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

- \(~29\) 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”

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

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

Glycemic control goal

Varies from HbA1c ≤ 6.5% to HbA1c ≤ 8%

Treatment Regimens

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

2013 American Diabetes Association’s Consensus Algorithm
# 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:
\[ s_t \in \{L \times M\} \cup D \]
- HbA1c states: \( \ell_t \in L = \{\ell(1), \ell(2), \ldots, \ell(k)\} \)
- Medication states: \( m_t \in M = \{(m_1,t, m_2,t, \ldots, m_n,t)|m_i,t \in \{0, 1\}\} \)
- Absorbing states: \( D = \{E^{\text{macro}}, E^{\text{micro}}, D^o\} \)

## Actions:
\( \alpha_t \)
The selection of medication(s) to initiate at each time epoch.
Absorbing states: $E^{\text{macro}}, E^{\text{micro}}, D^O$

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

Quality-adjusted life-years (QALYs)

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

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

Medication costs

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

- \(c_i\): the quarterly cost of using medication \(i\)
Value Functions

Outcome measures:

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

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

- \(\pi\): a treatment policy
- \(p(i)\): the probability of being in HbA1c state \(\ell(i)\) at diagnosis
### Data Sources

<table>
<thead>
<tr>
<th>Model Input</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition probabilities among HbA1c States</td>
<td>Ingenix dataset</td>
</tr>
<tr>
<td>Medication effects on HbA1c</td>
<td>Ingenix dataset</td>
</tr>
<tr>
<td>Probabilities of micro- and macro-vascular events</td>
<td>UKPDS outcomes model(^1)</td>
</tr>
<tr>
<td>Probability of death from other causes</td>
<td>CDC mortality tables</td>
</tr>
<tr>
<td>Disutility of medications</td>
<td>Sinha et al. 2010(^2)</td>
</tr>
<tr>
<td>Disutility of hyperglycemia</td>
<td>Kahn et al. 2010(^3)</td>
</tr>
<tr>
<td>Quarterly cost of medications</td>
<td>Bennett et al. 2011(^4), Yeaw et al. 2012(^5)</td>
</tr>
</tbody>
</table>

---

### Ingenix dataset

- An administrative claims data set with linked laboratory data

- \(~37,000\) 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 (∼ 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 (∼ 97,000 pairs)
Main Results

Expected QALYs prior to the First Event

Expected Medication Cost per QALY ($/QALY)

- met+sulf+insulin
- met+DPP-IV+insulin
- met+GLP-1+insulin
- met+insulin
Sensitivity Analysis on HbA1c Transition Probabilities

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

The absolute changes in the Expected QALYs (QALYs) for different factors are shown:
- Medication Disutility
- TPM (Method 1)
- TPM (Method 2)
- Medication Effect on HbA1c
- Medication Cost

Absolute changes in the Expected QALYs (QALYs):

-0.06 to 0.06
-0.04 to 0.07
-0.03 to 0.04
-0.01 to 0.01
-0.08 to 0.00
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 all points along the random direction fall in the standard simplex
- Reduce the line segment iteratively until finding a point in the uncertainty set

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

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:

Decision
Epoch:

HbA1c
State:

Absorbing
State:

\[ m_0 \]

\[ \ell_0 \]
Decision Process

Action: Medication
State: Decision
Epoch: HbA1c
State: Absorbing
State: m0

m0 → m1 → m1

0 → r0(ℓ0,m1) → 1

ℓ0 → ℓ1

0 1
m1
A0

Michigan
Brian Denton
Robust Medical Treatment Decisions
**Decision Process**

- **Action:** 
  - $A_0 \rightarrow m_0 \rightarrow m_1 \uparrow A_1 \rightarrow m_1 \rightarrow m_2 \uparrow A_2 \rightarrow m_2 \rightarrow m_3 \uparrow \cdots \uparrow A_{T-1} \rightarrow m_{T-1} \rightarrow m_T$

- **Medication State:** $m_0, m_1, m_2, \ldots, m_T$

- **Decision Epoch:** $0, 1, 2, \ldots, T$

- **HbA1c State:** $\rho_0(\ell_0, m_1), r_1(\ell_1, m_2), \ldots, r_{T-1}(\ell_{T-1}, m_T)$

- **Absorbing State:** $\mathcal{D}$

---

**Michigan**

Brian Denton

Robust Medical Treatment Decisions
Decision Process

Action: $A_0 \rightarrow m_1 \rightarrow A_1 \rightarrow m_1 \rightarrow m_2 \rightarrow A_2 \rightarrow m_2 \rightarrow m_3 \rightarrow \cdots \rightarrow A_{T-1} \rightarrow m_{T-1} \rightarrow m_T \rightarrow \varnothing$

Medication State: $m_0 \rightarrow m_1 \rightarrow m_1 \rightarrow m_2 \rightarrow m_2 \rightarrow m_3 \rightarrow \cdots \rightarrow m_T$

Decision Epoch: $r_0(\ell_0, m_1) \rightarrow r_1(\ell_1, m_2) \rightarrow \cdots \rightarrow r_{T-1}(\ell_{T-1}, m_T) \rightarrow R_T(\ell_T, m_T)$

HbA1c State: $\ell_0 \rightarrow \ell_1 \rightarrow \ell_2 \rightarrow \cdots \rightarrow \ell_{T-1} \rightarrow \ell_T$

Absorbing State: $D \rightarrow D \rightarrow D \rightarrow \cdots \rightarrow D$
The optimal value function of the RMDP-TM

\[ V_{t}^{\text{RMDP-TM}}(s_t) = \begin{cases} \max_{\pi \in \Pi} \min_{\theta_t \in \Theta_t} \mathbb{E}_{s_t} \left[ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k(s_k)) + \lambda^{T-t} r_T(s_T) \right], & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\ 0, & \text{otherwise.} \end{cases} \]

The optimal value function of the time-invariant RMDP

\[ V_{t}^{\text{TI-RMDP-TM}}(s_t) = \begin{cases} \max_{\pi \in \Pi} \min_{\theta \in \Theta} \mathbb{E}_{s_t} \left[ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k(s_k)) + \lambda^{T-t} r_T(s_T) \right], & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\ 0, & \text{otherwise.} \end{cases} \]
Uncertainty Set

Uncertainty Set of TPM:

\[ Q_t \triangleq [q_t, \ell_t(\ell_{t+1})] \in Q_t \]

Interval Model with Uncertainty Budget (IMUB)

\[ Q_{t,\ell_t}^{\text{IMUB}}(\Gamma) = \prod_{\ell_t \in \mathcal{L}} Q_{t,\ell_t}^{\text{IMUB}}(\Gamma) \]

Constraints for \( Q_{t,\ell_t}^{\text{IMUB}}(\Gamma) \)

- 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, \( \Gamma \)

\[ \sum_{i \in \mathcal{L}} (z^l_{t,\ell_t}(i) + z^u_{t,\ell_t}(i)) \leq \Gamma, \quad z^l_{t,\ell_t}(\ell_{t+1}) \cdot z^u_{t,\ell_t}(\ell_{t+1}) = 0 \]
Optimality Equations of the RMDP with IMUB

\[ \nu_{t}^{RMDP}(\ell_t, m_t) = \max_{\alpha_t(s_t) \in A_t} \left\{ r_t(s_t, \alpha_t(s_t)) + (1 - p^E_t(s_t, \alpha_t(s_t)))\lambda \right\} \times \min_{q_t, \ell_t \in Q_{t, t+1}(\Gamma)} \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_t}(\ell_{t+1}) \nu_{t+1}^{RMDP}(\ell_{t+1}, m_{t+1}) \}

inner problem

Robust Dynamic Programming (RDP) Algorithm

- 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

\[
\begin{align*}
\min & \quad \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) v_{t+1}^{\text{RMDP-TM}}(\ell_{t+1}, m_{t+1}) \\
\text{s.t.} & \quad q_{t,\ell_t}(\ell_{t+1}) = \hat{q}_{t,\ell}(\ell_{t+1}) - \delta^l_{t,\ell}(\ell_{t+1}) z^l_{t,\ell}(\ell_{t+1}) + \delta^u_{t,\ell}(\ell_{t+1}) z^u_{t,\ell}(\ell_{t+1}), \\
& \quad \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) = 1, \\
& \quad \sum_{\ell_{t+1} \in \mathcal{L}} (z^l_{t,\ell_t}(\ell_{t+1}) + z^u_{t,\ell_t}(\ell_{t+1})) \leq \Gamma_{t,\ell}, \\
& \quad z^l_{t,\ell_t}(\ell_{t+1}) \cdot z^u_{t,\ell_t}(\ell_{t+1}) = 0, \\
& \quad 0 \leq z^l_{t,\ell_t}(\ell_{t+1}), z^u_{t,\ell_t}(\ell_{t+1}) \leq 1, \\
& \quad 0 \leq q_{t,\ell_t}(\ell_{t+1}) \leq 1,
\end{align*}
\]
Proposition

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

\[
\begin{align*}
\min_{t} & \quad \sigma_{t}^{\text{IMUB-LP}}(s_{t}, \alpha_{t}(s_{t}), \Gamma_{t, t_{t}}) = \sum_{\ell_{t+1} \in L} q_{t, \ell_{t}}(\ell_{t+1}) v_{t+1}^{\text{RMDP}}(\ell_{t+1}, m_{t+1}) \\
\text{s.t.} & \quad \sum_{\ell_{t+1} \in L} q_{t, \ell_{t}}(\ell_{t+1}) = 1, \\
& \quad \sum_{\ell_{t+1} \in L} \left[ x_{t, \ell_{t}}^{l}(\ell_{t+1}) + x_{t, \ell_{t}}^{u}(\ell_{t+1}) \right] \leq \Gamma_{t, \ell_{t}}, \\
& \quad x_{t, \ell_{t}}^{u}(\ell_{t+1}) \geq \frac{q_{t, \ell_{t}}(\ell_{t+1}) - \hat{q}_{t, \ell_{t}}(\ell_{t+1})}{\delta_{t, \ell_{t}}(\ell_{t+1})}, \quad \forall \ell_{t+1} \in L \\
& \quad x_{t, \ell_{t}}^{l}(\ell_{t+1}) \geq \frac{\hat{q}_{t, \ell_{t}}(\ell_{t+1}) - q_{t, \ell_{t}}(\ell_{t+1})}{\delta_{t, \ell_{t}}(\ell_{t+1})}, \quad \forall \ell_{t+1} \in L, \\
& \quad 0 \leq x_{t, \ell_{t}}^{l}(\ell_{t+1}), x_{t, \ell_{t}}^{u}(\ell_{t+1}) \leq 1, \\
& \quad 0 \leq q_{t, \ell_{t}}(\ell_{t+1}) \leq 1, \quad \forall \ell_{t+1} \in L.
\end{align*}
\]
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 $\Gamma = |\mathcal{L}|$

1: For all $\ell(i)$, Set $y^l_i \leftarrow q^l_{t, \ell_t}(\ell(i))$, $y^u_i \leftarrow q^u_{t, \ell_t}(\ell(i))$, $c_i \leftarrow v^{\text{RMDP-TM}}_{t+1}(\ell(i), m_{t+1}(\alpha_t(s_t)))$

2: for $\tau = 1 \rightarrow |\mathcal{L}|$ do

3: $\sigma \leftarrow (1 - \sum_{i=1}^{|\mathcal{L}|} y^l_i)c_\tau + \sum_{i=1}^{|\mathcal{L}|} (y^u_i - y^l_i) \min\{c_i - c_\tau, 0\} + \sum_{i=1}^{|\mathcal{L}|} c_i y^l_i$

4: if ($\tau == 1$) or ($\tau > 1$ and $\sigma > \sigma^i_{\text{IMUB-LP}}(s_t, \alpha_t(s_t))$) then

5: $\sigma^i_{\text{IMUB-LP}}(s_t, \alpha_t(s_t)) \leftarrow \sigma$

6: else

7: Next $\tau$

8: end if

9: end for

10: Return $\sigma^i_{\text{IMUB-LP}}(s_t, \alpha_t(s_t))$
Proposition

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

(I): \( Q_{t}^{IM} = Q_{t'}^{IM}, \forall t, t' \in T \setminus \{ T \}, \)

(II): \( r_{t}(\ell_{t}, m_{t}, \alpha_{t}(s_{t})) \) is nonincreasing in \( \ell_{t}, \forall m_{t} \in \mathcal{M}, \alpha_{t}(s_{t}) \in \mathcal{A}_{t}, \) and \( t \in T \setminus \{ T \}, \) and \( r_{T}(\ell_{T}, m_{T}) \) is nonincreasing in \( \ell_{T}, \forall m_{T} \in \mathcal{M}, \)

(III): \( p_{t}^{E}(\ell_{t}, m_{t}, \alpha_{t}(s_{t})) \) is nondecreasing in \( \ell_{t}, \forall m_{t} \in \mathcal{M}, \alpha_{t}(s_{t}) \in \mathcal{A}_{t}, \) and \( t \in T \setminus \{ T \}, \) and

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

then

(a): \( v_{t}^{RMDP}(\ell_{t}, m_{t}) \) is nonincreasing in \( \ell_{t}, \forall m_{t} \in \mathcal{M}, t \in T \setminus \{ T \}, \) 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

<table>
<thead>
<tr>
<th>Solution methods</th>
<th>0 ((=) MDP)</th>
<th>1</th>
<th>2</th>
<th>\ldots</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backward induction</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>\ldots</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RDP algorithm+NLP</td>
<td>-</td>
<td>10417</td>
<td>10639</td>
<td>\ldots</td>
<td>10209</td>
<td>10891</td>
</tr>
<tr>
<td>RDP algorithm+LP</td>
<td>-</td>
<td>1141</td>
<td>1185</td>
<td>\ldots</td>
<td>1169</td>
<td>1129</td>
</tr>
<tr>
<td>RDP algorithm+ (O(</td>
<td>\mathcal{L}</td>
<td>)) algorithm</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>\ldots</td>
</tr>
</tbody>
</table>

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.
Nominal Performance
The expected total discounted reward under MLE

Worst-case Performance
The expected total discounted reward under the worst-case criterion when the TPM can vary over the entire uncertainty set
Treatment guideline: ADA's consensus algorithm with the glycemic control goal of HbA1c ≤ 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.

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

<table>
<thead>
<tr>
<th>Incorporate ambiguity in MDPs by allowing multiple models of parameters</th>
<th>Find single policy to maximize weighted value across models</th>
<th>Establish connections to stochastic programming</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="dice.png" alt="Dice" /></td>
<td><img src="scale.png" alt="Scale" /></td>
<td><img src="network.png" alt="Network" /></td>
</tr>
<tr>
<td>Develop exact solution methods</td>
<td>Design a fast heuristic supported by error bounds</td>
<td>Illustrate method on case study of prevention for heart disease</td>
</tr>
<tr>
<td><img src="target.png" alt="Target" /></td>
<td><img src="graph.png" alt="Graph" /></td>
<td><img src="heart.png" alt="Heart" /></td>
</tr>
</tbody>
</table>

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
Interval Matrix (IM) Model

\[ Q_{t,\ell_t}^{\text{IM}} = \left\{ q_{t,\ell_t} \in \mathbb{R}^{1 | \mathcal{L}|} \right\} \left| \begin{array}{c} \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) = 1, \\
q^l_{t,\ell_t}(\ell_{t+1}) \leq q_{t,\ell_t}(\ell_{t+1}) \leq q^u_{t,\ell_t}(\ell_{t+1}), & \forall \ell_{t+1} \in \mathcal{L}, \\
0 \leq q_{t,\ell_t}(\ell_{t+1}) \leq 1, & \forall \ell_{t+1} \in \mathcal{L}, \\
\end{array} \right\} \]

- \( q^l_{t,\ell_t}(\ell_{t+1}) = \tilde{q}_{\ell_t}(\ell_{t+1}) - \left[ \chi^2_{|\mathcal{L}|-1,1-\alpha/(2|\mathcal{L}|)} \frac{\tilde{q}_{\ell_t}(\ell_{t+1})(1-\tilde{q}_{\ell_t}(\ell_{t+1}))}{N_{\ell_t}} \right]^{1/2} \)

- \( q^u_{t,\ell_t}(\ell_{t+1}) = \tilde{q}_{\ell_t}(\ell_{t+1}) + \left[ \chi^2_{|\mathcal{L}|-1,1-\alpha/(2|\mathcal{L}|)} \frac{\tilde{q}_{\ell_t}(\ell_{t+1})(1-\tilde{q}_{\ell_t}(\ell_{t+1}))}{N_{\ell_t}} \right]^{1/2} \)

- \( \tilde{q}_{\ell_t}(\ell_{t+1}) \): maximum likelihood estimate (MLE) of \( q_{\ell_t}(\ell_{t+1}) \)

- \( N_{\ell_t} \): total number of patients in state \( \ell_t \)
Robust Dynamic Programming (RDP) Algorithm (Nilim et al. 2005)

1: \( v_T^{\text{RMDP-TM}}(\ell_T, m_T) \leftarrow r_T(\ell_T, m_T), \forall \ell_T \in \mathcal{L}, m_T \in \mathcal{M} \)
2: \textbf{for} \( t = T - 1 \rightarrow 0, \ell_t \in \mathcal{L}, \text{and} \ m_t \in \mathcal{M} \) \textbf{do}
3: \textbf{for} \( \alpha_t \in A_t(\ell_t, m_t) \) \textbf{do}
4: \hspace{1em} Solve the inner problem and calculate \( w_t(\ell_t, m_t, \alpha_t) = r_t(\ell_t, m_t, \alpha_t) + \lambda [1 - p_t^E(\ell_t, m_t, \alpha_t)] \sigma_t^*(\ell_t, m_t, \alpha_t, \Gamma_t, \ell_t) \)
5: \textbf{end for}
6: \textbf{end for}

7: Update Value Function: \( v_t^{\text{RMDP-TM}}(\ell_t, m_t) \leftarrow \max_{\alpha_t \in A_t(\ell_t, m_t)} \{ w_t(\ell_t, m_t, \alpha_t) \} \)
8: Update Optimal Action Set: \( A_t^*(\ell_t, m_t) \leftarrow \arg\max_{\alpha_t \in A_t(\ell_t, m_t)} \{ w_t(\ell_t, m_t, \alpha_t) \} \)
Interval Model with Uncertainty Budget (IMUB)

\[ Q_{t,\ell_t}^{\text{IMUB}} (\Gamma_{t,\ell_t}) = \{ q_{t,\ell_t} \in \mathbb{R}^{|\mathcal{L}|}_+ \mid \begin{align*}
q_{t,\ell_t}(\ell_{t+1}) &= \hat{q}_{t,\ell} (\ell_{t+1}) - \delta_{t,\ell_t}(\ell_{t+1}) z_{t,\ell_t}(\ell_{t+1}) \bigl( \ell_{t+1} + 1 \bigr) + \delta_{t,\ell_t}(\ell_{t+1}) z_{t,\ell_t}(\ell_{t+1}) 
\sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t}(\ell_{t+1}) &= 1, \\
\sum_{\ell_{t+1} \in \mathcal{L}} \left( z_{t,\ell_t}(\ell_{t+1}) + z_{t,\ell_t}(\ell_{t+1}) \right) &\leq \Gamma_{t,\ell_t}, \\
z_{t,\ell_t}(\ell_{t+1}) \cdot z_{t,\ell_t}(\ell_{t+1}) &= 0, \\
0 &\leq z_{t,\ell_t}(\ell_{t+1}), z_{t,\ell_t}(\ell_{t+1}) \leq 1, \\
0 &\leq q_{t,\ell_t}(\ell_{t+1}) \leq 1, \\
\hat{q}_{t,\ell} (\ell_{t+1}) &\text{: nominal value (the point estimate)} \\
\delta_{t,\ell_t}(\ell_{t+1}) &\text{: maximum left(right)-hand side variation} \\
z_{t,\ell_t}(\ell_{t+1}) &\text{: proportion of variation} \\
\Gamma_{t,\ell_t} &\text{: the total uncertainty budget on row } \ell_t \text{ of } Q_t \end{align*} \]
Simulation-based Analyses on Life Years Gained from Selected Population-based Prevention Programs

- Steele et al. 2004 (colorectal cancer screening)
- **100% reduction in indoor tanning**
  - Ekwueme et al. 2014 (cervical cancer screening)
- **80% reduction in indoor tanning**
  - Wright et al. 1998 (cervical cancer screening)
  - Maciosek et al. 2010 (breast cancer screening)
  - Maciosek et al. 2010 (colorectal cancer screening)
- **50% reduction in indoor tanning**
  - Maciosek et al. 2010 (cholesterol screening)
  - Maciosek et al. 2010 (influenza immunization)
- **Under 18 age restriction on indoor tanning**
- **20% reduction in indoor tanning**
  - Maciosek et al. 2010 (hypertension screening)

---


Optimal Value Functions of the RMDP

\[
V_{t}^{\text{RMDP}}(s_t) = \begin{cases} 
\max_{\pi \in \Pi} \min_{\theta_t \in \Theta_t} \mathbb{E}_{s_t}^{\theta_t} \left[ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k(s_k)) + \lambda^{T-t} r_T(s_T) \right], \\
0, \quad & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\
\end{cases}
\]

- **Treatment policy:** \( \pi \)
- **All admissible treatment policies:** \( \Pi \)
- **Adversaries policy:** \( \theta_t \triangleq (Q_t, Q_{t+1}, \ldots, Q_{T-1}) \)
- **All admissible adversary policies:** \( \Theta_t \triangleq \prod_{\tau=t}^{T-1} Q_{\tau} \)