COLLEGE OF ENGINEERING
INDUSTRIAL & OPERATIONS ENGINEERING
UNIVERSITY OF MICHIGAN

Using Longitudinal Data to Build Natural History Models

Lessons learned from modeling type 2 diabetes and prostate cancer

INFORMS Healthcare Conference, Rotterdam, 2017

Brian Denton
Department of Industrial and Operations Engineering
University of Michigan

Natural History Models

Natural history models can be used to simulate the natural course of **chronic diseases**.

- Randomized controlled trial (RCT) data is the **gold standard** but:
  - RCTs are expensive and time consuming and may not always be possible
  - Can test a very limited number of treatment options over short time periods

- **Observational data** is abundant but there are challenges to unlocking knowledge due to various sources of bias and confounding

Examples

- **Breast Cancer**: Maillart, L.M., Ivy, J.S., Ransom, S., Diehl, K. Assessing dynamic breast cancer screening policies. *Operations Research*, 56(6):1411–1427, 2008.

- **Liver Disease:** Alagoz, L.M. Maillart, A.J. Schaefer, and M.S. Roberts. Choosing among living-donor and cadaveric livers. *Management Science*, 53(11):1702–1715, 2007

- **Prostate Cancer**: Zhang, J, Denton, B.T., Balasubramanian, H, Shah, N., Inman, B., Optimization of prostate biopsy referral decisions. *M&SOM*, 14(4):529–547, 2012.

- **Lung Cancer** de Koning, Harry J., Rafael Meza, Sylvia K. Plevritis, Kevin Ten Haaf, Vidit N. Munshi, Jihyoun Jeon, Saadet Ayca Erdogan et al. "Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the US Preventive Services Task Force*." Annals of internal medicine* 160, no. 5 (2014): 311-320.

Agenda

- Estimating Markov decision process model parameters
- Examples:
  - Complete Observability: Type 2 diabetes
  - Partial Observability: Prostate cancer
- Conclusions

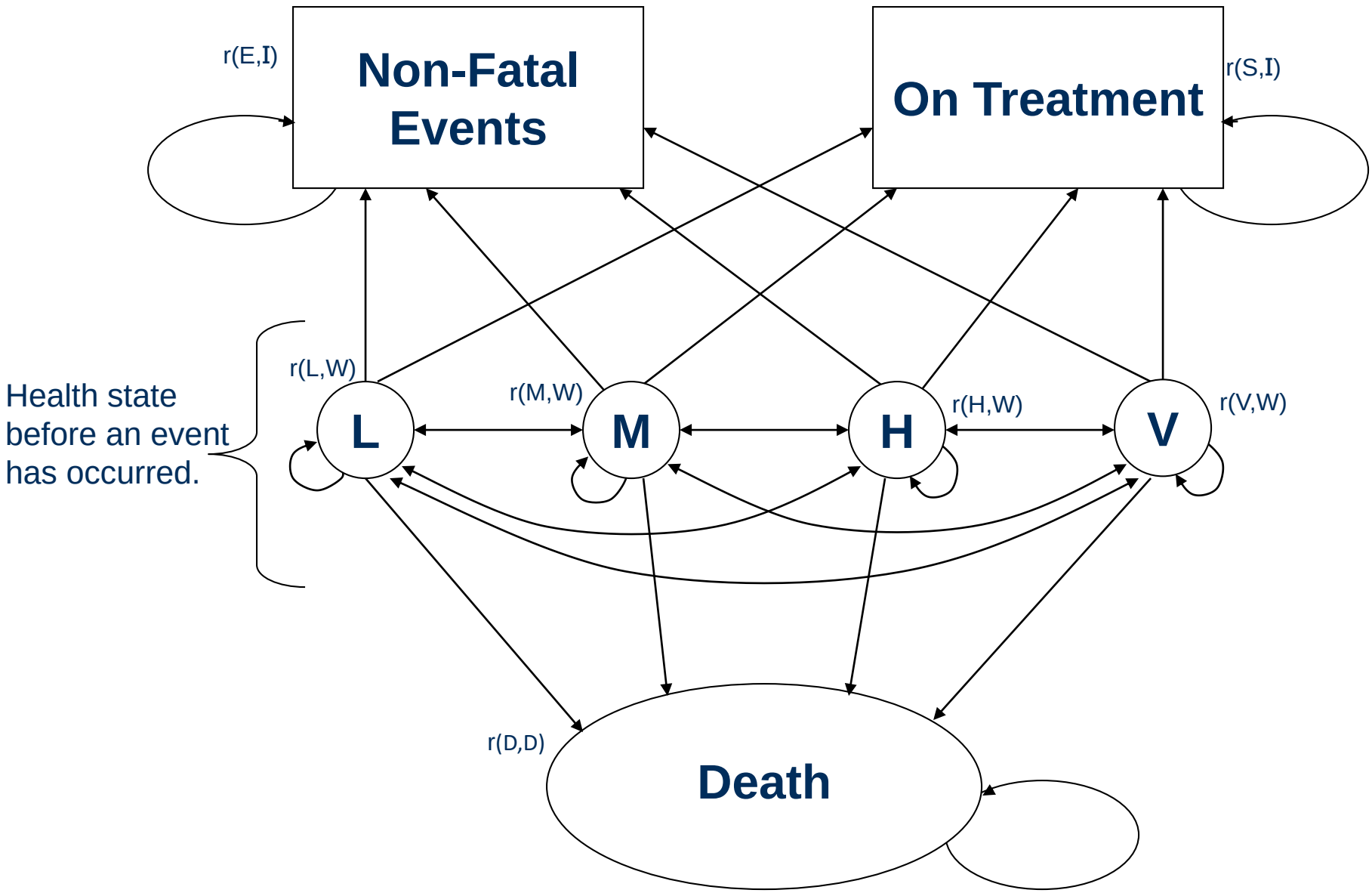
Markov Decision Processes

- Stages: $T = \{1, \dots, N\}$
- States: $S = \{s_1, \dots, s_R, s_{R+1}, \dots, s_{R+M}, D\}$
- Actions: $A(s)$
- Transition probabilities: $p_t(s'|s,a), \ \forall\, t \in \mathcal{T}$
- Rewards: $r_t(s,a)$
- Optimality Equations:

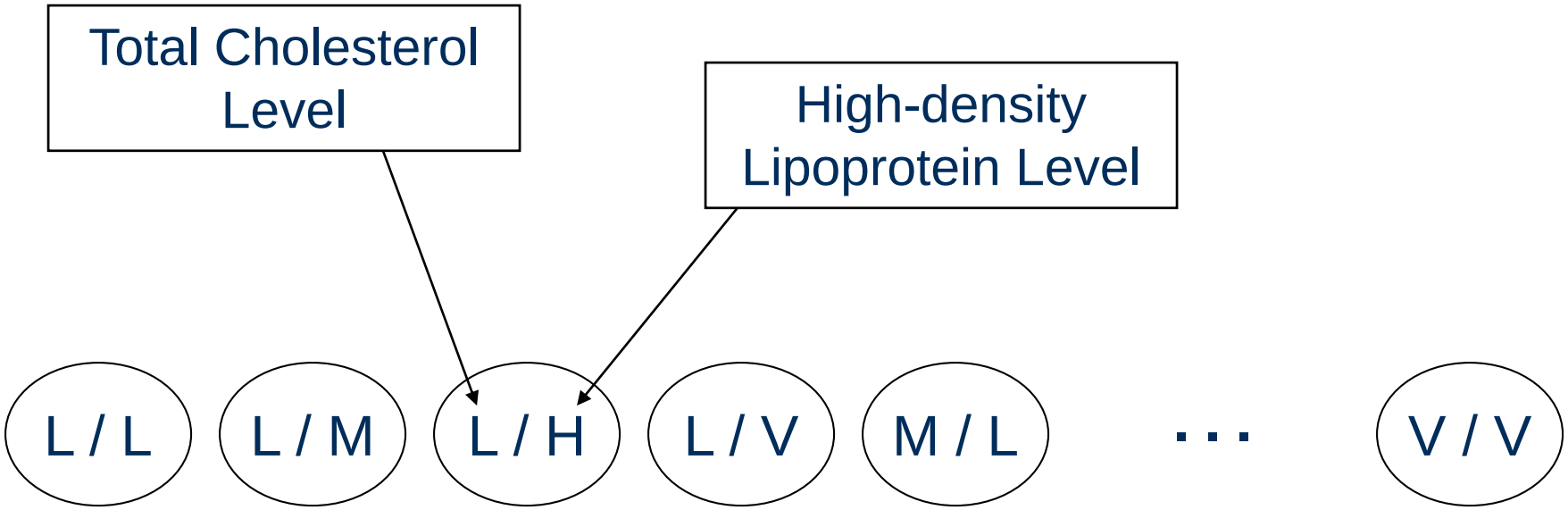
$$v_t(s_t) = \max_{a \in A(s)} \{r_t(s,a) + \lambda \sum_{j \in S} p_t(j|s_t,a) v_{t+1}(j)\}, \forall\, t \in T, s \in S$$

$$v_N(s_N) = R_N(s_N), \forall s_N \in S$$

State Transition Diagram



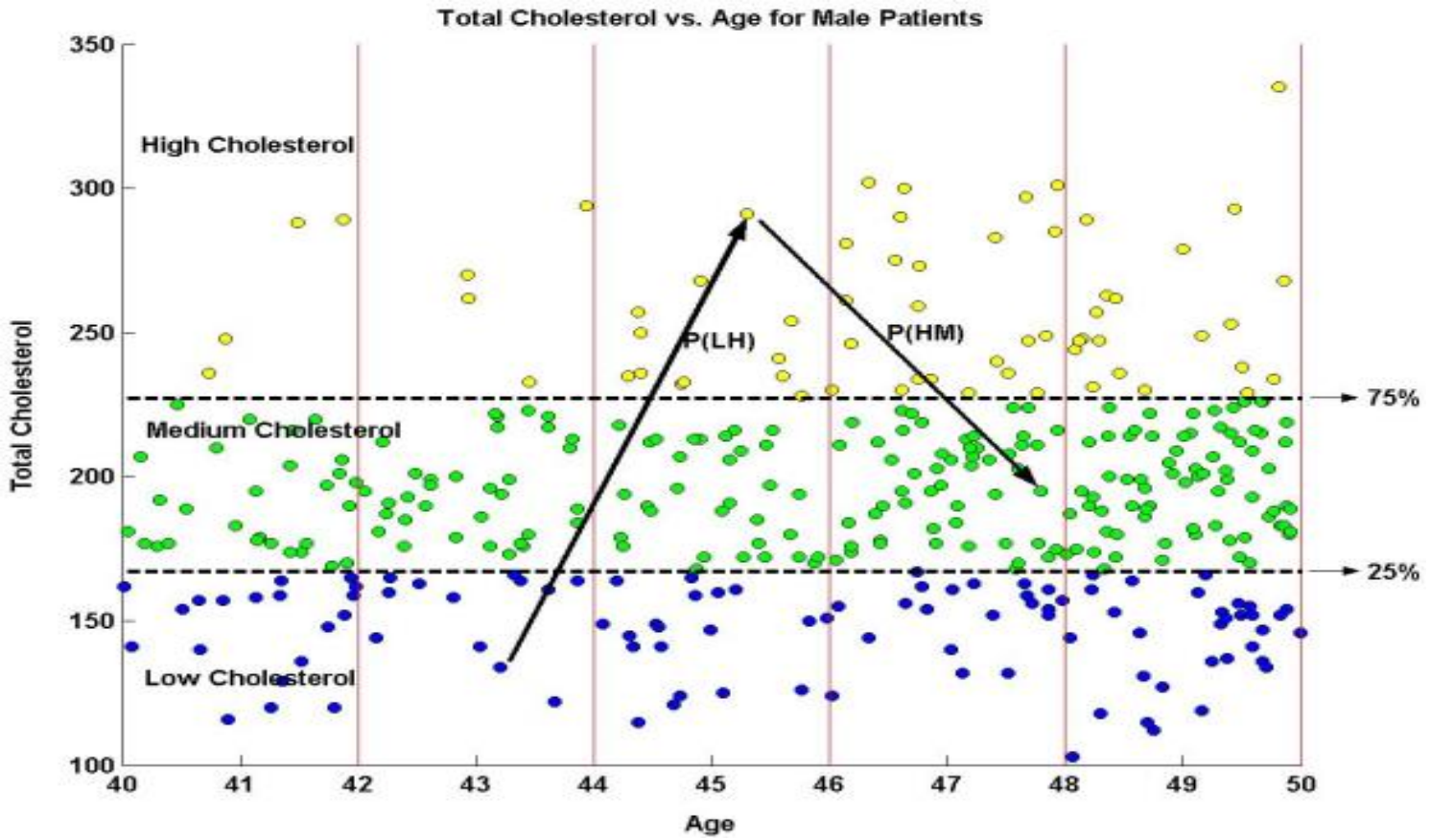
Defining Health States



TC and HDL have four possible levels each, so there are 16 states in total.

	L	M	H	V
TC	<160	160-200	200-240	>240
HDL	< 40	40-50	50-60	>60

Example: Total Cholesterol



Mason, J.E. et al. 2014. "Optimizing the Simultaneous Management of Blood pressure and Cholesterol for Type 2 Diabetes Patients*." European Journal of Operational Research*; 233(3) 727-738.

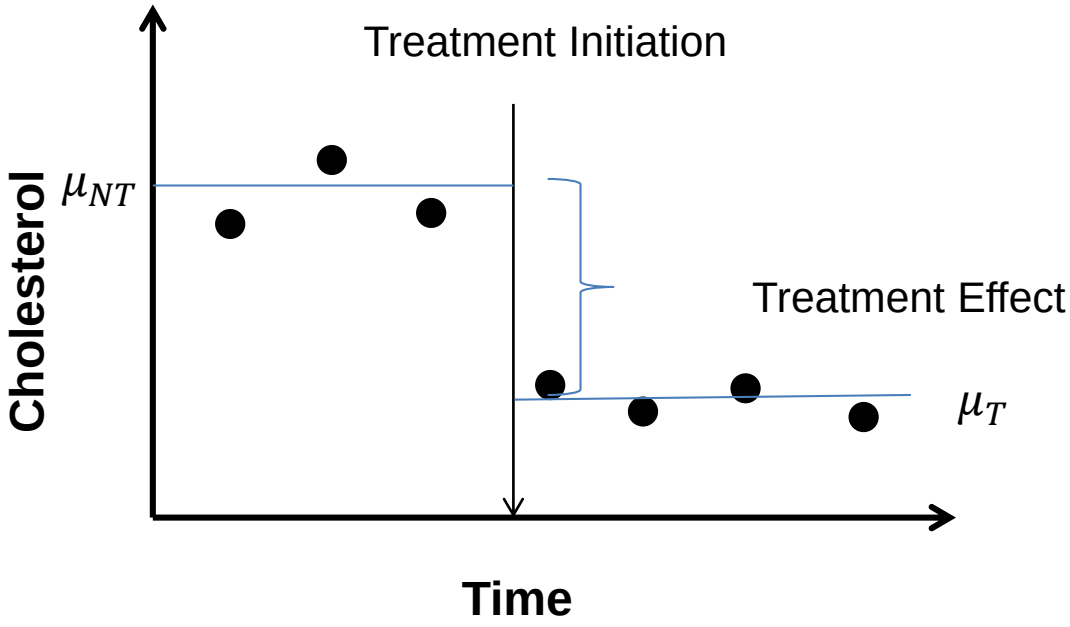
Computing Treatment Effects

- Published randomized trials
  - Adherence bias
  - Patients are carefully followed over short time periods
- Electronic medical record data
  - Selection bias: patients who are treated are typically “sicker”
  - Using a high risk population may over estimate treatment effects

Computing Treatment Effects

Treatment options:

- Statins
- Fibrates
- Ace Inhibitors
- ARBs
- Calcium Channel Blockers
- Thiazide



Addressing Confounding

- Linear Random-Effects Model:
  - assumes each individual patient's longitudinal measurements are correlated
  - individual effect is represented by adding a random intercept in our model

Propensity Score Matching:

- Reduces bias in the context of observational data by accounting for covariates that influence treatment

Estimating Treatment Effects

Step 1: Fit a generalized linear random effects model to estimate the probability of taking medication (propensity score) using other confounding variables:

Example:

Propensity Score (Statins) ~ Age + Gender +BMI+ Other Treatments + Individual effect

Step 2: Estimate the treatment effect by fitting a second regression model:

Example:

Cholesterol ~ Propensity Score (Statins) + Use of Statins

Part of Cholesterol explained by confounding variables

Part of the Cholesterol explained by treatment effect

Uncertainty Set

A combination of laboratory data and pharmacy claims data can be use to estimate transition probabilities between states

$$p(s'|s\ ,a) = \frac{n(s,s',a)}{\sum_{s'} n(s,s',a)}, \forall s', s, a$$

$1-\alpha$ confidence intervals for row s of the TPM:

$$[\hat{p}(s'|s,a) - S(\hat{p}(s'|s,a)L, \quad \hat{p}(s'|s,a) + S(\hat{p}(s'|s,a)L]$$

where

$$S(\hat{p}(s'|s,a))L = \left[\chi^2_{|s|-1,\alpha/2|S|)} \frac{\hat{p}(s'|s,a)\big(1-\hat{p}(s'|s,a)\big)}{N(s)}\right]^{\frac{1}{2}}$$

Gold, Ruth Z. "Tests auxiliary to χ 2 tests in a Markov chain." *The Annals of Mathematical Statistics* 34, no. 1 (1963): 56-74.

Event Probabilities

- Type 2 diabetes complications:
  - Framingham model
  - UKPDS model
  - Archimedes

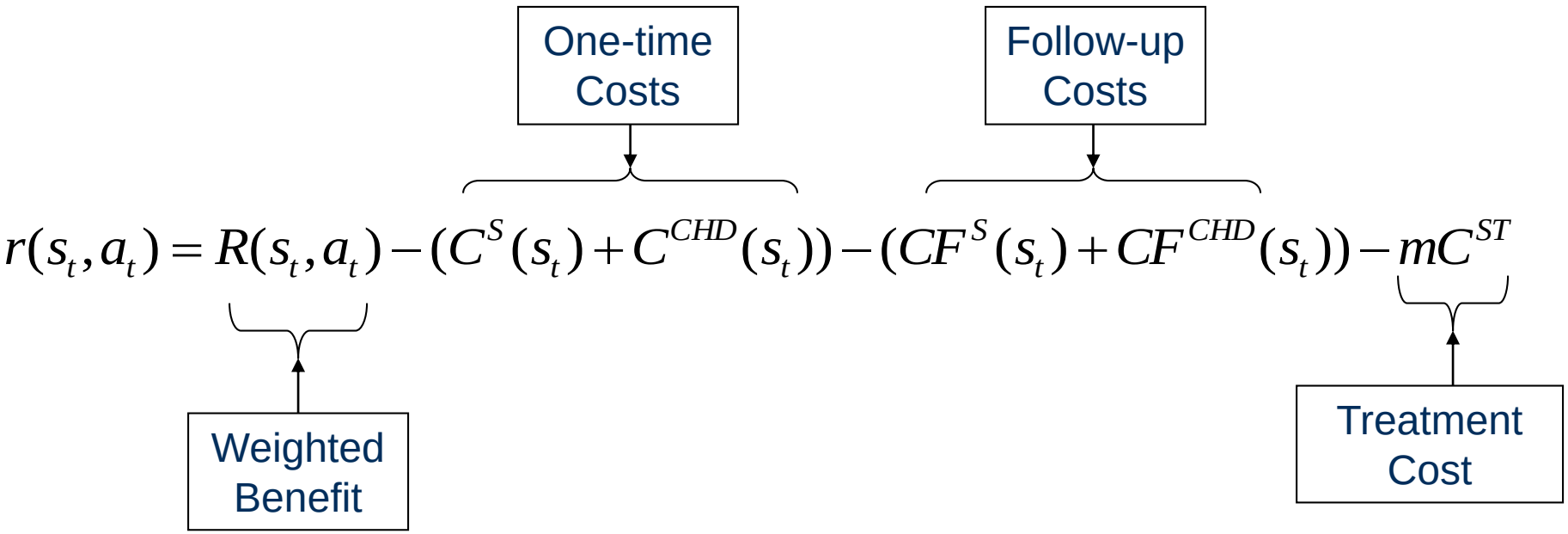
- Other cause mortality:
  - Life tables (e.g. U.S. Centers for Disease Control and Prevention)

Reward Data

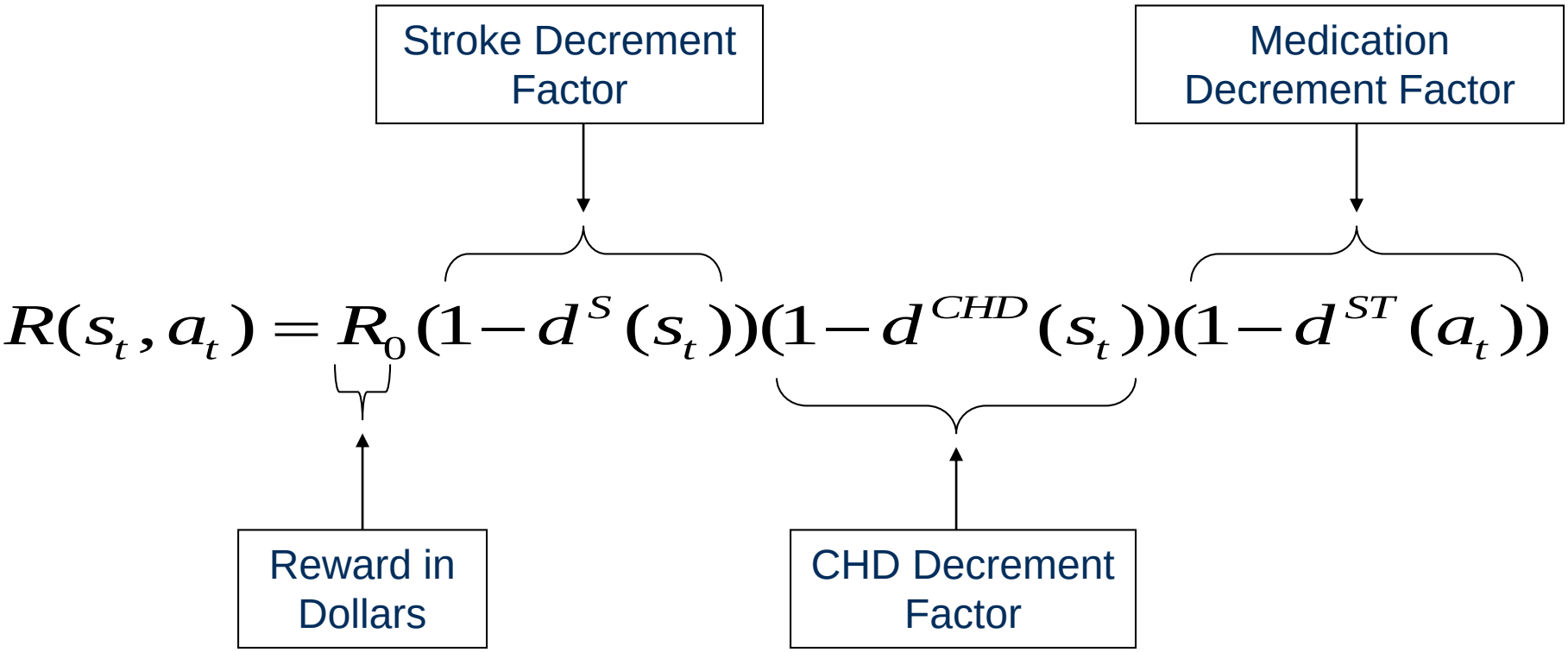
- Patient
  - Maximize expected quality adjusted life years (QALYs)
- Third-party Payer
  - Minimize expected costs of treatment and health services
- Society
  - Maximize a weighted combination of expected patient rewards for QALYs minus costs of treatment and health services

Societal Perspective

- Objective function includes rewards for quality adjusted life years (QALYs) and costs



Weighted Annual Benefit to the Patient



Reward Parameters

Parameter	Description	Value
C^S	Initial hospitalization cost for a stroke.	\$11,161
C^{CHD}	Initial hospitalization cost for a CHD event.	\$16,085
CF^S	Yearly follow-up cost for a stroke.	\$1,664
CF^{CHD}	Yearly follow-up cost for a CHD event.	\$2,576
C^{ST}	Cost of statin treatment.	\$360
R_0	Patient reward for a year of quality life.	\$100,000
d^S	Stroke utility decrement.	0.21
d^{CHD}	CHD utility decrement.	0.07
d^{ST}	Statins utility decrement.	0.03

Sources

- Systematic review of the literature via Pubmed
- Insurance claims data
- Pharmacy Redbook drug costs
- Cost Effectiveness Registry: https://research.tufts-nemc.org/cear4/Home.aspx

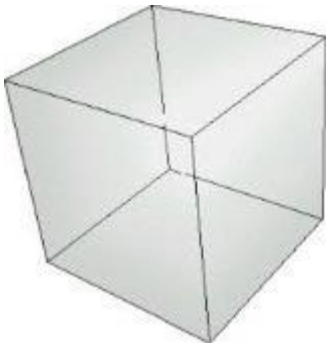
Model Validation

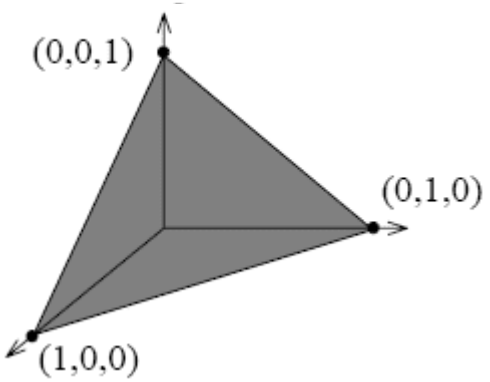
- Expert opinion
- Statistical methods:
  - Cross-validation
  - Boot-strapping
- Comparison to independent estimates of long-term outcomes (e.g. lifespan, time to event, prevalence)

TPM Sampling Method

Basic idea:

- Random-direction algorithm[1] for sampling random vectors over convex region
- Sample each row of the TPM independently from intersection of uncertainty set, U, and standard simplex, Δ





[1]: Smith, R.L, Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions, Operations Research, 32(6) p 1296-1308, 1984

Algorithm

$Choose\ initial\ point\ \boldsymbol{X_0}\ in\ the\ uncertainty\ set, \mathcal{U}$

$For\ j = 1, \dots, W + M\ samples$

$Sample\ \boldsymbol{d}\ such\ that\ \boldsymbol{X_{j-1}} + \lambda \boldsymbol{d} \in \Delta$

$Find\ \underline{\lambda}\ and\ \bar{\lambda}\ such\ that\ \boldsymbol{X_{j-1}} + \lambda \boldsymbol{d} \in \mathcal{U}$

$Sample\ \lambda\ uniformly\ in\ interval\ [\underline{\lambda}, \bar{\lambda}]$

$While(\boldsymbol{X_{j-1}} + \lambda \boldsymbol{d} \notin \mathcal{U})$

$if\ \lambda \geq 0\ then\ \bar{\lambda} \leftarrow \lambda$

$else\ \underline{\lambda} \leftarrow \lambda$

$Sample\ \lambda\ uniformly\ in\ interval\ [\underline{\lambda}, \bar{\lambda}]$

$End\ While$

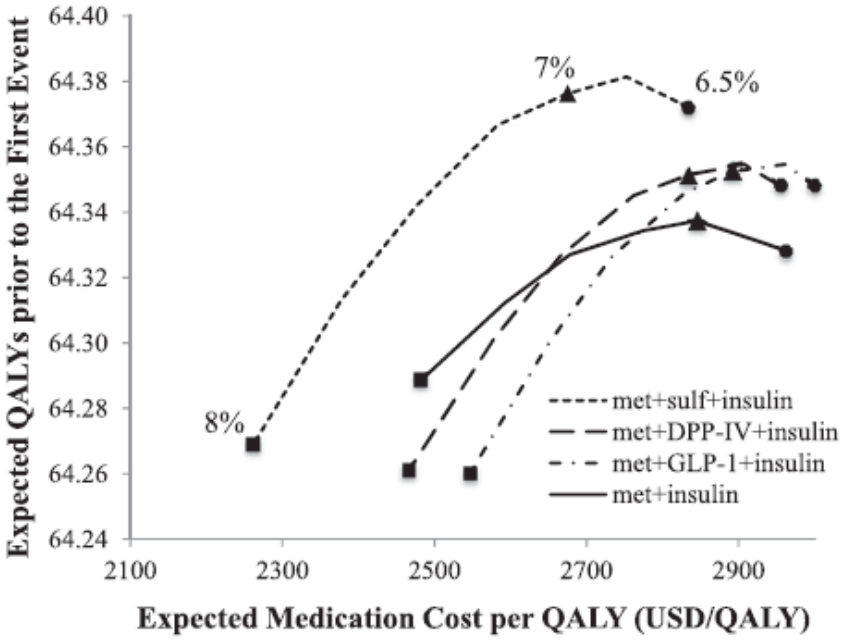
$\boldsymbol{X_j} \leftarrow \boldsymbol{X_{j-1}} + \lambda \boldsymbol{d};\ j \leftarrow j + 1$

$End\ For$

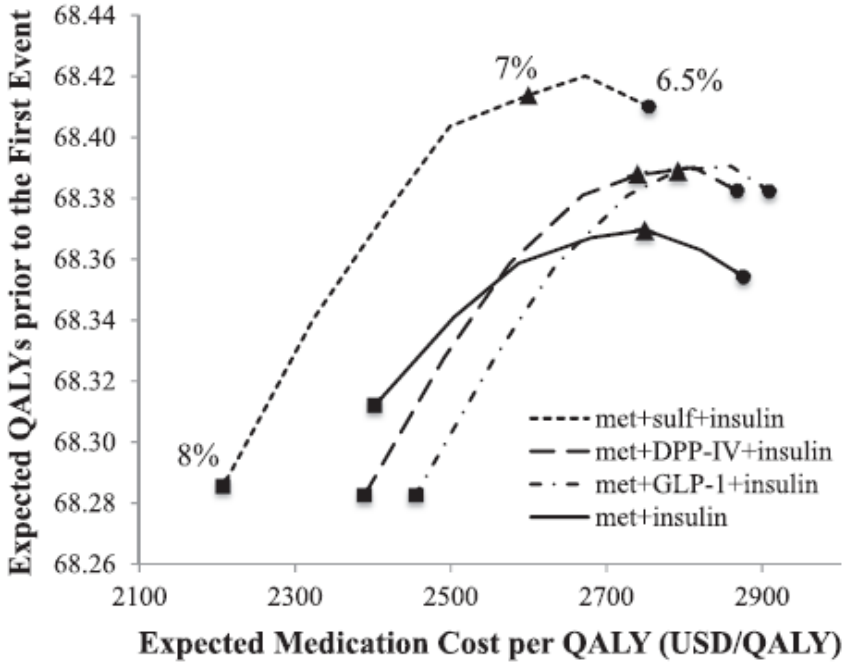
Zhang, Y., Mason, J.E., Wu, H., Wilson, J., Denton, B.T., "Conducting Probabilistic Sensitivity Analysis for Markov Decision Processes," Working Paper (under review), 2016.

Are Newer Drugs Better?

Men

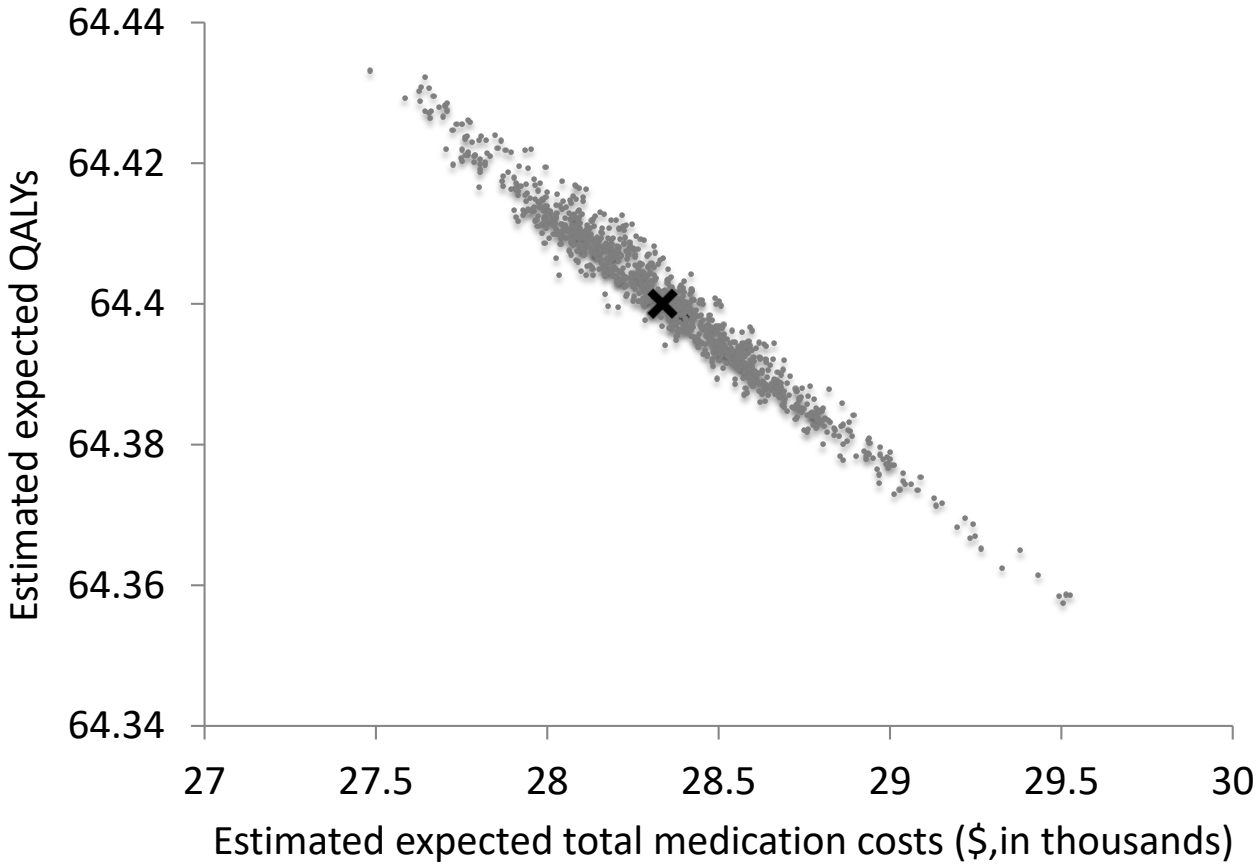


Women



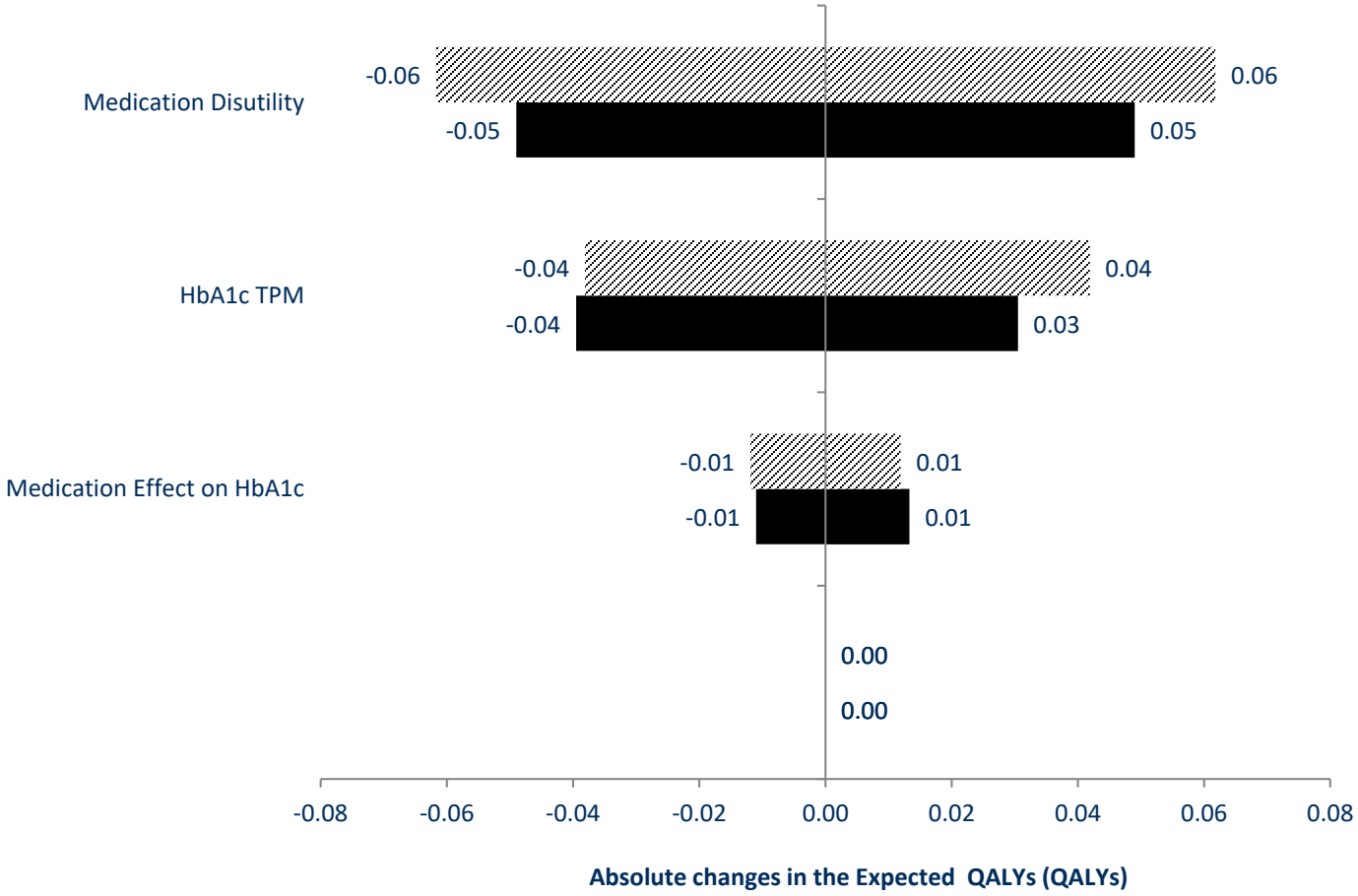
Zhang, Y., McCoy, R.G., Mason, J., Smith, S.A., Shah, N., Denton, B.T., "Second-Line Agents for Glycemic Control for Type 2 Diabetes: Are Newer Agents Better?," *Diabetes Care* 37:5 1338-1345, 2014.

Example: Sensitivity Analysis for Glycemic Control



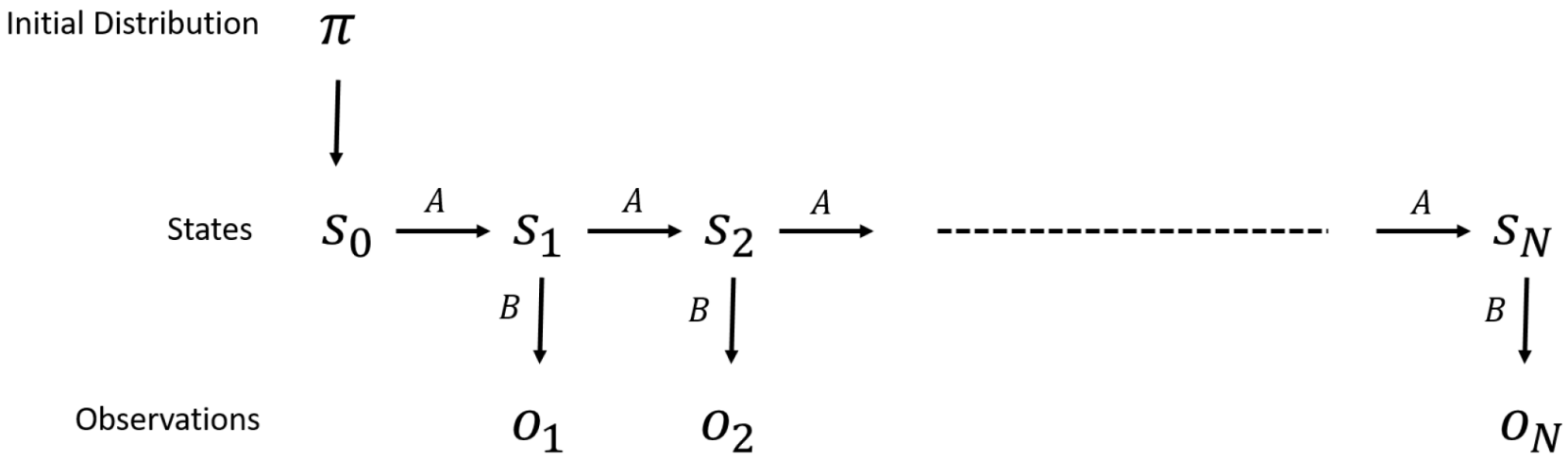
Zhang, Y., McCoy, R.G., Mason, J., Smith, S.A., Shah, N., Denton, B.T., "Comparative Analysis of Treatment Regimens for Glycemic Control in Diabetes Patients," *Diabetes Care* 37:5 1338-1345, 2014.

Example: Sensitivity Analysis for Glycemic Control



Partially Observable Problems

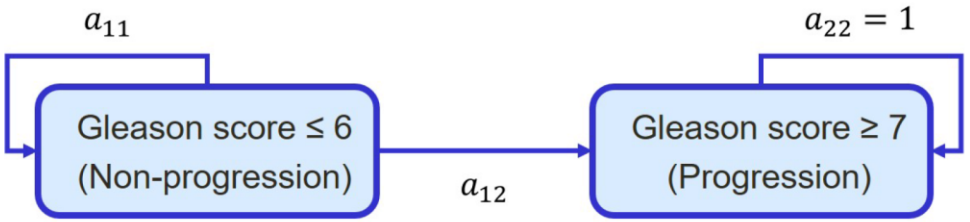
For some diseases the true disease status cannot be measured without invasive diagnostic tests



Example: Prostate Cancer

Active surveillance of prostate cancer involved regular biopsies to monitor disease severity

Biopsies are imperfect and sometimes “miss” the cancer



$$A = \begin{array}{c} \\ \text{GS}\leq 6 \\ \text{GS}\geq 7 \end{array} \begin{array}{c} \begin{array}{cc} \text{GS}\leq 6 & \text{GS}\geq 7 \end{array} \\ \begin{bmatrix} a_{11} & a_{12} \\ 0 & 1 \end{bmatrix} \end{array},$$

$$B = \begin{array}{c} \\ \text{GS}\leq 6 \\ \text{GS}\geq 7 \end{array} \begin{array}{c} \begin{array}{cc} - & + \end{array} \\ \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \end{array}, \quad \pi = \begin{array}{c} \begin{array}{cc} \text{GS}\leq 6 & \text{GS}\geq 7 \end{array} \\ \begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} \end{array},$$

Baum-Welch Algorithm

- Each patient in the dataset has an observation sequence:

$$O^{(1)} = [-, -, +]$$
$$O^{(2)} = [-, +]$$
$$O^{(3)} = [-, -, -, -]$$

- Baum-Welch algorithm finds parameter estimates, $\lambda = (A, B, \pi)$, that locally maximize:

$$P(O|\lambda) = \prod_{k=1}^{N} P(O^{(k)}|\lambda)$$

Rabiner, Lawrence R. "A tutorial on hidden Markov models and selected applications in speech recognition." *Proceedings of the IEEE* 77, no. 2 (1989): 257-286.

Baum-Welch Algorithm

- Each patient in the dataset has an observation sequence:

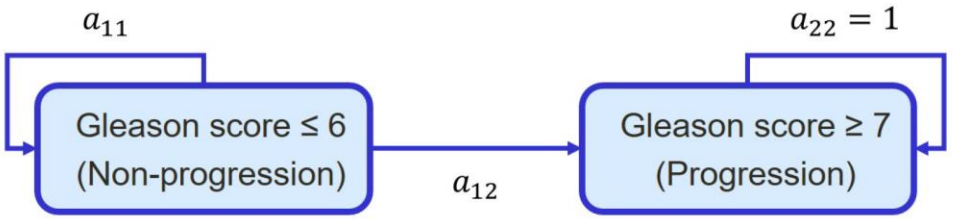
$$O^{(1)} = [-, -, +]$$
$$O^{(2)} = [-, +]$$
$$O^{(3)} = [-, -, -, -]$$

- Baum-Welch algorithm finds parameter estimates, $\lambda = (A, B, \pi)$, that locally maximize:

$$P(O|\lambda) = \prod_{k=1}^{N} P(O^{(k)}|\lambda)$$

Results

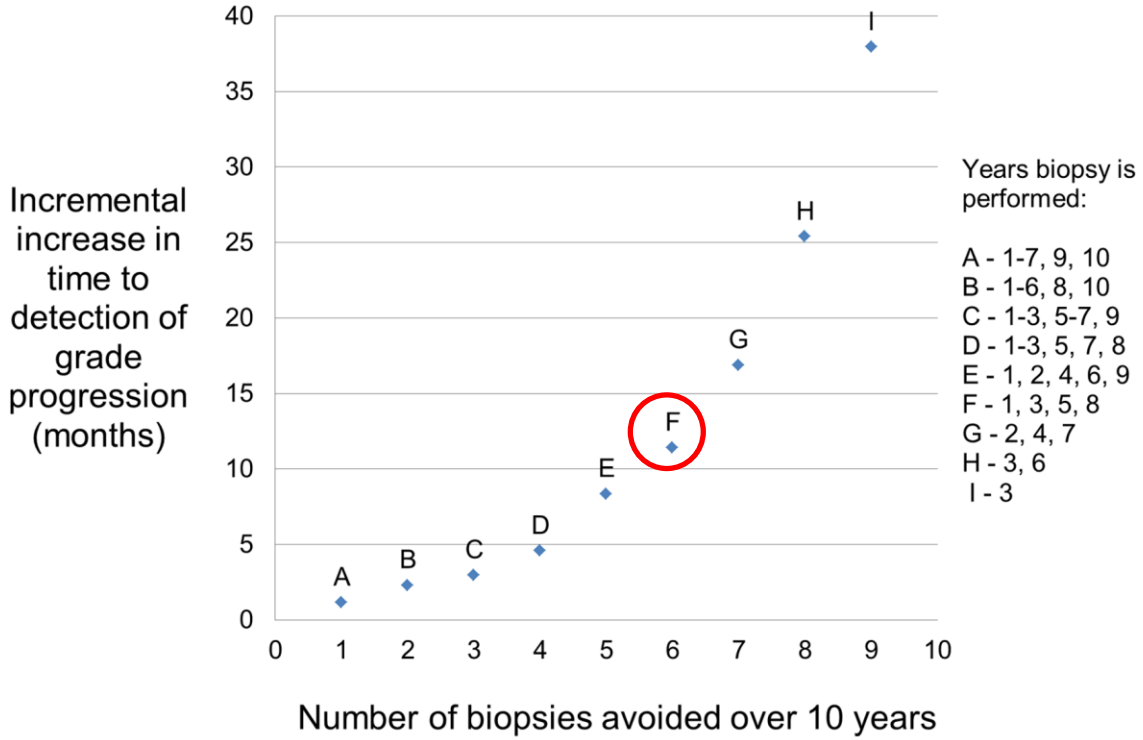
Results from a 10-year study of patients on active surveillance at Johns Hopkins University



- Percentage of patients with GS≤ 6 at diagnosis $= 90.22\%$
- Annual progression probability $= 3.97\%$
- Sensitivity biopsy $= 61.03\%$, Specificity biopsy $= 98.62\%$

Results

Eliminating 6 biopsies would only increase the average time to detection by 11.4 months relative to an annual biopsy schedule



Takeaway Messages

- Markov decision processes (MDPs) are increasingly used to study medical decisions
- Natural history models are the fundamental foundation for MDPs
- Little is know about the best ways to estimate MDP model parameters and mitigate bias from observational data

Acknowledgements

Christine Barnett, Research Triangle Institute
Jennifer Mason, University of Virginia
Lauren Steimle, University of Michigan
Jim Wilson, NC State University
Yuanhui Zhang, CDC
Nilay Shah, Mayo Clinic
Steven Smith, Mayo Clinic

This work was supported by the National Science Foundation CMMI 1462060. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Thanks!

Slides posted on my website:
http://umich.edu/~btdenton/presentations/

Brian Denton
Industrial and Operations Engineering
University of Michigan
btdenton@umich.edu

Recent Related Work

Zhang, J., Denton, B.T., Balasubramanian, H., Inman, B., Shah, N., “Optimization of Prostate Biopsy Referral Decisions,” *Manufacturing & Service Operations Management* , 14(4), 529-547, 2012.

Mason, J., Denton, B.T., Shah, N., Smith, S., “Optimizing the Simultaneous Management of Cholesterol and Blood Pressure Treatment Guidelines for Patients with Type 2 Diabetes,” European Journal of Operational Research, 233, 727-738, 2013.

Zhang, Y., McCoy, R.G., Mason, J., Smith, S.A., Shah, N., Denton, B.T., “Second-line agents for glycemic control for type 2 diabetes: are newer agents better?,” Diabetes Care, 37:5 1338-1345, 2014.

Zhang, Y., , Wu, H., Denton, B.T., Wilson, J.R., Lobo, J.M., “Conducting Probabilistic Sensitivity Analysis for Markov Decision Processes,” Working paper

Zhang, Y., Denton, B.T., “Robust Markov Decision Processes for Medical Treatment Decisions,” Working Paper, 2015 (available at Optimization Online: http://www.optimization-online.org/DB_HTML/2015/10/5134.html)

Steimle, L., Denton, B.T., “Screening and Treatment of Chronic Diseases,” Markov Decision Processes in Practice, Eds: Richard Boucherie and Nico M. van Dijk, Springer, New York, 2017