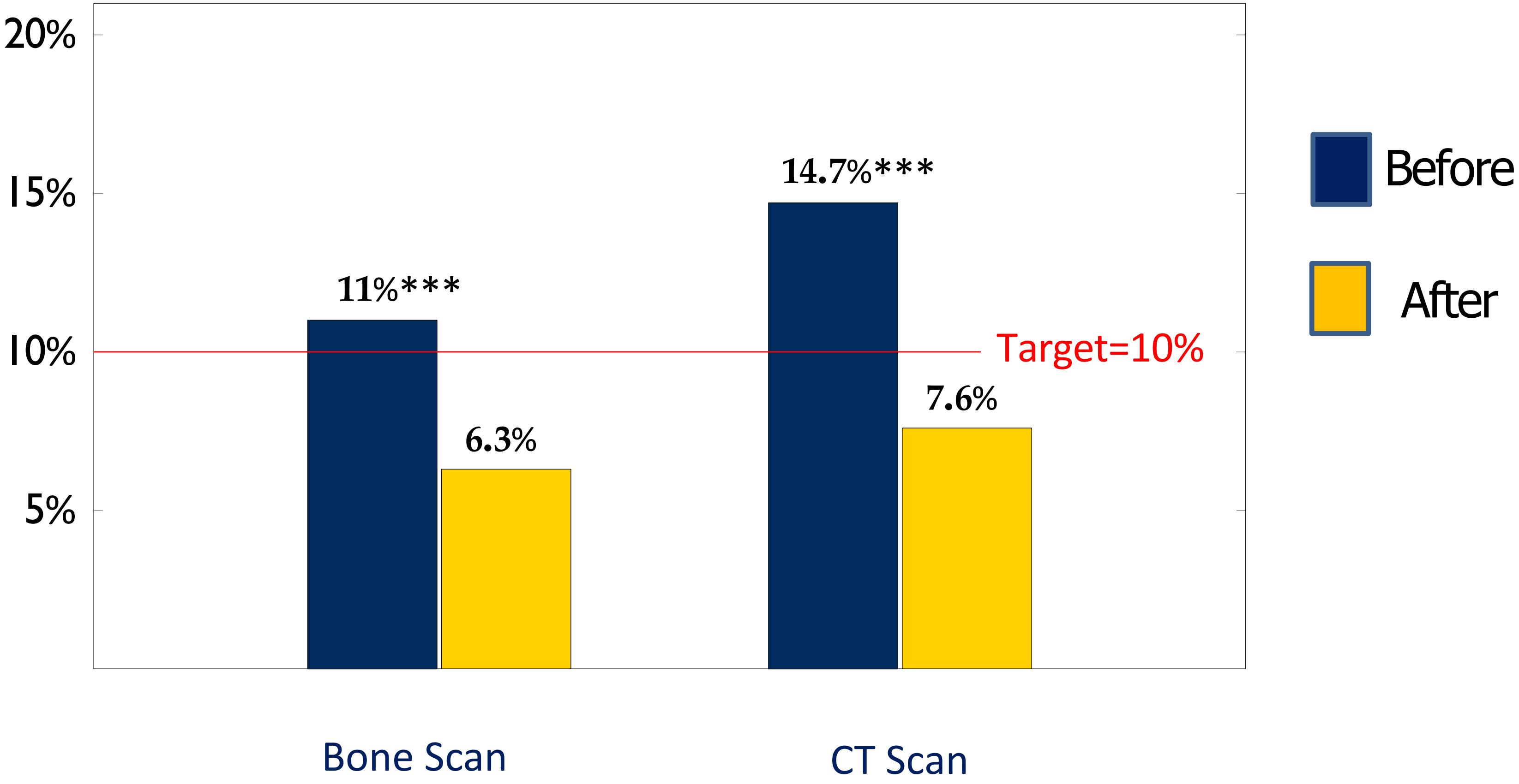
Optimized imaging guideline performance for varying δ





Michigan state-wide decrease in imaging

Imaging rates for patients not recommended by the guideline



Merdan, S., Barnett, C., Miller, D.C., Montie, J.E., Denton, B.T. "Data Analytics for Optimal Detection of Metastatic Prostate Cancer," *Operations Research,* 69 (3), 774-794, *2021*

3. Treatment

Setting: *Treatment of Type 2 diabetes*

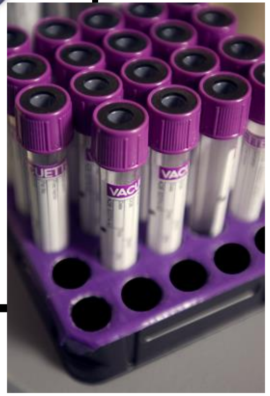
OR Challenge: *ambiguity in risk estimates; stochastic integer programming*

Markov decision process sequence of steps

Markov Process
$P(a_t)$

State
s_t

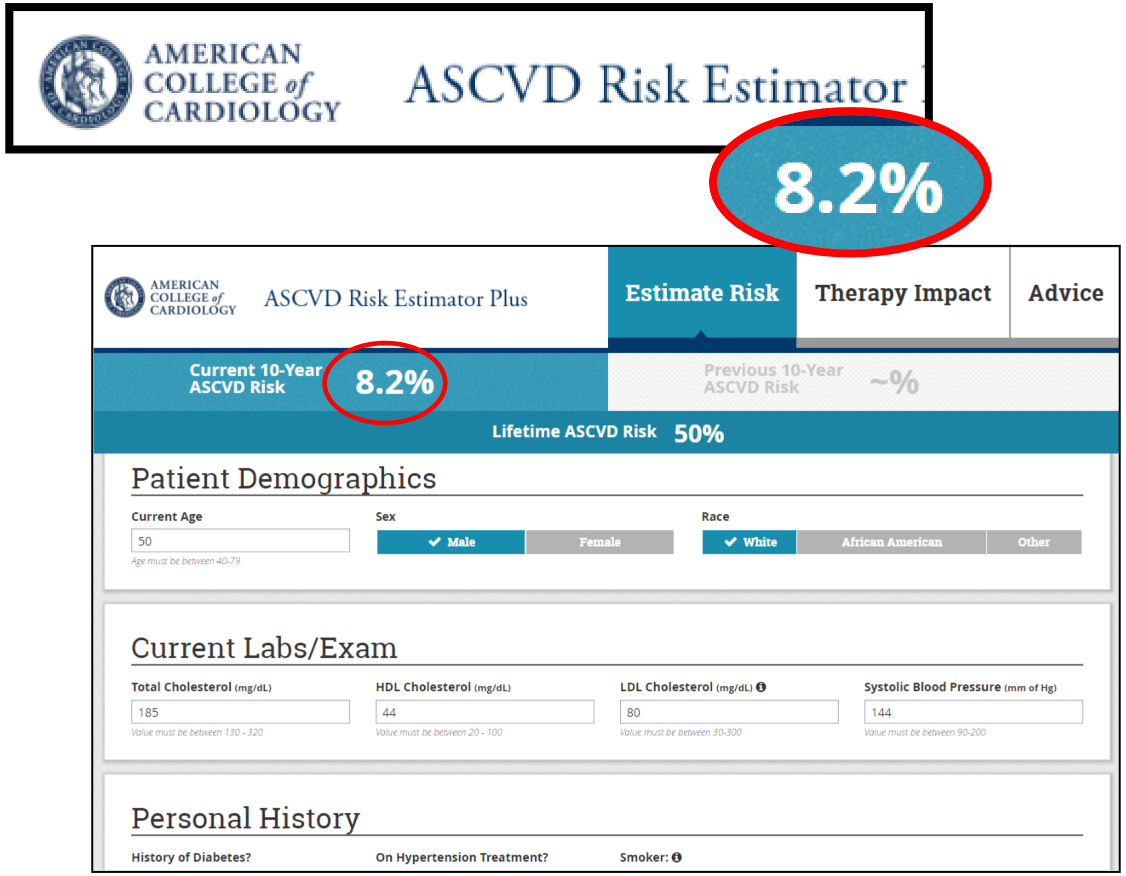
Statins
Fibrates
ACE Inhibitors
ARBs
Beta Blockers
Thiazide

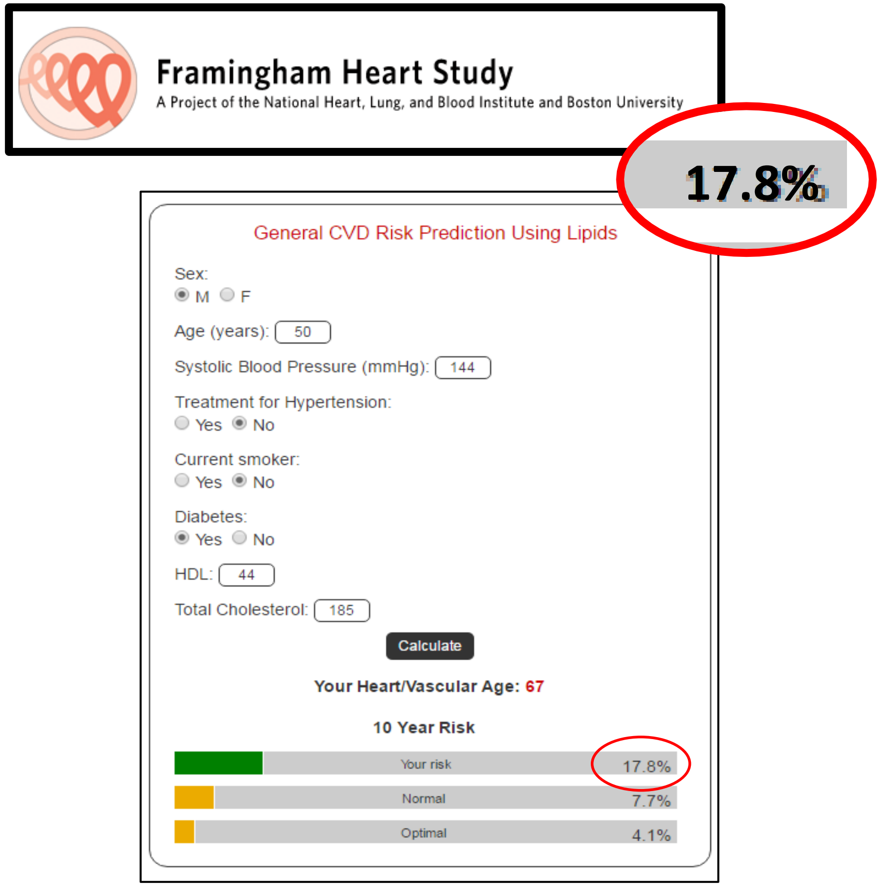


Action
a_t

Decision-Maker's Policy
$\pi(s_t)$

Well-established clinical studies give conflicting estimates about CVD risk





Multi-model Markov Decision Process notation

Generalizes a standard Markov decision process

- States, $\mathcal{S} \equiv \{1, \ldots, S\}$
- Decision epochs, $\mathcal{T} \equiv \{1, \ldots, T\}$
- Actions, $\mathcal{A} \equiv \{1, \ldots, A\}$
- Rewards, $R \in \mathbb{R}^{S \times A \times T}$

Finite set of models, $\mathcal{M} = \{1, \ldots, |\mathcal{M}|\}$

- Model m: An MDP $(\mathcal{S}, \mathcal{A}, \mathcal{T}, R, P^m)$
- Transition probabilities P^m are model-specific
- Model weights: $\lambda_1, \lambda_2, \ldots, \lambda_{|\mathcal{M}|}$

Steimle, L. N., Kaufman, D.L., and Denton B.T. "Multi-model Markov Decision Processes." *IISE Transactions, 2021.*

The **weighted value problem** seeks a single policy (π) that performs across all MDPs

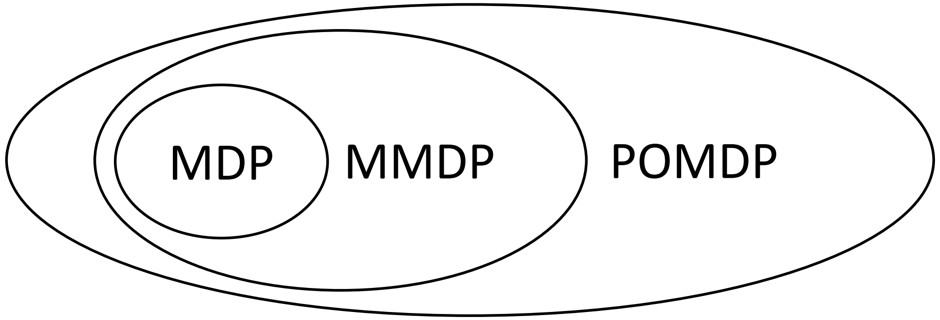
Performance of policy π in model m:

$$v^m(\pi) = \mathbb{E}^{\pi, P^m}\left[\sum_{t=1}^{T} r_t(s_t, a_t) + r_{T+1}(s_{T+1})\right]$$

Weighted value problem:

$$W^* = \max_{\pi \in \Pi} \sum_{m \in \mathcal{M}} \lambda_m v^m(\pi)$$

The weighted value problem is hard



The MMDP is a special case of a partially-observable MDP.

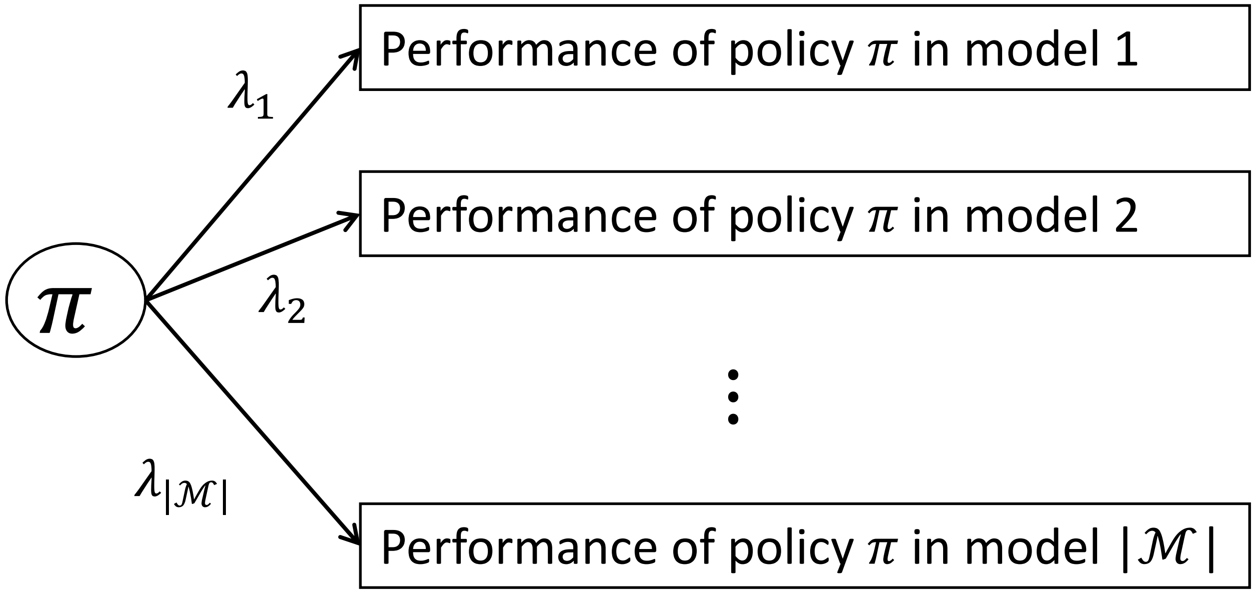
Proposition: The optimal policy may be history-dependent.

Proof by contradiction

Proposition: In general, the Weighted Value Problem is PSPACE-hard.

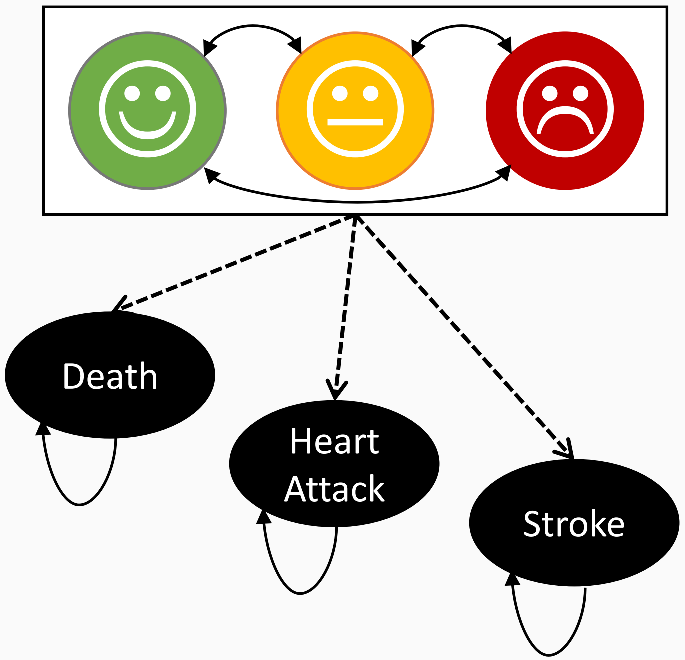
Reduction from *Quantified Satisfiability*

MMDPs can be formulated as two-stage stochastic integer program



Stochastic program	MMDP
Scenarios	Model of MDP
Binary first-stage decision variables	Policy
Continuous second-stage decision variables	MDP model value functions

Example: treatment for cardiovascular disease for patients with type 2 diabetes



Multi-model Markov decision process

- 4,096 states
- 64 actions: combinations of medication
- 40 decision epochs
- 2 models

Case study data

- Longitudinal data from Mayo Clinic
- Framingham, ACC risk calculators
- Disutilities from medical literature

Mason, J. E., Denton, B. T., Shah, N. D., & Smith, S. A. (2014). Optimizing the simultaneous management of blood pressure and cholesterol for type 2 diabetes patients. *European Journal of Operational Research*, *233*(3), 727-738.

A comparison of MMDP policy to MDP policies that ignore model ambiguity

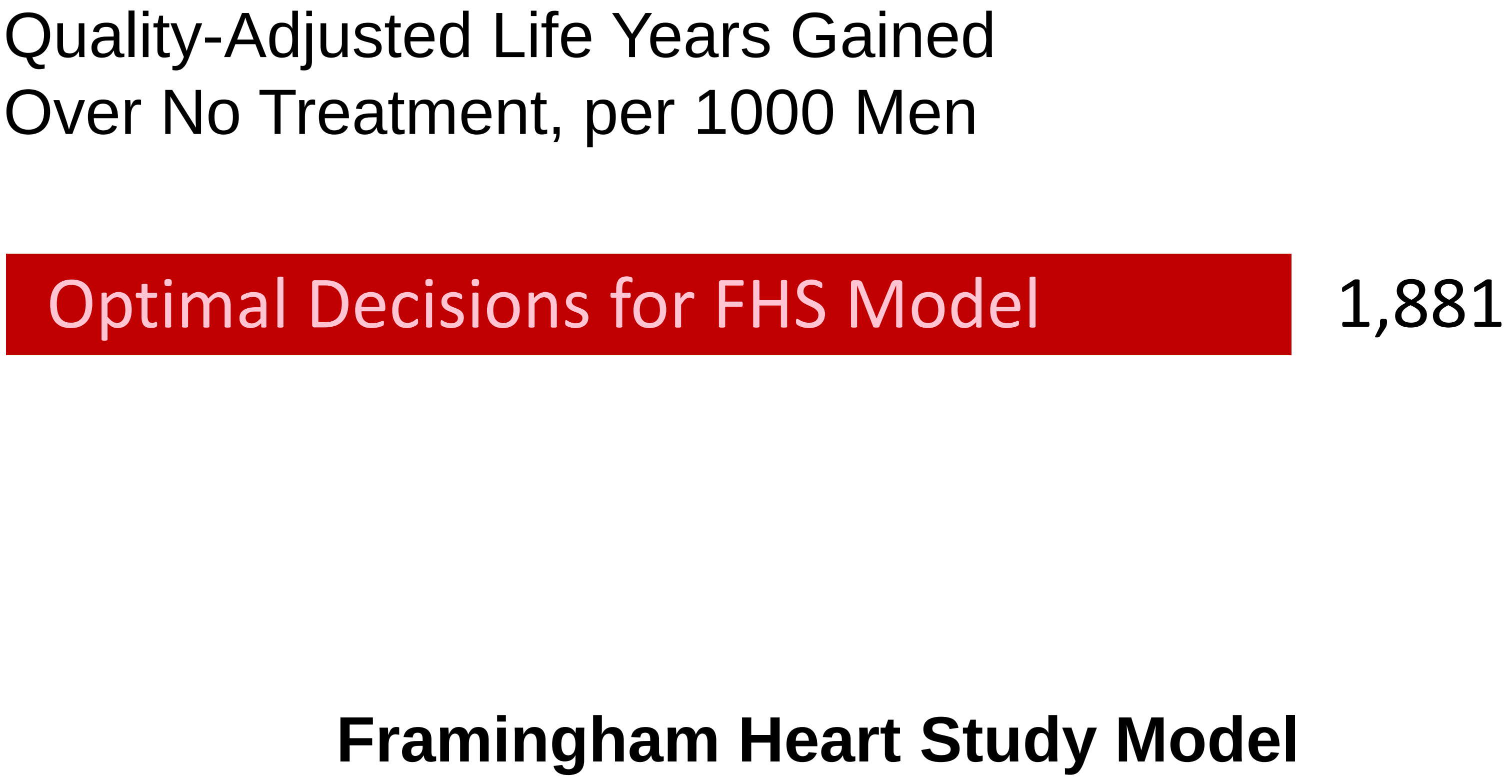
Quality-Adjusted Life Years Gained
Over No Treatment, per 1000 Men

Optimal Decisions for FHS Model

MMDP Decisions

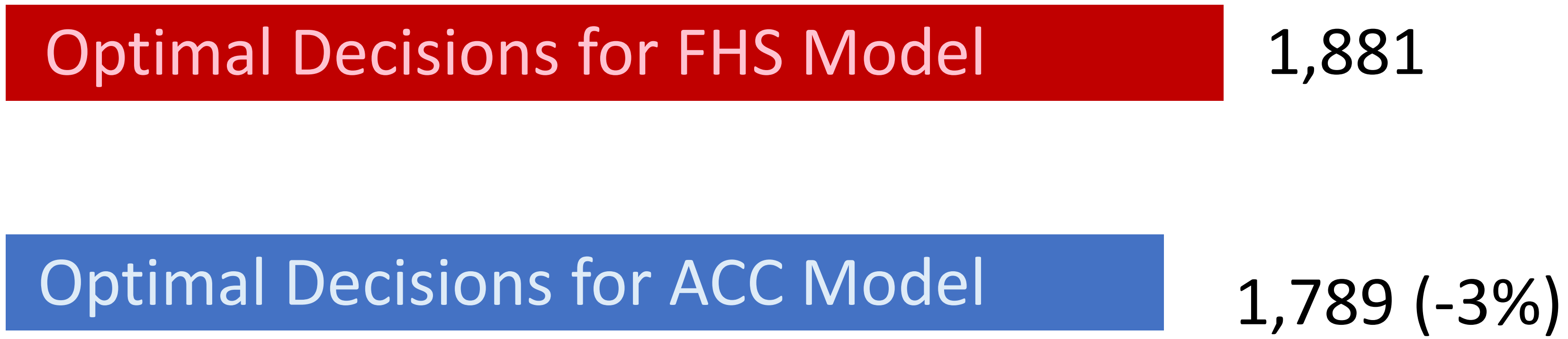
Optimal Decisions for ACC Model

In some cases, ignoring ambiguity has relatively minor implications



In some cases, ignoring ambiguity has relatively minor implications

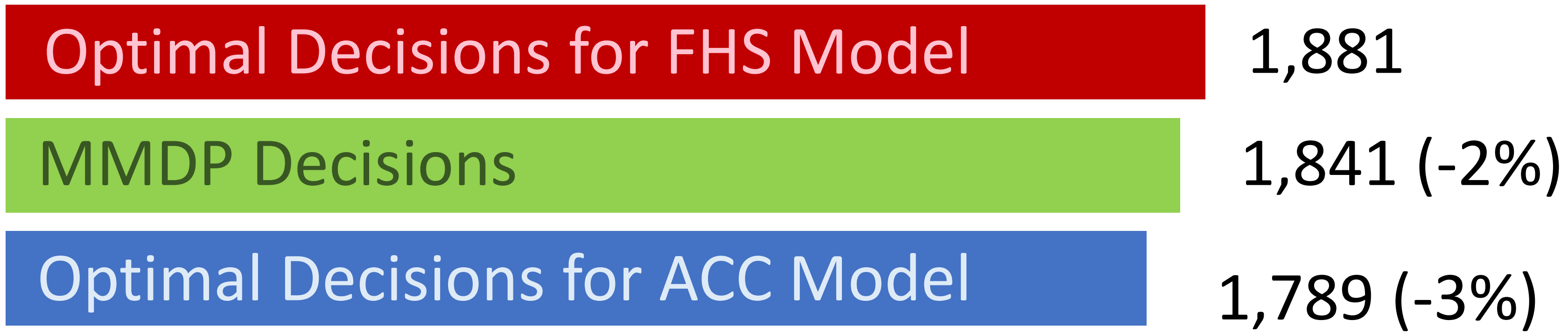
Quality-Adjusted Life Years Gained
Over No Treatment, per 1000 Men



Framingham Heart Study Model

In some cases, ignoring ambiguity has relatively minor implications

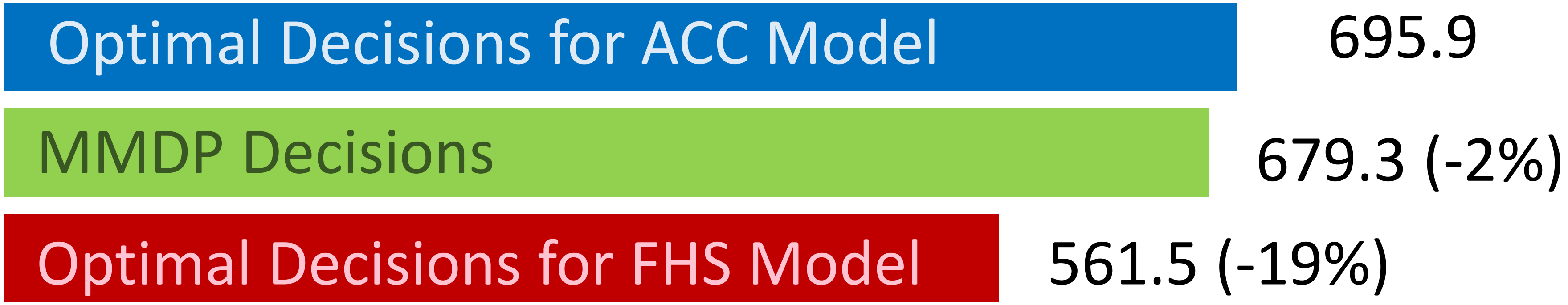
Quality-Adjusted Life Years Gained
Over No Treatment, per 1000 Men



Framingham Heart Study Model

But in other cases, ignoring ambiguity can have major implications

Quality-Adjusted Life Years Gained
Over No Treatment, per 1000 Men



American College of Cardiology Model

Recent articles on MMDPs and extensions

Models for chronic disease to help resolve model ambiguity

1. **Steimle, L.**, Kauffman, D., Denton, B.T., "Multi-model Markov Decision Processes: A New Method for Mitigating Parameter Ambiguity," IISE *Transactions*, 53(10):1124-39, 2021

2. **Steimle, L**., Ahluwalia, V., Kamdar, C., Denton, B.T., "Decomposition Methods for Solving Multi-model Markov Decision Processes," *IISE Transactions,* 53 (12), 1295-1310, 2021

A recent study that addresses ambiguity for active surveillance of prostate cancer:

Li, W., Denton, B.T., "Multi-model Partially Observable Markov Decision Processes," Working Paper, 2023, (available on Optimization Online)

Recap

1. <u>Prevention</u> of cardiovascular events; Markov decision process (MDP) with sparse data

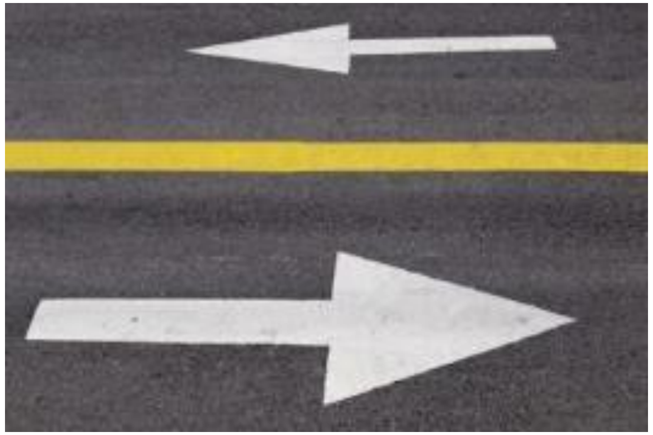
2. <u>Diagnosis</u> of cancer; machine learning, selection bias, and class imbalance

3. <u>Treatment</u> of diabetes; MMDP, stochastic programming, ambiguity in risk models

Parting Thoughts

- Observational data present challenges and opportunities for Medicine and OR

- OR is still underutilized in medicine; there are many unexplored opportunities at the intersection of OR & Analytics



Acknowledgments

Students

Christine Barnett, PhD

Weiyu Li, PhD

Selin Merdan, PhD

Daniel Otero, PhD

Erkin Otles, PhD

Lauren Steimle, PhD

Zheng Zhang, PhD

Collaborators

Rod Hayward, MD

David Kauffman, PhD

Mariel Lavieri, PhD

David C. Miller, MD

James E. Montie, MD

Todd Morgan, MD

Jeremy Sussman, MD

COLLEGE OF ENGINEERING
INDUSTRIAL & OPERATIONS ENGINEERING
UNIVERSITY OF MICHIGAN

Brian Denton
Industrial and Operations Engineering
University of Michigan

btdenton@umich.edu

Find these slides and related articles on my website