Robust Optimal Control for Medical Treatment Decisions–An Application to Type 2 Diabetes

Brian Denton

(Work with Yuanhui Zhang, Research Triangle Institute, and Lauren Steimle, University of Michigan)

Department of Industrial and Operations Engineering
University of Michigan

February 1, 2018

Type 2 Diabetes

Prevalence

- $\sim$29 million people had diabetes in U.S. [1]
- Approximately 9.3% of the population
- 90-95% have type 2 diabetes

[1] Data from the 2012 National Diabetes Fact Sheet

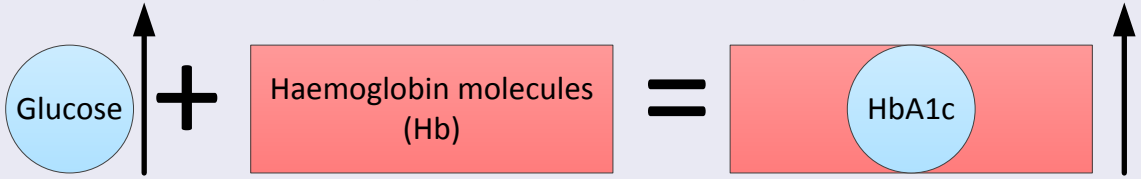
Type 2 Diabetes

Prevalence

- ∼29 million people had diabetes in U.S. [1]
- Approximately 9.3% of the population
- 90-95% have type 2 diabetes

Glycemic Control

- The focus of managing type 2 diabetes
- Clinically measured by glycated haemoglobin (HbA1c)



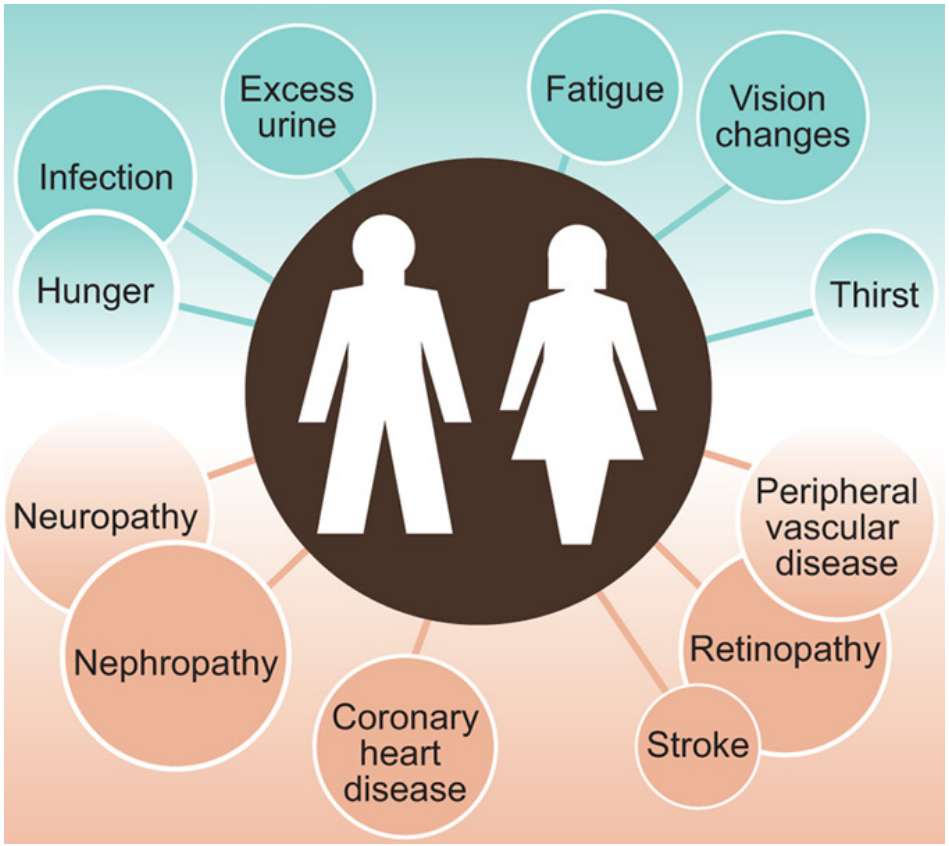
$$\text{HbA1c level } (\%) = \frac{\text{HbA1c concentration}}{\text{Total Hb concentration}}$$

[1] Data from the 2012 National Diabetes Fact Sheet

Importance of Glycemic Control

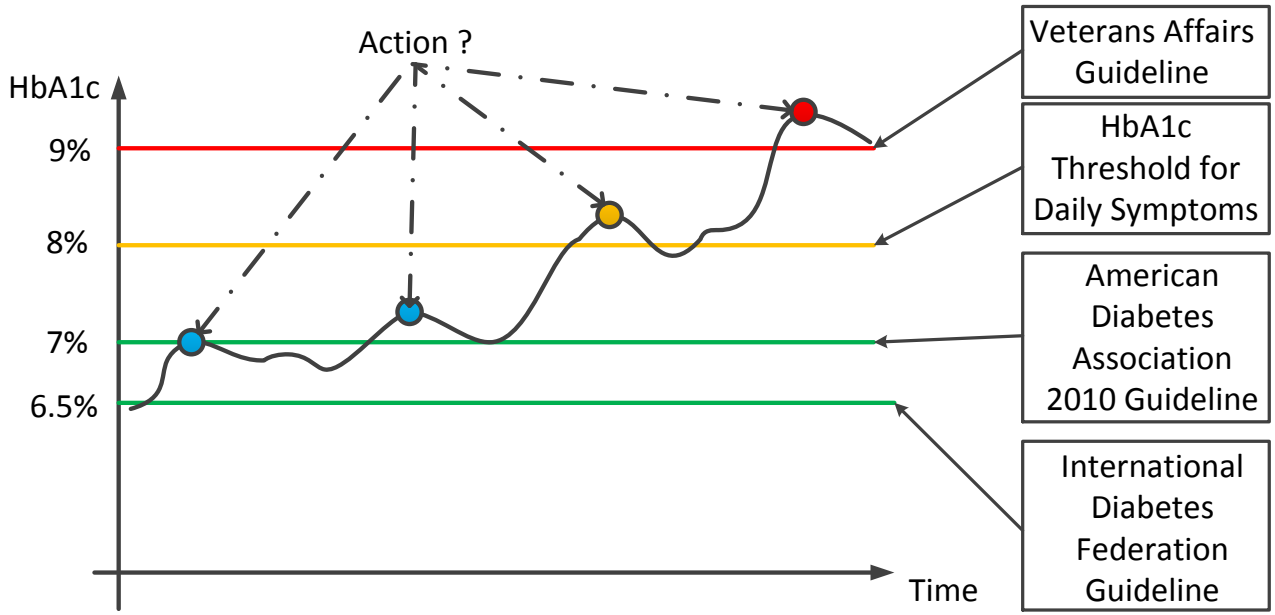
High blood sugar causes damage throughout the body

- Daily symptoms
  - Thirst
  - Fatigue
  - Blurred vision
- Microvascular events:
  - Kidney Disease
  - Blindness
  - Amputation
- Macrovascular events:
  - Heart Attack
  - Stroke



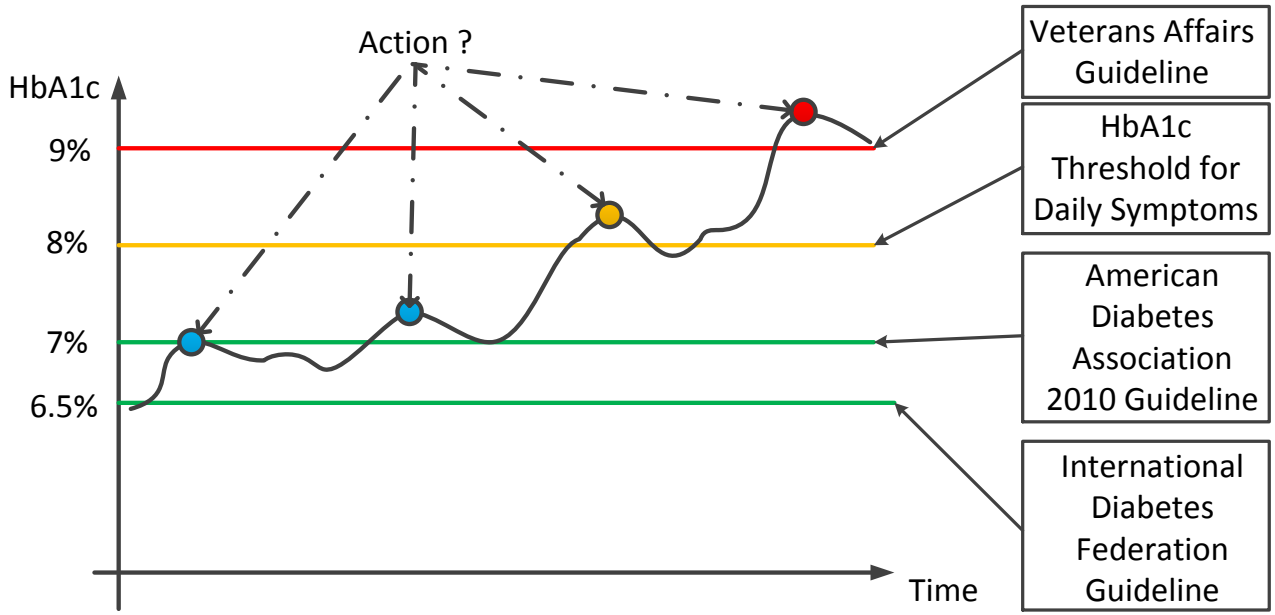
Challenges in Glycemic Control

1: A patient's HbA1c level changes stochastically over time
2: Treatment guidelines are not consistent and are "one size fits all"



Challenges in Glycemic Control

1: A patient's HbA1c level changes stochastically over time
2: Treatment guidelines are not consistent and are "one size fits all"



Why not keep HbA1c as low as possible?

The Action to Control Cardiovascular Risk in Diabetes trial was halted in February, 2008 because of high mortality rate in intensive control arm.

Challenges in Glycemic Control (cont'd)

3: Many glycemic control medications are available with differences in treatment effects, side effects and costs

Oral medications:

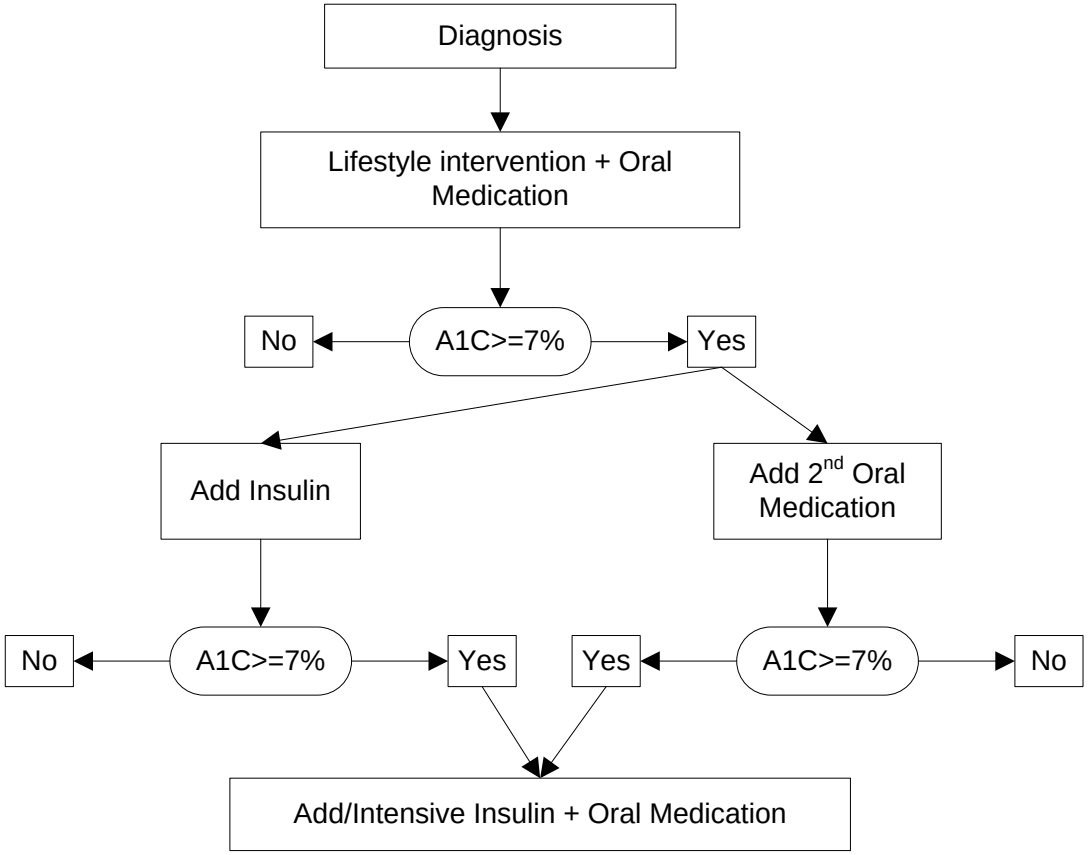
- Metformin (met)
- Sulfonylurea (sulf)
- DPP-4 inhibitor (DPP-4)

Injectable medications:

- Insulin
- GLP-1 agonist (GLP-1)

DPP-4 inhibitor: Dipeptidyl peptidase-4 inhibitor
GLP-1 agonist: Glucagon-like peptide-1 receptor agonists

Current Treatment Guidelines and Regimens



2013 American Diabetes Association's Consensus Algorithm

Glycemic control goal

Varies from HbA1c $\leq$ 6.5% to HbA1c $\leq$ 8%

Treatment Regimens

- met+sulf+insulin
- met+DPP-4+insulin
- met+GLP-1+insulin
- met+insulin

Markov Decision Process Model Components

Time horizon: $t = \{1, 2, \ldots, T\}$

- 1: age at diagnosis
- T: age 100 years
- Discretized into 3-month time intervals

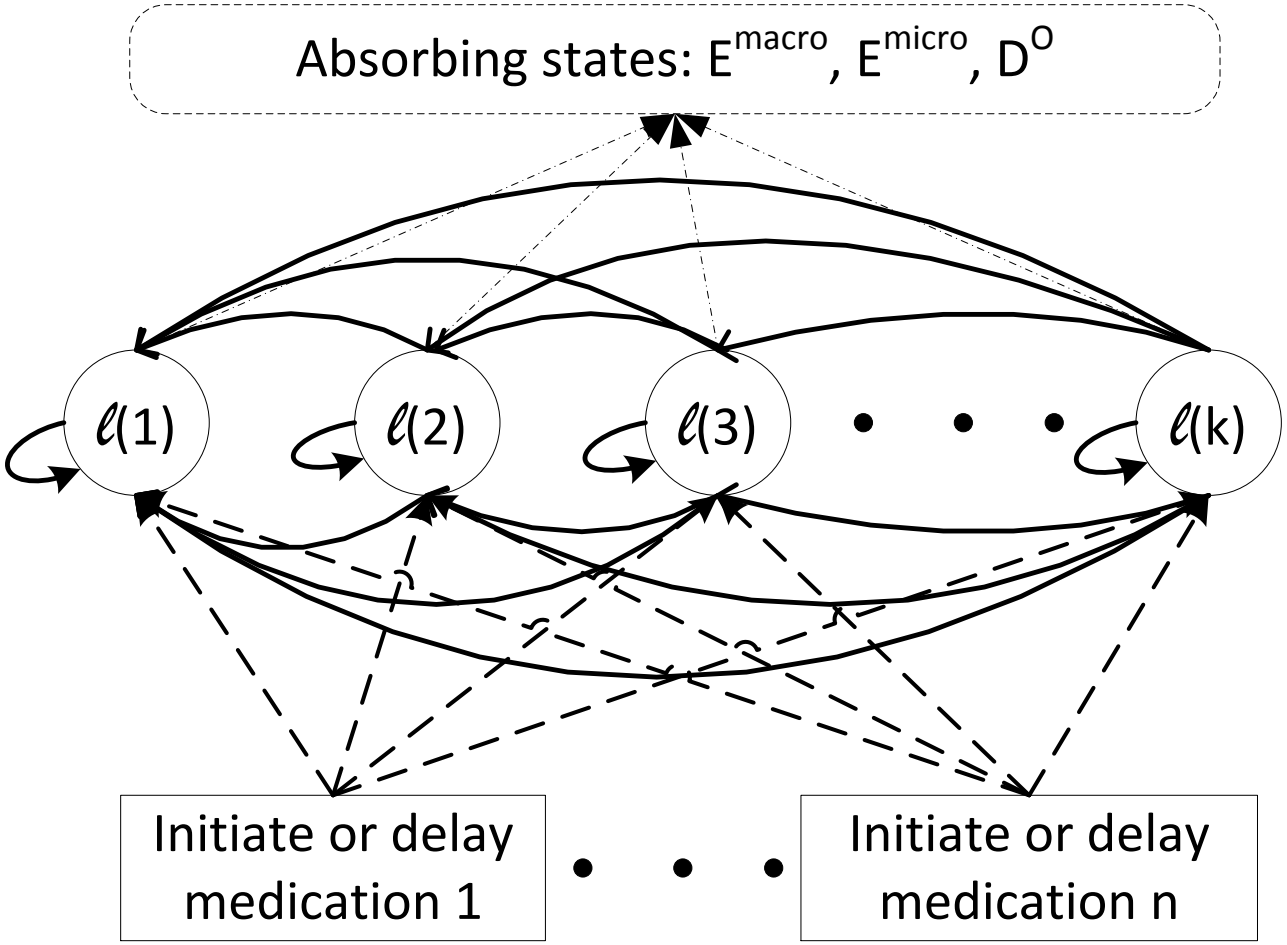
States: $s_t \in \{\mathcal{L} \times \mathcal{M}\} \bigcup \mathcal{D}$

- HbA1c states: $\ell_t \in \mathcal{L} = \{\ell(1), \ell(2), \ldots, \ell(k)\}$
- Medication states: $\boldsymbol{m}_t \in \mathcal{M} = \{(m_{1,t}, m_{2,t}, \ldots, m_{n,t}) | m_{i,t} \in \{0, 1\}\}$
- Absorbing states: $\mathcal{D} = \{E^{\text{macro}}, E^{\text{micro}}, D^{\text{o}}\}$

Actions: α_t

The selection of medication(s) to initiate at each time epoch

One-step Transition Diagram



- Probabilities of transitioning into absorbing states: $p_t^{\mathcal{D}}(s_t, \alpha_t)$
- HbA1c state transition probabilities: $q_{t,\ell_t}(\ell_{t+1})$

One-period Rewards

Quality-adjusted life-years (QALYs)

$$r_t^Q(s_t, \alpha_t) = \begin{cases} 0.25\,(1 - D^{\text{hyper}}(s_t, \alpha_t))(1 - D^{\text{med}}(\alpha_t)), & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\ 0, & \text{otherwise.} \end{cases}$$

- $D^{\text{hyper}}(s_t, \alpha_t)$: disutility of daily symptoms
- $D^{\text{med}}(\alpha_t)$: disutility of using medications

Medication costs

$$r_t^C(s_t, \alpha_t) = \begin{cases} \sum_{i=1}^{n} c_i\, \alpha_{i,t} & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\ 0, & \text{otherwise.} \end{cases}$$

- c_i: the quarterly cost of using medication i

Value Functions

Outcome measures:

- Expected QALYs prior to the first event
- Expected total medication costs

$$V(\pi) = \sum_{i=1}^{k} p(i) \mathbb{E}^{\pi}_{\ell(i)} \left[\sum_{t=1}^{T-1} \lambda^{t-1} r_t(s_t, \alpha_t) + \lambda^{T-1} r_T(s_T) \right]$$

- π: a treatment policy
- $p(i)$: the probability of being in HbA1c state $\ell(i)$ at diagnosis

Data Sources

Model Input	Source
Transition probabilities among HbA1c States	Ingenix dataset
Medication effects on HbA1c	Ingenix dataset
Probabilities of micro- and macro-vascular events	UKPDS outcomes model[1]
Probability of death from other causes	CDC mortality tables
Disutility of medications	Sinha et al. 2010[2]
Disutility of hyperglycemia	Kahn et al. 2010[3]
Quarterly cost of medications	Bennett et al. 2011[4],Yeaw et al. 2012[5]

Ingenix dataset

- An administrative claims data set with linked laboratory data
- $\sim$37,000 patients diagnosed with type 2 diabetes

[1] Clarke et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes, 2006
[2] Sinha et al., Cost and consequences associated with newer medications for glycemic control in T2DM, 2010
[3] Kahn et al., Age at initiation and frequency of screening to detect type 2 diabetes, 2010
[4] Bennett et al., Comparative effectiveness and safety of medications for type 2 diabetes, 2011
[5] Yeaw et al., Cost of self-monitoring of blood glucose in US among patients on an insulin regimen for diabetes, 2012

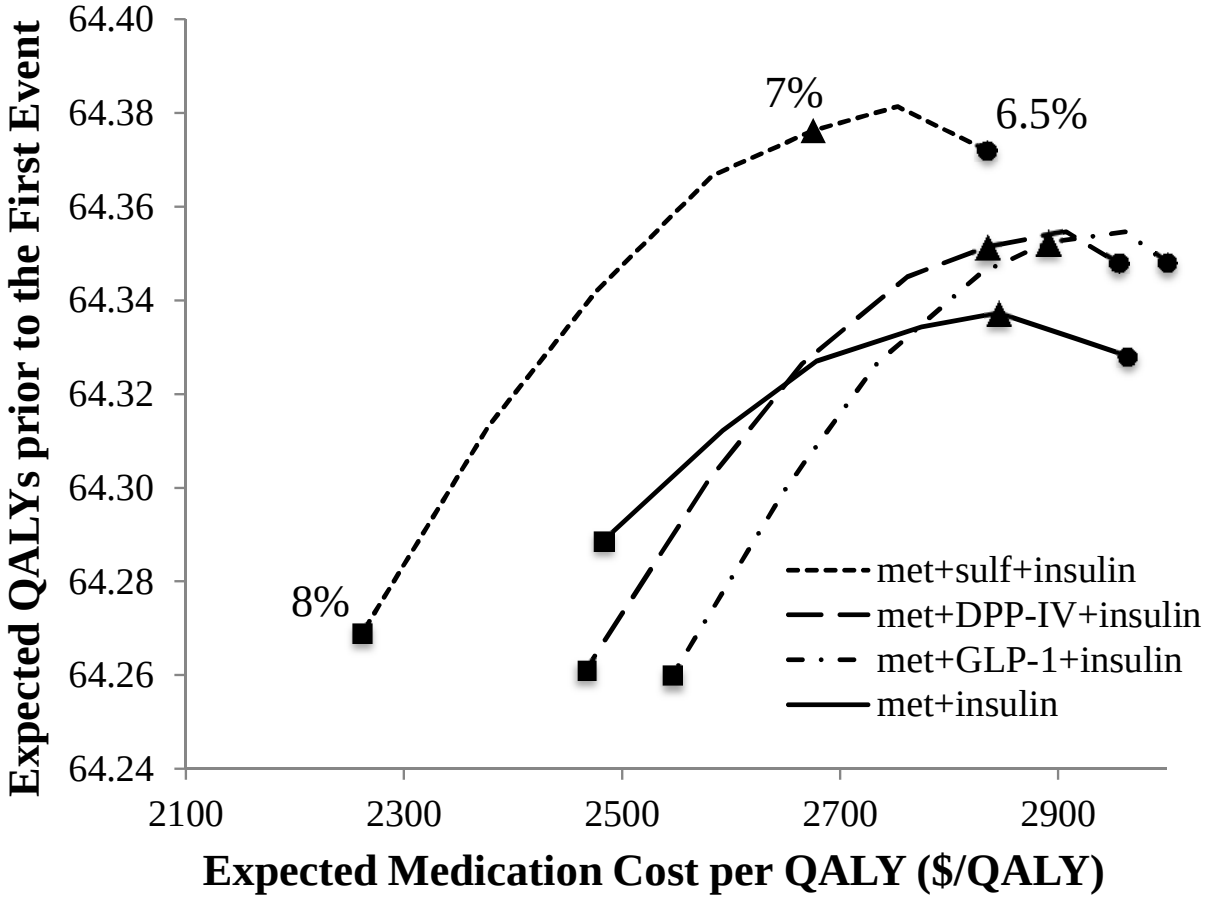
Method for Estimating HbA1c Transition Probabilities

1: Extract HbA1c records at 3 month intervals ($\sim 30,000$ pairs)
2: Adjust HbA1c values for the effects of medication
3: Discretize natural HbA1c values into 10 HbA1c states
4: Estimate the maximum likelihood estimate for the transition probabilities

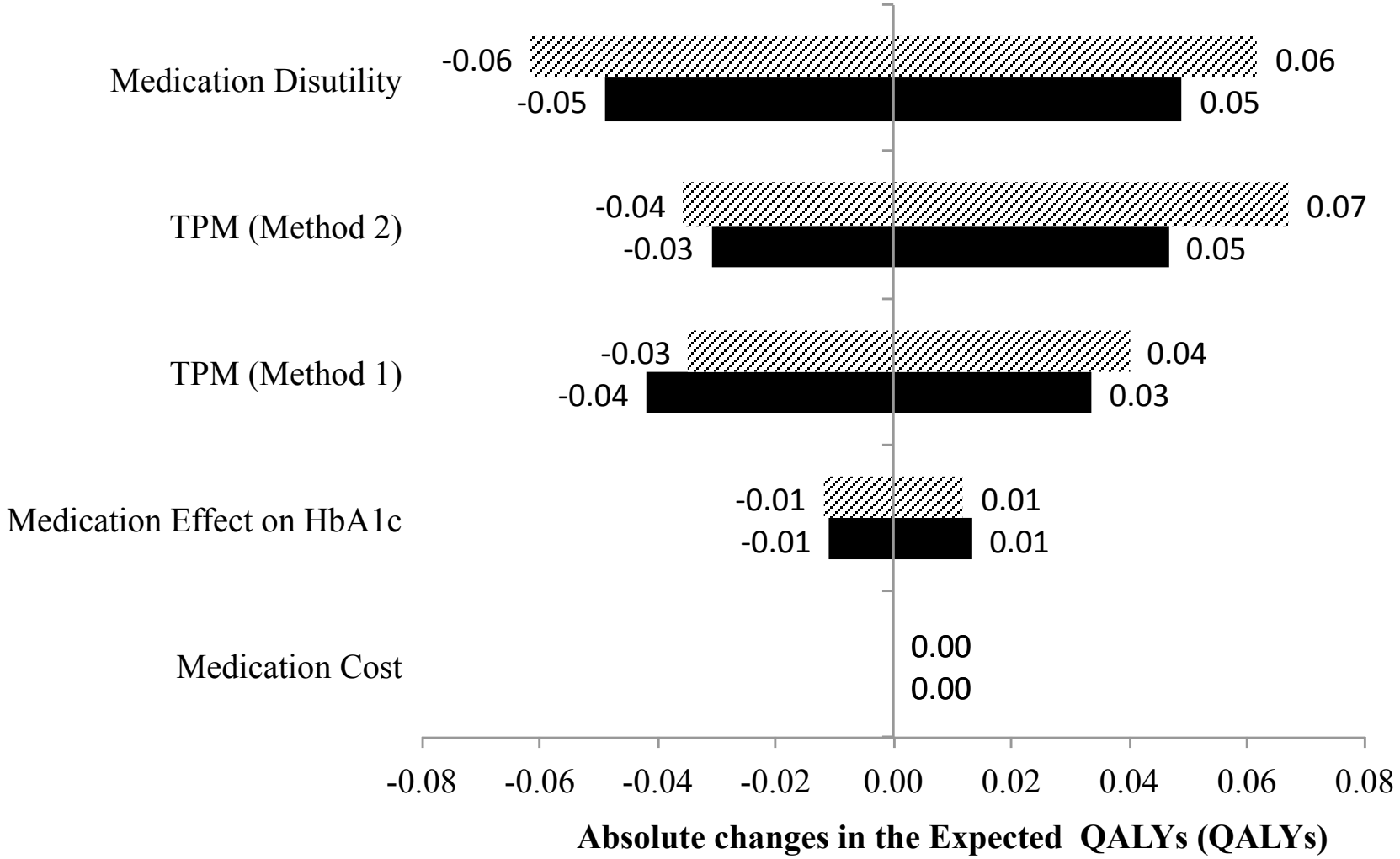
Validation

- Cross validation
- Test set: all two consecutive HbA1c records if the period between the tests was over 3 months ($\sim 97,000$ pairs)

Main Results



Sensitivity Analysis on HbA1c Transition Probabilities



The solid bar represents males, and the hatched bar represents females. TPM: transition probability matrix

TPM Sampling Algorithm

Basic idea: Hit-and-Run (Smith, 1984)

- Algorithm for sampling random vectors over convex region
- Sample each row of the TPM independently

Procedure

- Start with a point in the uncertainty set (e.g. MLE)
- Sample a random direction in the uncertainty set
- Determine the bounds so that the all points along the random direction fall in the standard simplex
- Reduce the line segment iteratively until finding a point in the uncertainty set

Zhang, Y., Mason, J.E., Wu, H., Wilson, J., Denton, B.T., "Conducting Probabilistic Sensitivity Analysis for Markov Decision Processes," *Healthcare Management Science* (in press), 2017.

Conclusions

- Treatment effects of medications are over-estimated by randomized trials
- Use of sulfonylurea as the second-line medication dominates other treatment regimens
- The ADA's glycemic control threshold of HbA1c $\leq 7\%$ results in higher QALYs than other published thresholds
- Results are sensitive to uncertainty in transition probability matrices

Zhang, Y., McCoy, R.G., Mason, J.E., Smith, S.A., Shah, N.D., Denton, B.T., "Second-line agents for glycemic control for type 2 diabetes: Are newer agents better?" Diabetes Care, 2014;37(5):1338-45

Extensions to Treatment Optimization

A robust Markov decision process model (RMDP) that assumes an adversarial context:

- Address uncertainty in transition probabilities by using an *uncertainty set* and *uncertainty budget*
- Generate policies that perform well with respect to maximum likelihood estimates of transition probabilities and are insensitive to uncertainty

Decision Process

Action:

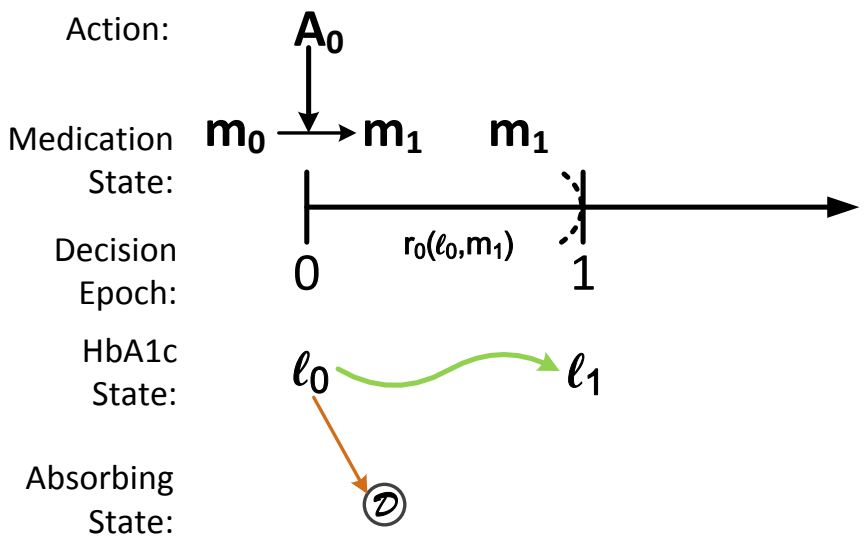
Medication State: $\mathbf{m_0}$

Decision Epoch: 0

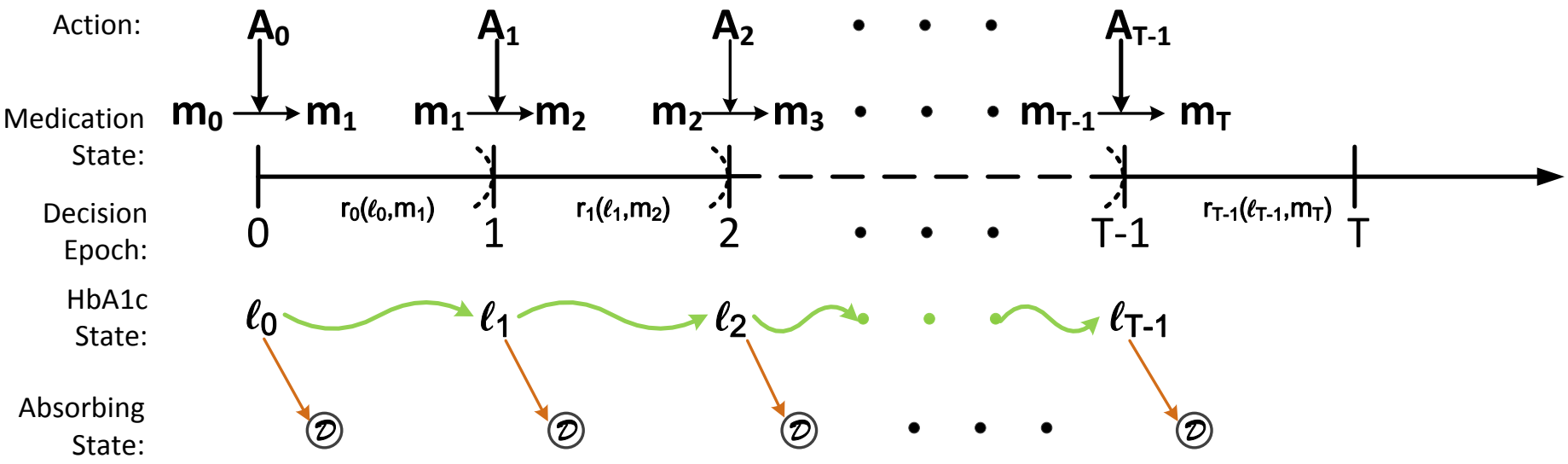
HbA1c State: ℓ_0

Absorbing State:

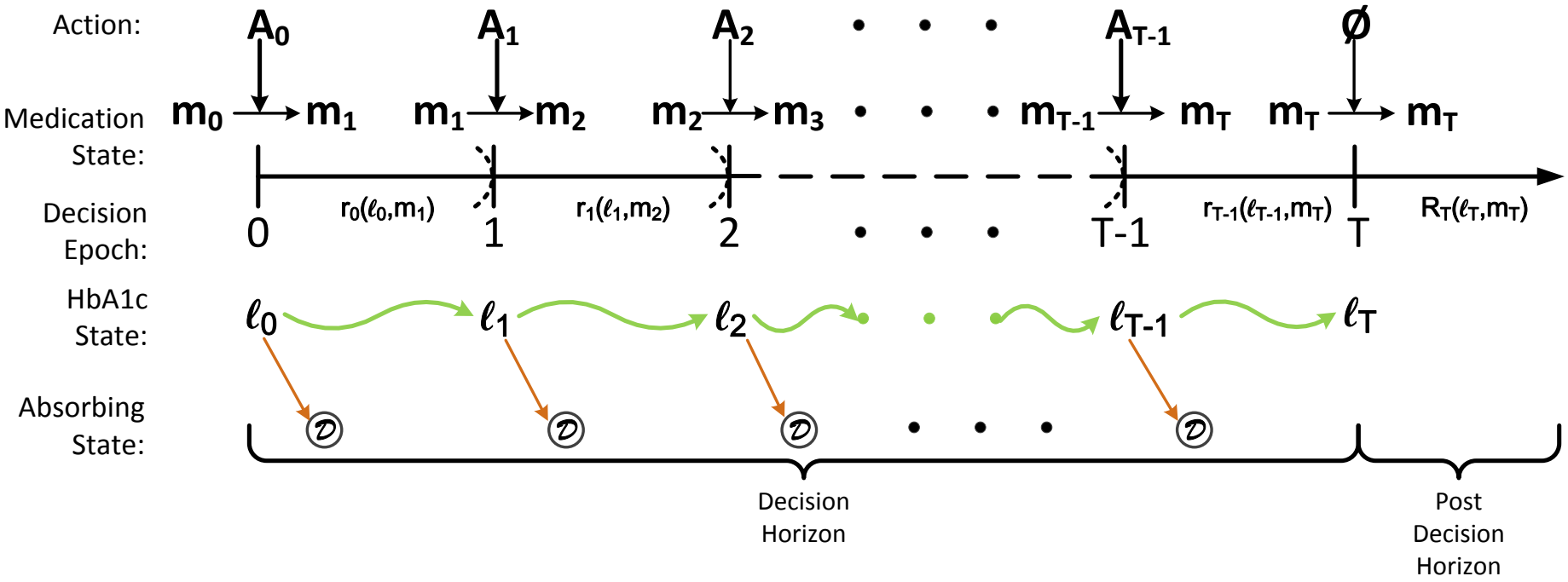
Decision Process



Decision Process



Decision Process



RMDP vs TI-RMDP

The optimal value function of the RMDP-TM

$$v_t^{\text{RMDP-TM}}(s_t) = \begin{cases} \max\limits_{\boldsymbol{\pi}\in\boldsymbol{\Pi}} \min\limits_{\boldsymbol{\theta}_t\in\Theta_t} \mathbb{E}_{s_t}^{\boldsymbol{\theta}_t}\left[\sum\limits_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \boldsymbol{\alpha}_k(s_k)) + \lambda^{T-t} r_T(s_T)\right], & \forall\, s_t \in \mathcal{L}\times\mathcal{M}, \\ 0, & \text{otherwise.} \end{cases}$$

The optimal value function of the time-invariant RMDP

$$v_t^{\text{TI-RMDP-TM}}(s_t) = \begin{cases} \max\limits_{\boldsymbol{\pi}\in\boldsymbol{\Pi}} \min\limits_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{s_t}^{\boldsymbol{\theta}}\left[\sum\limits_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \boldsymbol{\alpha}_k(s_k)) + \lambda^{T-t} r_T(s_T)\right], & \forall\, s_t \in \mathcal{L}\times\mathcal{M}, \\ 0, & \text{otherwise.} \end{cases}$$

Uncertainty Set

Uncertainty Set of TPM:

$Q_t \triangleq [q_{t,\ell_t}(\ell_{t+1})] \in \mathcal{Q}_t$

Interval Model with Uncertainty Budget (IMUB)

$$\mathcal{Q}_t^{\text{IMUB}}(\Gamma) = \prod_{\ell_t \in \mathcal{L}} \mathcal{Q}_{t,\ell_t}^{\text{IMUB}}(\Gamma)$$

Constraints for $\mathcal{Q}_{t,\ell_t}^{\text{IMUB}}(\Gamma)$

- The row vector is in the standard simplex
- Each element lies in its statistical confidence interval
- The total variation from MLE can not exceed the *uncertainty budget*, Γ

$$\sum_{i \in \mathcal{L}} \left(z_{t,\ell_t}^{l}(i) + z_{t,\ell_t}^{u}(i) \right) \leq \Gamma, \; z_{t,\ell_t}^{l}(\ell_{t+1}) \cdot z_{t,\ell_t}^{u}(\ell_{t+1}) = 0$$

Optimality Equations of the RMDP with IMUB

$$v_t^{\text{RMDP}}(\ell_t, \boldsymbol{m}_t) = \max_{\alpha_t(s_t)\in\mathcal{A}_t} \Bigg\{ r_t(s_t, \alpha_t(s_t)) + (1 - p_t^{\text{E}}(s_t, \alpha_t(s_t)))\lambda \times \underbrace{\min_{\boldsymbol{q}_{t,\ell_t}\in\mathcal{Q}_{t,\ell_t}^{\text{IMUB}}(\Gamma)} \sum_{\ell_{t+1}\in\mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) v_{t+1}^{\text{RMDP}}(\ell_{t+1}, \boldsymbol{m}_{t+1})}_{\textit{inner problem}} \Bigg\}$$

Robust Dynamic Programming (RDP) Algorithm [2]

- A backward-induction based algorithm
- The inner problem, a nonlinear program, needs to be solved $|\mathcal{L}| \cdot |\mathcal{M}| \cdot (T-1)$ times

[2] Nilim et al. Robust control of markov decision processes with uncertain transition matrices, *Operations Research*, 2005

Inner Problem: IMUB-NLP

$$
\begin{aligned}
\min \quad & \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) v_{t+1}^{\text{RMDP-TM}}(\ell_{t+1}, \boldsymbol{m}_{t+1})) \\
\text{s.t.} \quad & q_{t,\ell_t}(\ell_{t+1}) = \hat{q}_{t,\ell}(\ell_{t+1}) - \delta^l_{t,\ell}(\ell_{t+1}) z^l_{t,\ell}(\ell_{t+1}) + \delta^u_{t,\ell}(\ell_{t+1}) z^u_{t,\ell}(\ell_{t+1}), \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \ell_{t+1} \in \mathcal{L}, \\
& \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) = 1, \\
& \sum_{\ell_{t+1} \in \mathcal{L}} \left(z^l_{t,\ell_t}(\ell_{t+1}) + z^u_{t,\ell_t}(\ell_{t+1}) \right) \leq \Gamma_{t,\ell}, \\
& z^l_{t,\ell_t}(\ell_{t+1}) \cdot z^u_{t,\ell_t}(\ell_{t+1}) = 0, \qquad \forall \ell_{t+1} \in \mathcal{L}, \\
& 0 \leq z^l_{t,\ell_t}(\ell_{t+1}), z^u_{t,\ell_t}(\ell_{t+1}) \leq 1, \qquad \forall \ell_{t+1} \in \mathcal{L}. \\
& 0 \leq q_{t,\ell_t}(\ell_{t+1}) \leq 1, \qquad \forall \ell_{t+1} \in \mathcal{L},
\end{aligned}
$$

Inner problem formulation: IMUB-LP

Proposition

The following linear reformulation of the inner problem, called IMUB-LP, is equivalent to IMUB-NLP

$$
\begin{aligned}
\min \quad & \sigma_t^{\text{IMUB-LP}}(s_t, \boldsymbol{\alpha}_t(s_t), \Gamma_{t,\ell_t}) = \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) v_{t+1}^{\text{RMDP}}(\ell_{t+1}, \boldsymbol{m}_{t+1})) \\
\text{s.t.} \quad & \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) = 1, \\
& \sum_{\ell_{t+1} \in \mathcal{L}} [x^l_{t,\ell_t}(\ell_{t+1}) + x^u_{t,\ell_t}(\ell_{t+1})] \leq \Gamma_{t,\ell_t}, \\
& x^u_{t,\ell_t}(\ell_{t+1}) \geq \frac{q_{t,\ell_t}(\ell_{t+1}) - \hat{q}_{t,\ell_t}(\ell_{t+1})}{\delta^u_{t,\ell_t}(\ell_{t+1})}, && \forall \ell_{t+1} \in \mathcal{L} \\
& x^l_{t,\ell_t}(\ell_{t+1}) \geq \frac{\hat{q}_{t,\ell_t}(\ell_{t+1}) - q_{t,\ell_t}(\ell_{t+1})}{\delta^l_{t,\ell_t}(\ell_{t+1})}, && \forall \ell_{t+1} \in \mathcal{L}, \\
& 0 \leq x^l_{t,\ell_t}(\ell_{t+1}), x^u_{t,\ell_t}(\ell_{t+1}) \leq 1, && \forall \ell_{t+1} \in \mathcal{L}. \\
& 0 \leq q_{t,\ell_t}(\ell_{t+1}) \leq 1, && \forall \ell_{t+1} \in \mathcal{L}.
\end{aligned}
$$

Fast Algorithm

Proposition

An optimal solution to the inner problem with $\Gamma = |\mathcal{L}|$ can be found using an algorithm with complexity, $O(|\mathcal{L}|)$.

A fast algorithm to solve the inner problem with $\Gamma = |\mathcal{L}|$

For all $\ell(i)$, Set $y_i^l \leftarrow q_{t,\ell_t}^l(\ell(i))$, $y_i^u \leftarrow q_{t,\ell_t}^u(\ell(i))$, $c_i \leftarrow v_{t+1}^{\text{RMDP-TM}}(\ell(i), \boldsymbol{m}_{t+1}(\boldsymbol{\alpha}_t(s_t)))$
for $\tau = 1 \rightarrow |\mathcal{L}|$ **do**
    $\sigma \leftarrow (1 - \sum_{i=1}^{|\mathcal{L}|} y_i^l) c_\tau + \sum_{i=1}^{|\mathcal{L}|} (y_i^u - y_i^l) \min\{c_i - c_\tau, 0\} + \sum_{i=1}^{|\mathcal{L}|} c_i y_i^l$
    if ($\tau == 1$) or ($\tau > 1$ and $\sigma > \sigma_t^{\text{IMUB-LP}}(s_t, \boldsymbol{\alpha}_t(s_t))$) **then**
        $\sigma_t^{\text{IMUB-LP}}(s_t, \boldsymbol{\alpha}_t(s_t)) \leftarrow \sigma$
    else
        Next τ
    end if
end for
Return $\sigma_t^{\text{IMUB-LP}}(s_t, \boldsymbol{\alpha}_t(s_t))$

Proposition

For RMDP with IM model, if the following conditions hold:

(I): $\mathcal{Q}_t^{IM} = \mathcal{Q}_{t'}^{IM}, \forall t, t' \in \mathcal{T}\backslash\{T\}$,

(II): $r_t(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t(s_t))$ *is nonincreasing in* ℓ_t, $\forall\, \boldsymbol{m}_t \in \mathcal{M}$, $\boldsymbol{\alpha}_t(s_t) \in \mathcal{A}_t$, *and* $t \in \mathcal{T}\backslash\{T\}$, *and* $r_T(\ell_T, \boldsymbol{m}_T)$ *is nonincreasing in* ℓ_T, $\forall\, \boldsymbol{m}_T \in \mathcal{M}$,

(III): $p_t^E(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t(s_t))$ *is nondecreasing in* ℓ_t, $\forall\, \boldsymbol{m}_t \in \mathcal{M}$, $\boldsymbol{\alpha}_t(s_t) \in \mathcal{A}_t$, *and* $t \in \mathcal{T}\backslash\{T\}$, *and*

(IV): $Q_{T-1}^{WC,|\mathcal{L}|}$, *has the increasing failure rate (IFR) property*

then

(a): $v_t^{RMDP}(\ell_t, \boldsymbol{m}_t)$ *is nonincreasing in* ℓ_t, $\forall\, \boldsymbol{m}_t \in \mathcal{M}, t \in \mathcal{T}\backslash\{T\}$, *and*

(b): *the optimal policy of nature is stationary.*

Numerical Experiments

Research questions

- Can the RMDP be solved for the diabetes treatment problem?
- What are the benefits of using robust optimal solutions?

Computation Time Comparison

	RMDPs with the uncertainty budget Γ equals to					
Solution methods	0 (= MDP)	1	2	...	9	10
Backward induction	2	-	-	...	-	-
RDP algorithm+NLP	-	10417	10639	...	10209	10891
RDP algorithm+LP	-	1141	1185	...	1169	1129
RDP algorithm+ $O(\|\mathcal{L}\|)$ algorithm	-	-	-	...	-	4

Computation time (in seconds) are presented in terms of central processing unit (CPU) time

- Linear reformulation of the inner problem can significantly reduce the CPU for solving the RMDP
- The RMDP with $\Gamma = 10$ can be solved as efficiently as solving MDP using the backward induction algorithm

Treatment Policy Performance Measures

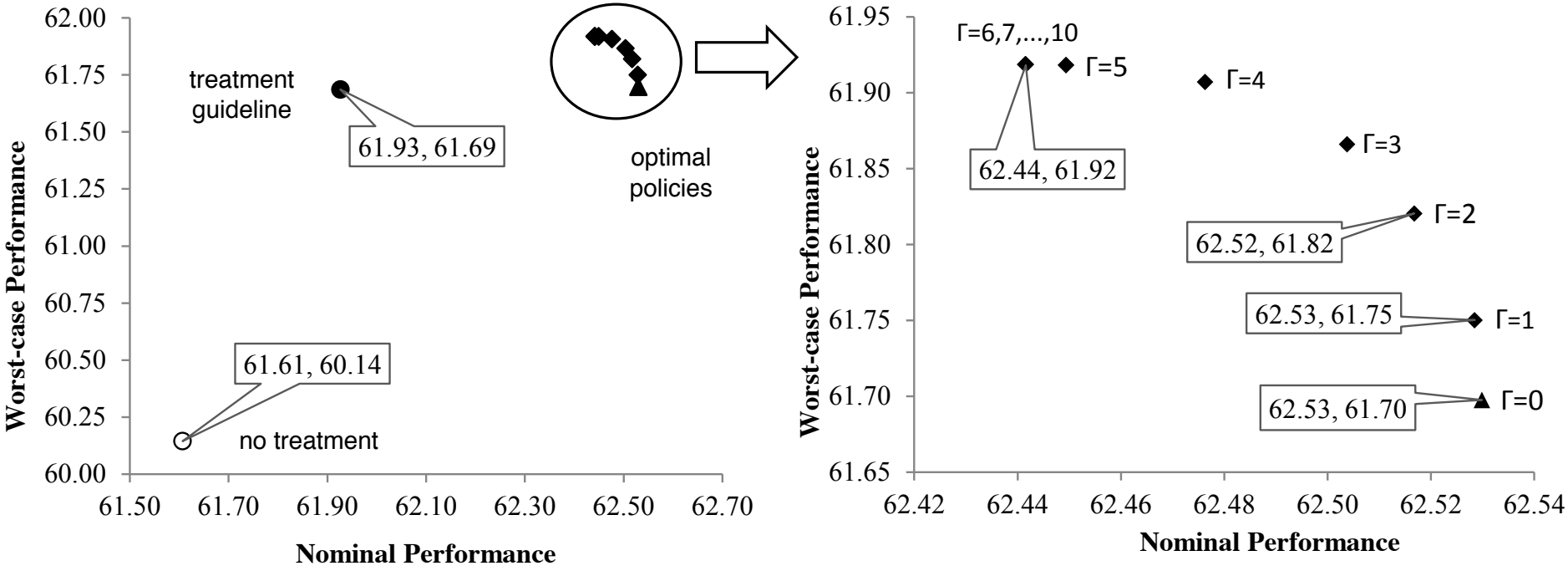
Nominal Performance

The expected total discounted reward under MLE

Wost-case Performance

The expected total discounted reward under the worst-case criterion when the TPM can vary over the entire uncertainty set

Treatment Policy Performance Comparison



Treatment guideline: ADA's concensus algorithm with the glycemic control goal of HbA1c $\leq$ 7%.

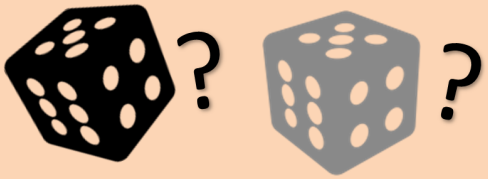
Conclusions

- RMDP with IMUB provides a new approach for controlling robustness of medical treatment decisions with respect to uncertainty in transition probability matrices
- The RMDP with IMUB can be solved efficiently with the proposed solution methods
- RMDP-based optimal policy could provide guidance for clinicians and policy makers when making treatment policies

Zhang, Y., Steimle, L.N., Denton, B.T., "Robust Markov Decision Processes for Medical Treatment Decisions," Working Paper, 2017. Available at *Optimization Online*.

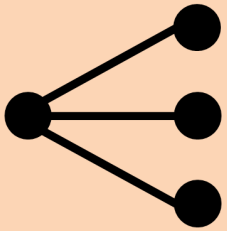
Multi-model Markov Decision Processes: A New Method for Mitigating Parameter Ambiguity

Incorporate ambiguity in MDPs by allowing multiple models of parameters



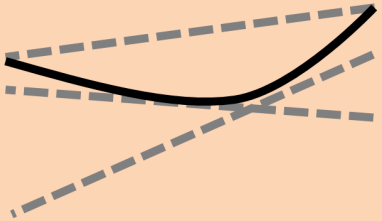
Find single policy to maximize weighted value across models

Establish connections to stochastic programming



Develop exact solution methods

Design a fast heuristic supported by error bounds



Illustrate method on case study of prevention for heart disease

Steimle, L.E., Kaufman, D., and Denton, B.T. “Multi-model Markov decision processes,” Working paper, 2018. Available at *Optimization Online*.

Acknowledgments

Lauren Steimle, PhD Student University of Michigan

Yuanhui Zhang, PhD Research Triangle Institute

Jennifer Mason, PhD University of Virginia

Nilay Shah, PhD Mayo Clinic

Steven Smith, MD Mayo Clinic

Jim Wilson, PhD NC State University

This project is funded in part by the National Science Foundation through Grant Number: CMMI 1536444.

Brian Denton, Department of Industrial and Operations Engineering, University of Michigan

Email: btdenton@umich.edu

Homepage: `http://umich.edu/~btdenton`

Backup Slides

Interval Matrix (IM) Model

$$\mathcal{Q}^{\text{IM}}_{t,\ell_t} = \left\{ \boldsymbol{q}_{t,\ell_t} \in \mathbb{R}_+^{|\mathcal{L}|} \;\middle|\; \begin{array}{ll} \sum_{\ell_{t+1}\in\mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) = 1, & \\ q^l_{t,\ell_t}(\ell_{t+1}) \le q_{t,\ell_t}(\ell_{t+1}) \le q^u_{t,\ell_t}(\ell_{t+1}), & \forall \ell_{t+1} \in \mathcal{L}, \\ 0 \le q_{t,\ell_t}(\ell_{t+1}) \le 1, & \forall \ell_{t+1} \in \mathcal{L}, \end{array} \right\}$$

- $q^l_{t,\ell_t}(\ell_{t+1}) = \tilde{q}_{\ell_t}(\ell_{t+1}) - \left[\chi^2_{|\mathcal{L}|-1,\alpha/(2|\mathcal{L}|)} \frac{\tilde{q}_{\ell_t}(\ell_{t+1})(1-\tilde{q}_{\ell_t}(\ell_{t+1}))}{N_{\ell_t}}\right]^{\frac{1}{2}}$
- $q^u_{t,\ell_t}(\ell_{t+1}) = \tilde{q}_{\ell_t}(\ell_{t+1}) + \left[\chi^2_{|\mathcal{L}|-1,1-\alpha/(2|\mathcal{L}|)} \frac{\tilde{q}_{\ell_t}(\ell_{t+1})(1-\tilde{q}_{\ell_t}(\ell_{t+1}))}{N_{\ell_t}}\right]^{\frac{1}{2}}$
- $\tilde{q}_{\ell_t}(\ell_{t+1})$: maximum likelihood estimate (MLE) of $q_{\ell_t}(\ell_{t+1})$
- N_{ℓ_t}: total number of patients in state ℓ_t

Robust Dynamic Programming (RDP) Algorithm

Robust Dynamic Programming (RDP) Algorithm (Nilim et al. 2005)

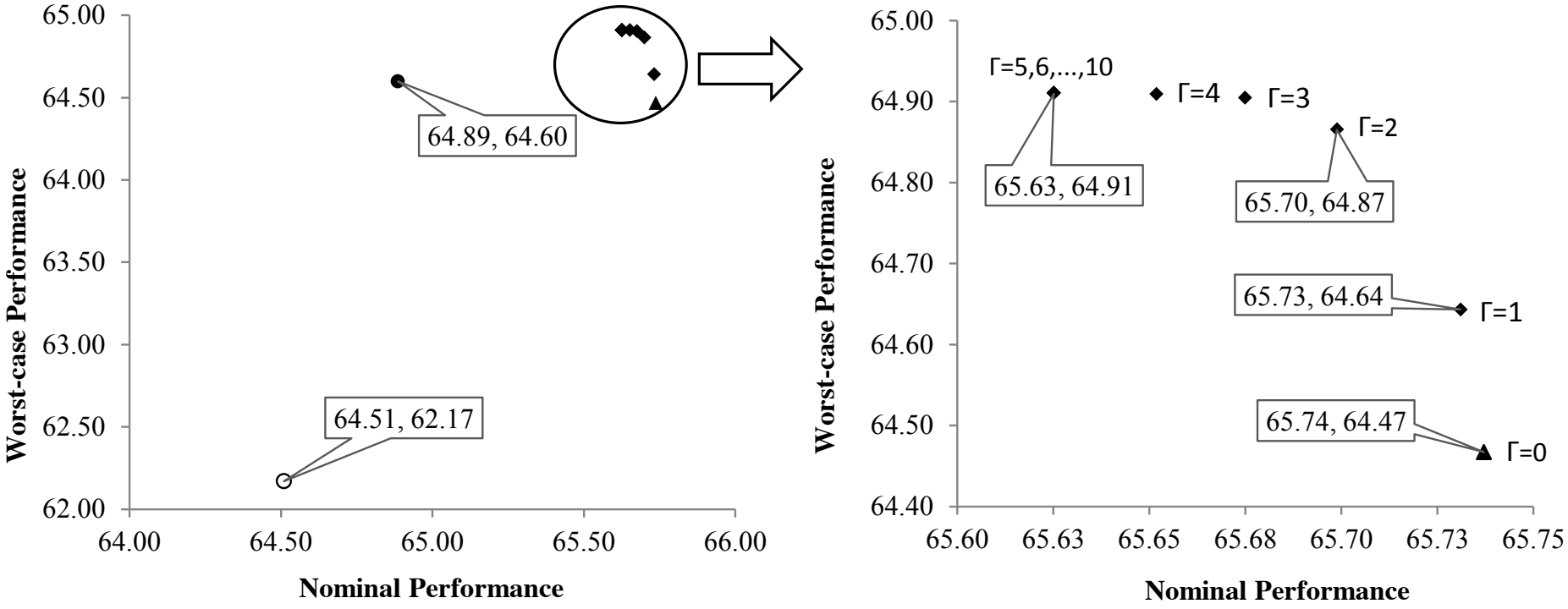
$v_T^{\text{RMDP-TM}}(\ell_T, \boldsymbol{m}_T) \leftarrow r_T(\ell_T, \boldsymbol{m}_T),\ \forall \ell_T \in \mathcal{L}, \boldsymbol{m}_T \in \mathcal{M}$
for $t = T - 1 \rightarrow 0$, $\ell_t \in \mathcal{L}$, and $\boldsymbol{m}_t \in \mathcal{M}$ **do**
    for $\boldsymbol{\alpha}_t \in \mathcal{A}_t(\ell_t, \boldsymbol{m}_t)$ **do**
        Solve the inner problem and calculate
        $w_t(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t) = r_t(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t) + \lambda[1 - p_t^E(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t)]\sigma_t^*(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t, \Gamma_{t,\ell_t})$
    end for
    Update Value Function: $v_t^{\text{RMDP-TM}}(\ell_t, \boldsymbol{m}_t) \leftarrow \max_{\boldsymbol{\alpha}_t \in \mathcal{A}_t(\ell_t, \boldsymbol{m}_t)} \{w_t(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t)\}$
    Update Optimal Action Set: $\mathcal{A}_t^*(\ell_t, \boldsymbol{m}_t) \leftarrow \operatorname{argmax}_{\boldsymbol{\alpha}_t \in \mathcal{A}_t(\ell_t, \boldsymbol{m}_t)} \{w_t(\ell_t, \boldsymbol{m}_t, \boldsymbol{\alpha}_t)\}$
end for

Interval Model with Uncertainty Budget (IMUB)

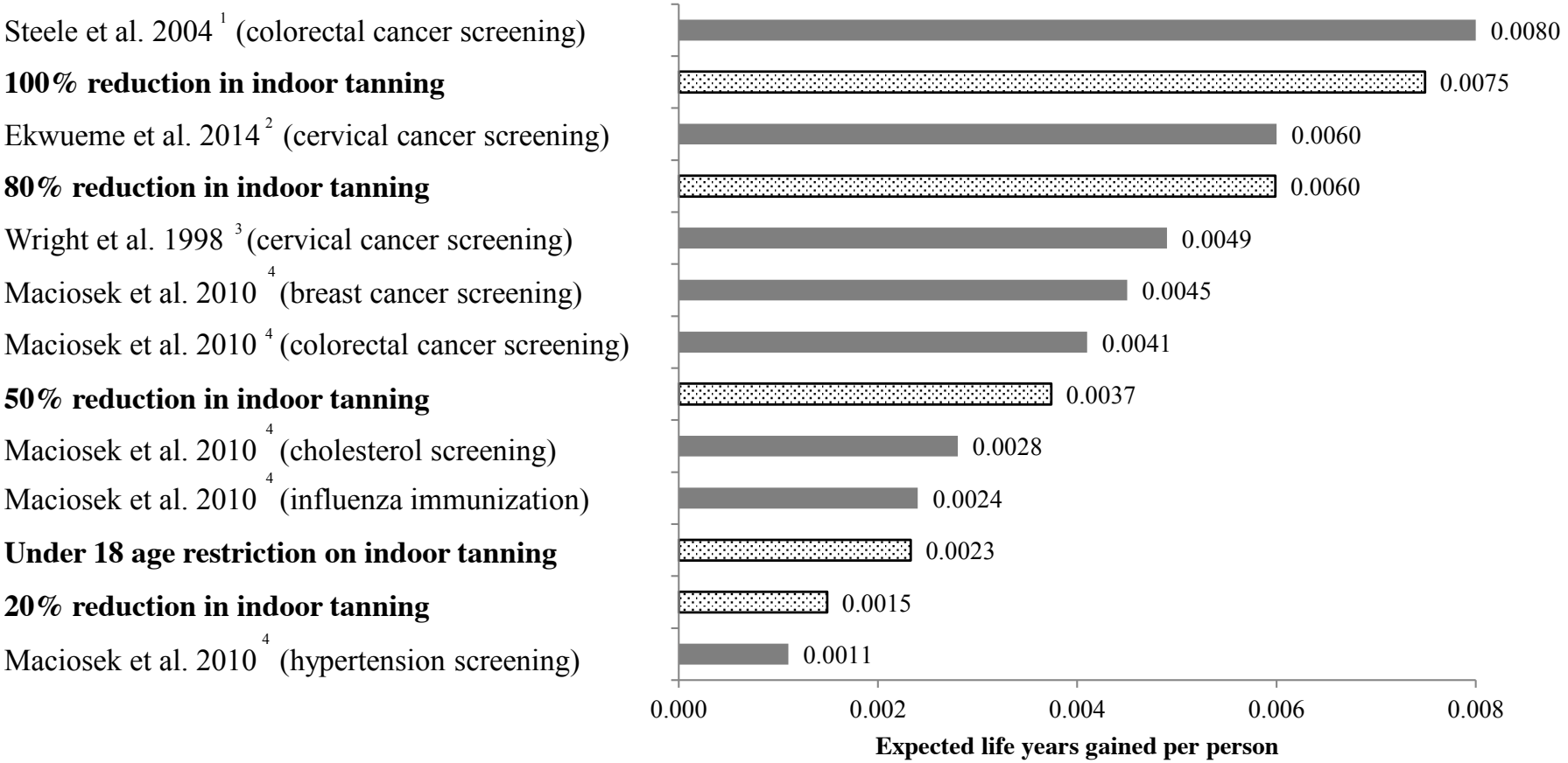
$$\mathcal{Q}^{\text{IMUB}}_{t,\ell_t}(\Gamma_{t,\ell_t}) = \left\{ \boldsymbol{q}_{t,\ell_t} \in \mathbb{R}_+^{|\mathcal{L}|} \;\middle|\; \begin{array}{ll} q_{t,\ell_t}(\ell_{t+1}) = \hat{q}_{t,\ell}(\ell_{t+1}) - \delta^l_{t,\ell_t}(\ell_{t+1}) z^l_{t,\ell_t}(\ell_{t+1}) & \\ \qquad\qquad + \delta^u_{t,\ell_t}(\ell_{t+1}) z^u_{t,\ell_t}(\ell_{t+1}), & \forall \ell_{t+1} \in \mathcal{L}, \\ \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) = 1, & \\ \sum_{\ell_{t+1} \in \mathcal{L}} \left(z^l_{t,\ell_t}(\ell_{t+1}) + z^u_{t,\ell_t}(\ell_{t+1}) \right) \leq \Gamma_{t,\ell_t}, & \\ z^l_{t,\ell_t}(\ell_{t+1}) \cdot z^u_{t,\ell_t}(\ell_{t+1}) = 0, & \forall \ell_{t+1} \in \mathcal{L}, \\ 0 \leq z^l_{t,\ell_t}(\ell_{t+1}), z^u_{t,\ell_t}(\ell_{t+1}) \leq 1, & \forall \ell_{t+1} \in \mathcal{L} \\ 0 \leq q_{t,\ell_t}(\ell_{t+1}) \leq 1, & \forall \ell_{t+1} \in \mathcal{L}, \end{array} \right\}$$

- $\hat{q}_{t,\ell}(\ell_{t+1})$: nominal value (the point estimate)
- $\delta^{l(u)}_{t,\ell_t}(\ell_{t+1})$: maximum left(right)-hand side variation
- $z^{l(u)}_{t,\ell_t}(\ell_{t+1})$: proportion of variation
- Γ_{t,ℓ_t}: the total uncertainty budget on row ℓ_t of Q_t

Treatment Policy Performance Comparison-Female



Simulation-based Analyses on Life Years Gained from Selected Population-based Prevention Programs



[1] Steele *et al*. Methods and Economic Considerations: Group 1 Report. ESGE/UEGF Colorectal Cancer--Public Awareness Campaign. The Public/Professional Interface Workshop: Oslo, Norway, June 20 - 22, 2003. fulfillment corporate. *Endoscopy* 2004;36(4):349-53

[2] Ekwueme *et al*, Impact of the National Breast and Cervical Cancer Early Detection Program on Cervical Cancer Mortality among Uninsured Low-Income Women in the U.S., 1991–2007. *American Journal of Preventive Medicine* 47, no. 3 : 300-08.

[3] Wright *et al*. Gains in life expectancy from medical interventions--standardizing data on outcomes. *N Engl J Med* 1998;339(6):380-6.

[4] Maciosek *et al*. Greater use of preventive services in U.S. health care could save lives at little or no cost. *Health Aff (Millwood)* 2010;29(9):1656-6

Optimal Value Functions of the RMDP

$$v_t^{\text{RMDP}}(s_t) = \begin{cases} \max\limits_{\boldsymbol{\pi}\in\boldsymbol{\Pi}} \min\limits_{\boldsymbol{\theta}_t\in\Theta_t} \mathbb{E}_{s_t}^{\boldsymbol{\theta}_t}\left[\sum\limits_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \boldsymbol{\alpha}_k(s_k)) + \lambda^{T-t} r_T(s_T)\right], & \\ & \forall\, s_t \in \mathcal{L}\times\mathcal{M}, \\ 0, & \text{otherwise.} \end{cases}$$

- Treatment policy: π
- All admissible treatment policies: $\boldsymbol{\Pi}$
- Adversaries policy: $\theta_t \triangleq (Q_t, Q_{t+1}, \ldots, Q_{T-1})$
- All admissible adversary policies: $\Theta_t \triangleq \prod\limits_{\tau=t}^{T-1} \mathcal{Q}_\tau$