Active Surveillance of Prostate Cancer in the Context of Model Ambiguity

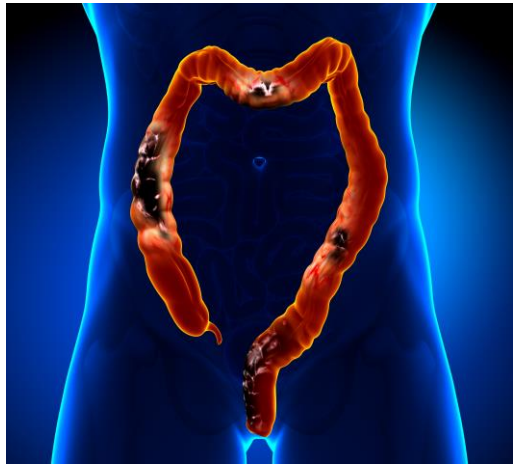
Brian Denton, Weiyu Li

HCSE 2023, Portugal

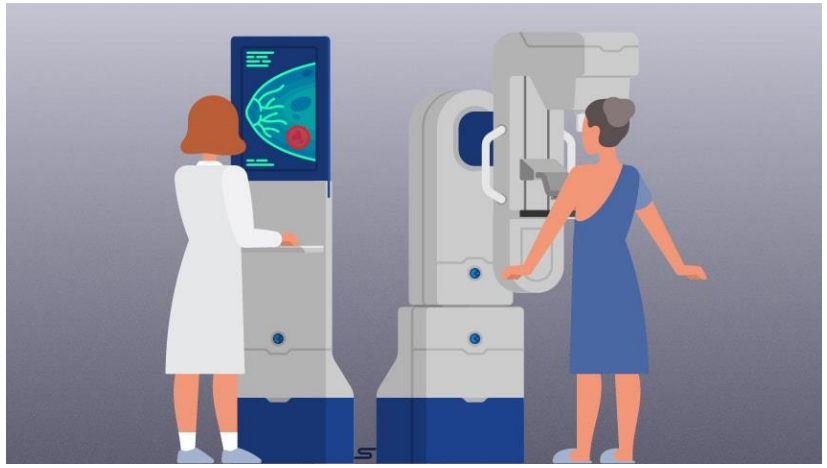
9/14/2023

Early detection of cancer

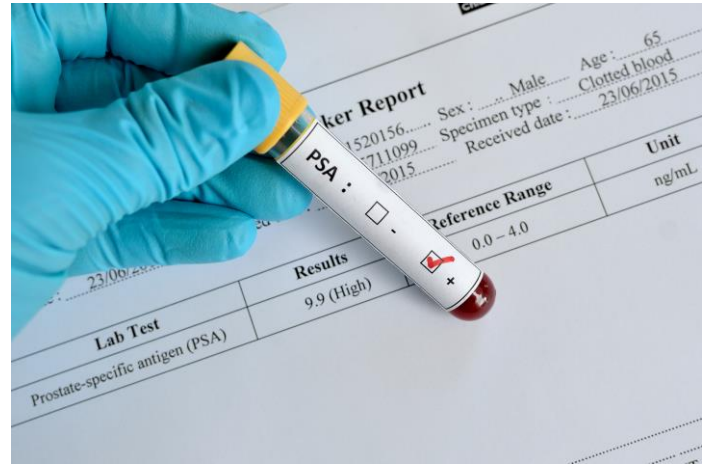
Early detection offers the opportunity for definitive treatment and cure, but many "low-risk" cancers are detected along the way



Colonoscopy Screening for Colorectal Cancer



Mammography Screening for Breast Cancer



Prostate-specific Antigen (PSA) Test for Prostate Cancer

Source: Harvard Health Publications + Michigan Health Lab

Cancer surveillance is challenging

- **Unobservable cancer state**, and stochastic transitions over time
- **Imperfect** diagnostic tests with false positive and false negative outcomes
- **Conflicting goals** to balance benefits and harms

Active surveillance of prostate cancer

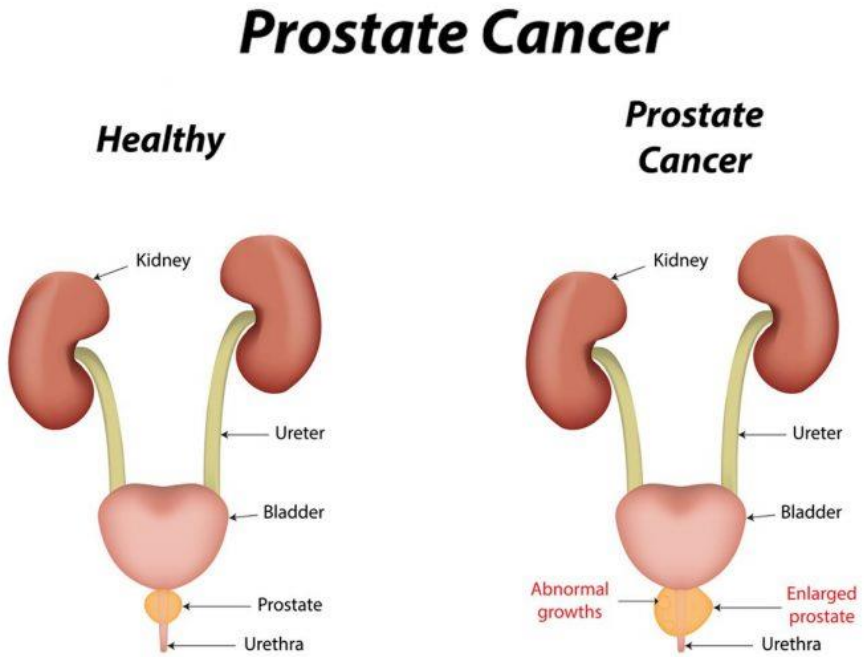
Active Surveillance (AS): monitoring “low-risk” patients with prostate specific antigen (PSA) tests and prostate biopsies

PSA Test

- Blood test with almost no direct harm
- Recommended every 6 months
- Very high rate of false positives and false negatives

Biopsy

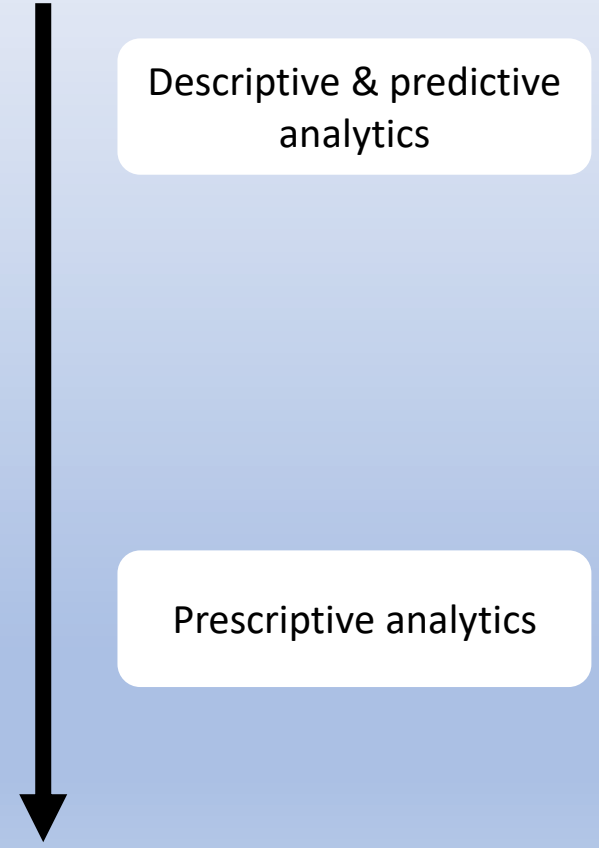
- Sampling the tissue with needles
- Painful and harmful
- Much more reliable than PSA test, but still imperfect



Source: joshya - Fotolia

Progression of Data-Driven Models for Active Surveillance of Prostate Cancer

- **Natural history modeling by HMM**
- **POMDP for individualized sequential decision-making**
- **Multi-POMDP model for addressing model ambiguity**



Urologists disagree on the best strategy for when to biopsy patients

Study	Recommended Biopsy Plan
Johns Hopkins (JH)	Annual biopsy
University of California, San Francisco (UCSF)	Biopsy 1 year after diagnosis, then every 1 to 2 years
Prostate Cancer Research International Active Surveillance (PRIAS) / University of Toronto (U of T)	Biopsy 1 year after diagnosis, then every 3 years

Data from surveillance studies

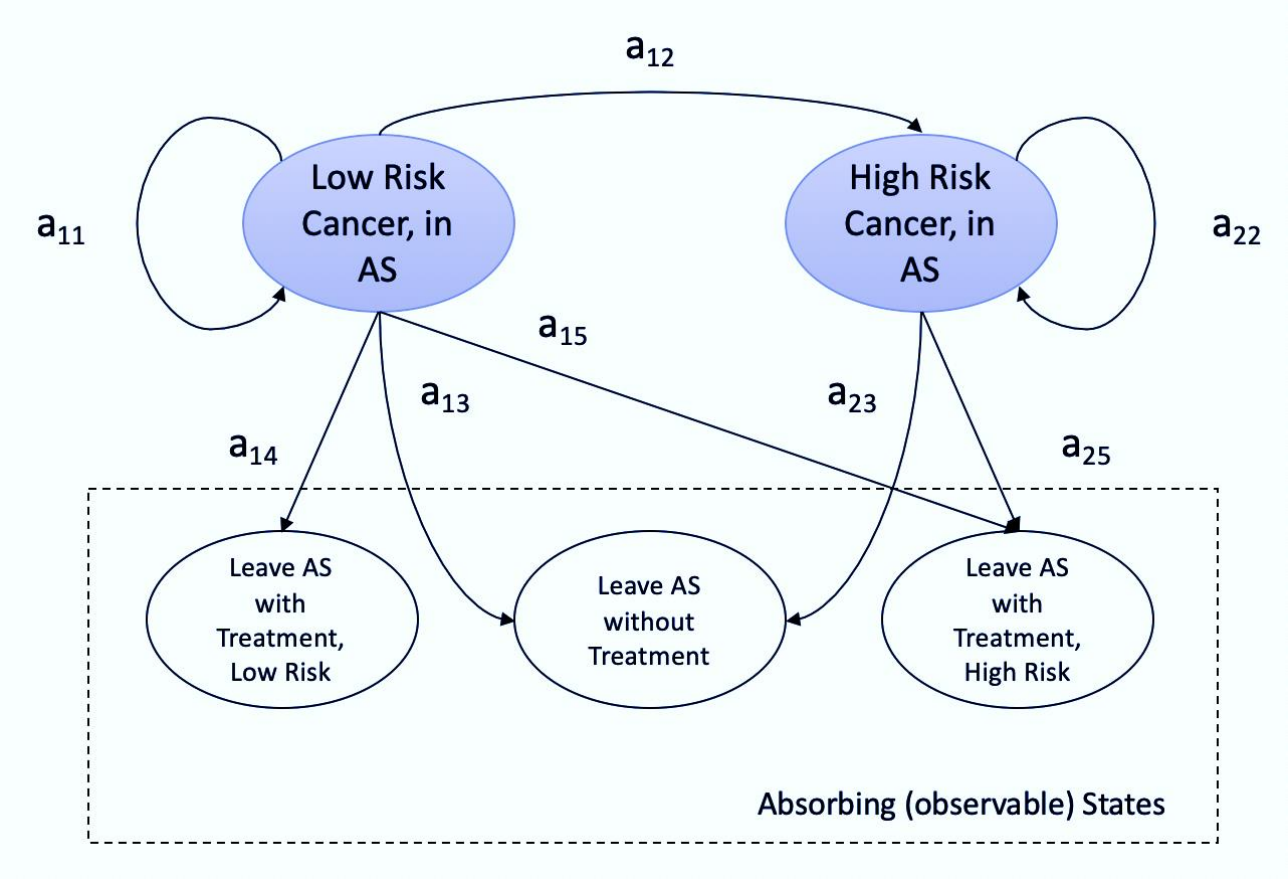
Movember Global Prostate Cancer Database:

- Includes 15,101 patients from 25 established AS cohorts worldwide
- Longitudinal observations of patients' clinical and demographic data

We picked 4 major AS cohorts (9,021 patients):

- Johns Hopkins (JH) Hospital
- University of California San Francisco (UCSF) Medical Center
- University of Toronto (U of T) medical center
- Prostate Cancer Research International Active Surveillance (PRIAS) project

Hidden Markov Model



- Time periods: annual time periods from the start to the end of the AS
- Initial distribution (at diagnosis)
  $$\phi = (\phi_1, 1 - \phi_1)$$
- Transition probability matrix:
  $$P = [\mathbb{P}(s_{t+1}|s_t)]$$
- Observation at time t:
  $$O_t = (X_t, Y_t)$$ (PSA, Biopsy)
- Observation probability
  $$F = [\mathbb{P}(o_t|s_t)]$$

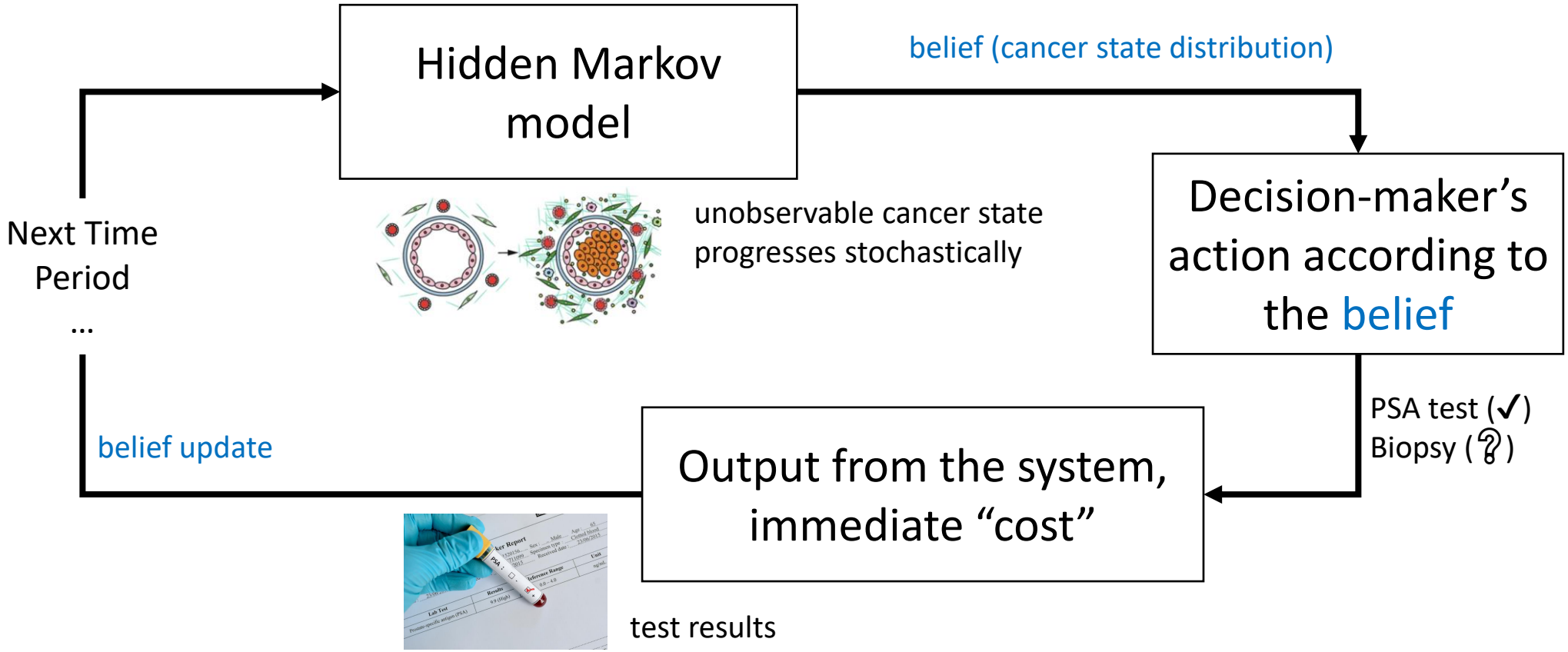
HMM model estimates for 4 different studies

Study	Mis-classification Error at Diagnosis	Annual Grade Progression Rate	Biopsy Sensitivity (True Pos.)	Biopsy Specificity (True Neg.)
JH	5.83% (0.74%)	6.91% (0.43%)	71.84% (0.53%)	99.72% (0.61%)
UCSF	8.09% (0.75%)	12.17% (0.85%)	74.31% (0.89%)	99.25% (0.80%)
U of T	7.74% (0.83%)	10.16% (0.79%)	79.49% (0.69%)	99.62% (0.75%)
PRIAS	6.53% (0.44%)	8.41% (0.73%)	76.14% (0.72%)	99.20% (0.95%)

Estimated parameters (and bootstrapped standard errors) by the HMMs in different studies.

Li, W. Denton, B.T., et al. “Comparison of biopsy under-sampling and annual progresssoin using hidden Markov models to learn fom prostate cancer active surveillance studies,” *Cancer Medicine*, 2020

POMDP for sequential decision-making



Goal: Minimize the weighted sum of harm from biopsies and delays in detection of high-risk cancer

Cancer Progression Figure by Alice Santi

Optimal value function & policy

The optimal policy $\pi^* = (\pi_0^*, \dots, \pi_T^*)$ achieves the maximum value function

$$V_t^{\pi^*}(b_t) \coloneqq \arg\max_{\pi} \mathbb{E}^{m,\pi}\left[\sum_{k=t}^{T} \gamma^{k-t} r^m(s_k, a_k, o_k) | b_t^m\right], \forall b_t, \forall t < T, \forall m$$

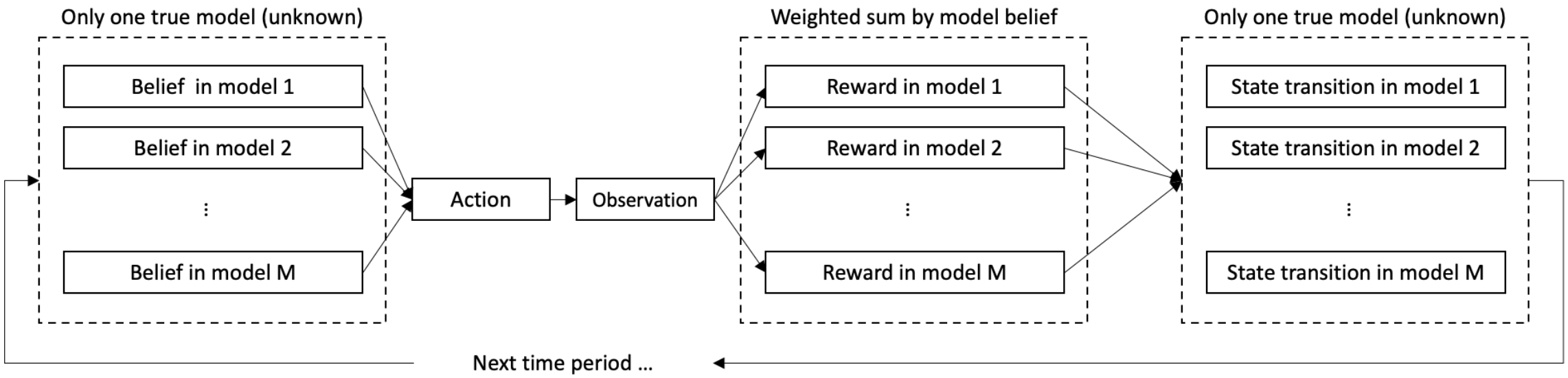
with the boundary condition

$$V_T^{\pi^*}(b_T) \coloneqq \arg\max_{\pi} r(b_T, \pi(b_T)), \forall b_T$$

where $\gamma \in [0,1]$ is a discount factor for future rewards.

Li W., Denton B.T., Morgan T.M.. "Optimizing Active Surveillance for Prostate Cancer Using Partially Observable Markov Decision Processes," *European Journal of Operational Research,* 2022.

Multi-Model POMDP



- True model is unknown
- A single policy is sought to minimize weighted costs

Results

- Objective: find the optimal timing of biopsies when the true model of each patient is unknown

- Cost function: weighted sum of
  1. number of biopsies to conduct
  2. delay time in detecting cancer progression

- non-informative prior model weights: $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1/4$

Comparison of the optimal policies when applying different models

Center	Minimum cost of the optimal policy (regret %)				
	JH model	UCSF model	U of T model	PRIAS model	MPOMDP model
JH	2.74 (0)	2.92 (6.50%)	3.84 (40.42%)	3.01 (9.89%)	2.87 (4.80%)
UCSF	2.54 (5.35%)	2.41 (0)	2.95 (22.45%)	2.68 (11.33%)	2.49 (3.33%)
U of T	2.65 (12.34%)	2.42 (2.39%)	2.36 (0)	2.77 (17.54%)	2.40 (1.72%)
PRIAS	2.59 (4.19%)	2.63 (5.54%)	3.11 (24.71%)	2.49 (0)	2.54 (2.03%)

Table 4.7: The optimal value (minimum cost) function in different AS studies when applying different policies.

- For each study center, the best model is always the “true” POMDP
- For each study center, the MPOMDP model dominates all “wrong” POMDPs
- The regret of a “wrong” model can be very high

Take Away Messages

- Be careful relying on a single study; you might pick the “wrong horse”
- Model ambiguity is much more important than statistical variation in model parameters
- See this working paper for more details: →

- Or contact me at btdenton@umich.edu

Acknowledgments

Students

Weiyu Li, PhD

Erkin Otles, PhD

Lauren Steimle, PhD

Zheng Zhang, PhD

Medical Collaborators

James E. Montie, MD

Todd Morgan, MD

Scott Tomlins, MD, PhD

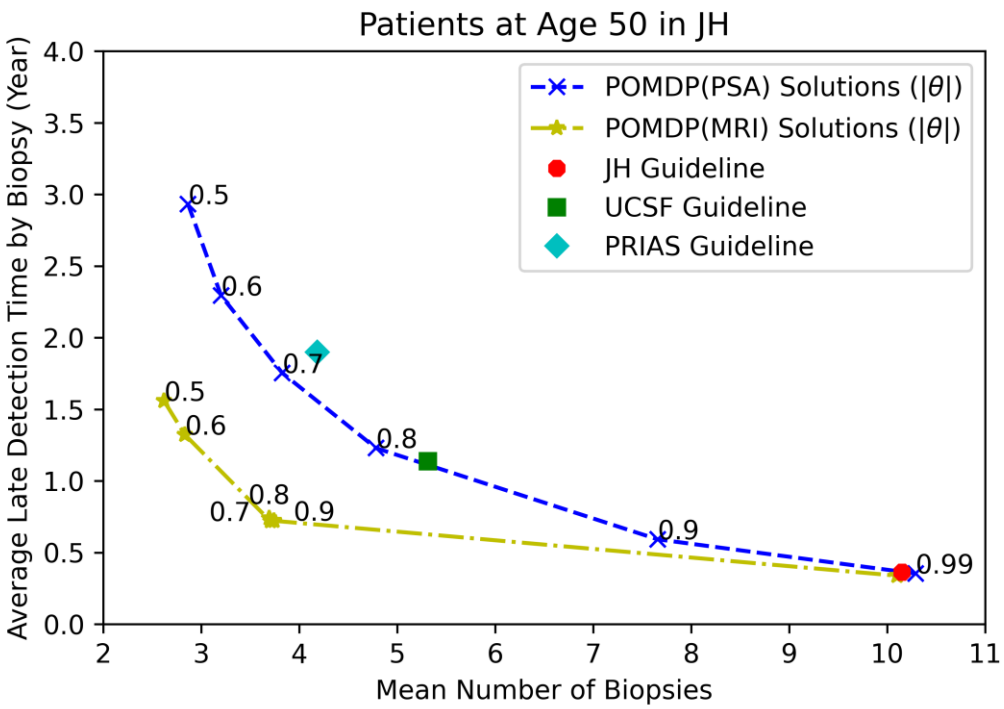
John Wei, MD

Backup slides

Bonus Example!

- Magnetic resonance imaging (MRI), is being adopted by some centers

- We conducted an experiment using MRI with the PSA test to show its potential impact.



The comparison between policies given by two AS-POMDP (PSA and MRI) models and current biopsy guidelines in the JH center.

Approximate methods: initialization with a lower bound

At each time t, recall the optimal value function

$$V_t(b_t) = \max_{\alpha \in \mathcal{A}} b_t \cdot \alpha, \forall b_t, \forall t$$

Instead of solving it for all $b_t \in [0,1]^M$, only sample a subset of belief points $B_t \subset [0,1]^M$, and solve $V_t(b_t)$ for $b_t \in B_t$, which gives the lower bound $\bar{V}_t, \ \forall t$.

The Optimal Value Function

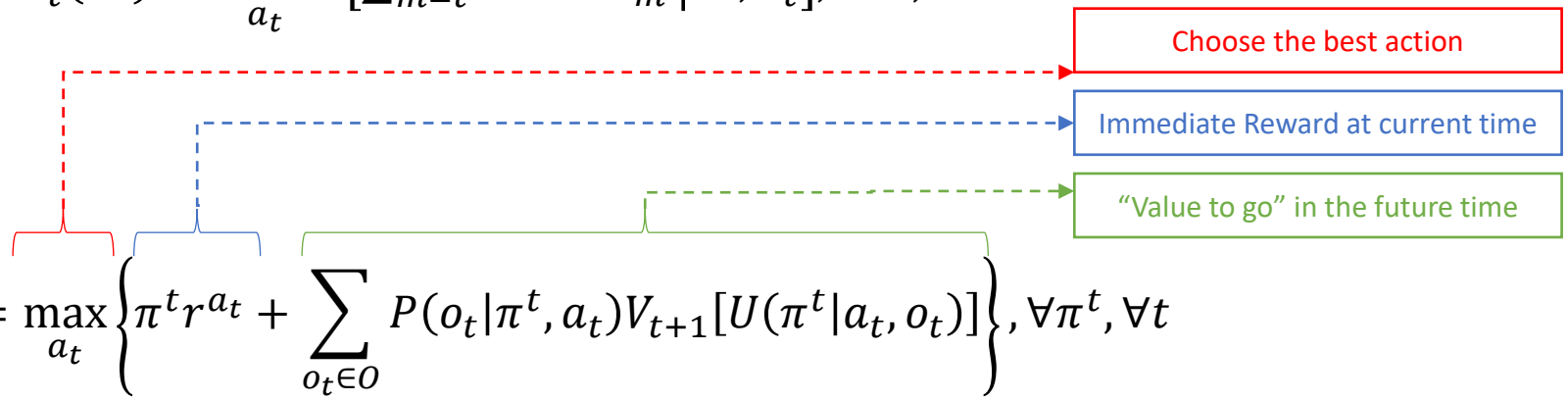
- Belief in high-risk cancer state

$$\pi^t \coloneqq P(S_t = \text{High Risk}), \forall t$$

- Optimal value function

$$V_t(\pi^t) \coloneqq \max_{a_t} \mathrm{E}\left[\sum_{m=t}^{\text{End}} \text{Reward}_m \,\middle|\, \pi^t, a_t\right], \forall \pi^t, \forall t$$

- Optimal equation



$$V_t(\pi^t) = \max_{a_t} \left\{ \pi^t r^{a_t} + \sum_{o_t \in O} P(o_t|\pi^t, a_t) V_{t+1}[U(\pi^t|a_t, o_t)] \right\}, \forall \pi^t, \forall t$$

where $U(\pi^t|a_t, o_t)$ is the updated belief calculated by the *Bayes formula*

- *Incremental pruning* algorithm (with approximation) to solve the POMDP (Cassandra et al. (1997))

Reference: Cassandra, A., Littman, M.L. and Zhang, N.L., 1997, August. Incremental pruning: A simple, fast, exact method for partially observable Markov decision processes. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence* (pp. 54-61). Morgan Kaufmann Publishers Inc..

The optimal individualized strategy

For each patient at time $t+1$, given the belief in high-risk cancer state $b_t, \forall t$:

1. Choose the optimal action given by:

value function

$$a_t^*(b^t) = \arg\max_{a_t} \mathrm{E}\left[\sum_t^{\mathrm{End}} \mathrm{Reward}_t \,\middle|\, b^t, a_t\right],$$

and observe the output o_t

2. Update his belief of being in high-risk cancer state: $b^{t+1} = P(S_{t+1} = \mathrm{high{-}risk} | b^t, a^t, o_t)$

… → previous belief b_t → choose action $a_t^*(b_t)$ → observe test results o_t → update belief b_{t+1} → …

Some related work in optimization under parameter ambiguity

- Robust optimization
  - e.g., robust DP (Iyengar 2005), Robust MDP (Nilim et al. 2005)
  - over conservative by optimizing the worst-case performance
- Multistage stochastic programming
  - e.g., Birge 2011 (textbook)
  - hard to tackle when the number stages increases
- Multi-model MDP
  - e.g., Steimle et al. 2018
  - only works for fully observable Markov processes

Weighted Value Problem

The objective is to find the optimal policy that performs well over all M POMDP models:

$$\max \sum_{m=1}^{M} \lambda_m \times \underbrace{\mathbb{E}[\text{future rewards in model } m | \text{curent beliefs, policy}]}_{\text{value function in POMDP models}}$$

There could be other risk measures:

- maximize the worst-case reward (robust optimization)
- minimize conditional value-at-risk (CVaR)
- probability measures

Baum-Welch Algorithm for Parameter Estimation

Given the observation sequences

$$O^{(1)} = \left(O_1^{(1)}, \dots, O_{T_1}^{(1)}\right), \dots, O^{(N)} = \left(O_1^{(N)}, \dots, O_{T_N}^{(N)}\right),$$

Baum-Welch algorithm, or equivalently *the EM (expectation-maximization) algorithm* estimates the model

$$\lambda = (A, B, \phi, c, \mu, \sigma)$$

that locally maximizes the likelihood function

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

Note: The estimated parameters quantify the most important factors in AS, which are the essential elements for decision making.

Reference: 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.