DPSCREEN: DYNAMIC PERSONALIZED SCREENING

KARTIK AHUJA[†], WILLIAM ZAME[†], AND MIHAELA VAN DER SCHAAR[*†], [†]UCLA, [*]UNIVERSITY OF OXFORD

MOTIVATION AND GOAL

Motivation

- Screening helps timely diagnosis (e.g., cancers)
- For every 1000 mammograms, 500 false positives, and 200 unnecessary biopsy
- Clinical guidelines are **not personalized!**
  - Yearly after the age between 45-54 in USA, Biennial after the age of 40 in Canada, Japan
- Disease models are complex, MISCAN-COLON
- Screening works use simplistic disease models

Goal

- Develop a **general framework** for screening
  - Personalized to patient feature and history
  - Applicable to many diseases

RELATED WORKS

Screening in Operations Research and Statistics

- Partially Observable Markov Decision Process **POMDP** (Ayer et.al.)
- **Bayesian Optimal Design** (Rizopoulos et.al.)
  - **Pros:** Principled search for optimal policies
  - **Cons: Markov/Semi-Markov, Stationary**

Screening in Medical Literature

- **Stochastic simulation** based (Frazie et.al.)
  - **Pros:** No assumptions on disease models
  - **Cons: Not personalized**, Compare a fixed set of policies chosen by experts

This work **No assumptions on disease models! Principled search for optimal policies!**

MODEL AND PROBLEM FORMULATION

- **Disease model:** Finite state stochastic process, one absorbing disease state, $\vec{S}$ is path of the stochastic process, $Pr(\vec{S} = \vec{s}|\boldsymbol{x})$ is probability of the path, $\boldsymbol{x} \in \mathcal{X}$ feature vector (age, gender, etc.), states are hidden
- **Diagnostic test:** $Z(t)$: test outcome at time t, $Pr(Z(t) = z|\vec{s}(t), \boldsymbol{x})$: probability of test outcome
- **External information:** $Y(t)$: observation by the patient, $Pr(Y(t) = y|\vec{s}(t), \boldsymbol{x})$: probability of the observation, if $Y(t) \in \mathcal{Y}$ patient goes to the clinician
- **Screening policy:** $\pi : \mathcal{H} \times \mathcal{X} \to \{1, ., T\}$: map from observation history, features to next arrival time
- **Costs:** **Cost of screening :** 1 (normalized),
Cost of delay in detection: $C(t_d - t_D; t_D)$, t_D/t_d is the time of incidence/detection
- **Optimal screening policy:** Minimizer of the weighted sum of the aggregate discounted screening costs and the delay costs, with weight w, discount factor δ, and the set of arrival times $\mathcal{T}_s$

$$\pi^* = \arg\min \overbrace{(1-w)E\Big[\sum_{t\in\mathcal{T}_s} \delta^t\Big]}^{\text{Screening costs}} + \overbrace{wE\Big[C(t_d - t_D; t_D)\Big]}^{\text{Delay costs}} \tag{1}$$

- **Challenges**: **Standard POMDP and POSMDP cannot be used**
  - Disease model is not Markov/Semi-Markov
  - Time between decision epochs depends on scheduled date, external information and the state path

STEPS TO DPSCREEN

Idea 1. **Define belief over entire path and not current state.** $\boldsymbol{b}(\vec{s}, l)$, $l = 1/0 \implies$ diagnosed/not

- **Bayesian belief update.**

$$\hat{b}(\vec{s}, l) = Pr(\vec{s}, l|\boldsymbol{b}, \underbrace{\boldsymbol{\tau}}_{\text{Scheduled visit}}, \underbrace{[y, z, \tilde{\tau}]}_{\text{Observation}}, x)$$

- **Bellman equation.**

$$V(\boldsymbol{b}, t) = \max_{\boldsymbol{\tau}} \Big[\sum_{\vec{s},z,l} \boldsymbol{b}(\vec{s}, l) Pr(z|\vec{s}, l)\tilde{C}(\vec{s}, l, t, z) + \sum_{z,y,\tilde{\tau}} Pr(z, y, \tilde{\tau}|\boldsymbol{\tau}, \boldsymbol{b}) V(\hat{\boldsymbol{b}}, t + \tilde{\tau})\Big]$$

$\tilde{C}$: cost for screening and cost of delay per epoch

- **Result.** **Value function is piecewise linear**

$$V(\boldsymbol{b}, t) = \max_{\boldsymbol{\alpha}\in\Gamma(t)} \boldsymbol{\alpha}^t \boldsymbol{b}$$

- Optimal $\boldsymbol{\alpha}$ vector decides the optimal policy
- If $\Gamma(t+1)$ is known, then a recursion derived from Bellman equation can determine $\Gamma(t)$

$$\Gamma(t) = \mathcal{R}^{exact}\big[\Gamma(t+1)\big] \tag{2}$$

$|\Gamma(t)|$ exponential in $t \implies \mathcal{R}^{exact}$ **intractable!**

- Point-based value iteration **PBVI**

$$\Gamma(t) = \mathcal{R}^{PBVI}\big[\Gamma(t+1)\big] \tag{3}$$

$|\boldsymbol{b}|$ large $\implies \mathcal{R}^{PBVI}$ **intractable!**

PROJECTED PBVI

Idea 2. Dimensionality reduction:

- Sample K i.i.d. paths from disease model
- Project the beliefs over the sampled subset

Idea 3. Basis set of policies:

- Random exploration
- Clinical guidelines/policies from existing works

Intuition: Ensure better performance than basis

- **Belief set construction**:
  - Use stochastic simulations to generate outcomes using basis set
  - Construct belief set $\bar{B}$: Bayesian update conditional on the sampled subset

- **Projected PBVI**: One α vector per point in $\bar{B}$

$$\hat{\Gamma}(t) = \mathcal{R}^{PPBVI}_{\bar{B}}\big[\hat{\Gamma}(t+1)\big] \tag{4}$$

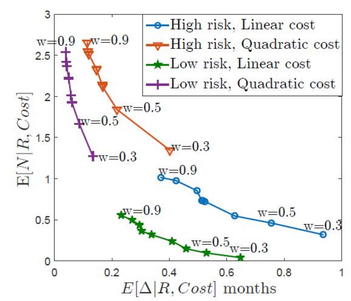
$$\hat{V}(\boldsymbol{b}, t) = \max_{\boldsymbol{\alpha}\in\hat{\Gamma}(t)} \boldsymbol{\alpha}^t \boldsymbol{b} \tag{5}$$

- **Approximation error**: $|\hat{V}(\boldsymbol{b}, t) - V(\boldsymbol{b}, t)| \leq \Omega(\bar{B})$, $\Omega(\bar{B})$ worst case sampling density
- **Computational complexity:** $\mathcal{O}(T^3|\bar{B}|^2 K|\mathcal{Y}||\mathcal{Z}|)$
- **Robustness:** Errors in model estimation?
Optimal policy is locally constant over the space of models, i.e small errors don't matter!
- Optimize hyperparameters, i..e, sampling policy and the basis set, is future work

ILLUSTRATIVE EXPERIMENTS

- **Dataset**: Deidentified breast cancer dataset of 45,000 women. At least one mammogram/woman.
- **Features:** Number of family members with breast cancer history, age, bmi, menopause
- **Disease model**: **Pre-incidence.** Two state Markov model for the onset of breast cancer.
Post-incidence. Universal tumor growth law. Tumor growth $\implies$ lumps develop $\implies$ self-arrivals
- **Model Estimation:** Parameters of the disease model estimated using standard MCMC methods.
- **Benchmarks:** Annual and Biennial screening
- **Metrics:** $E[\Delta|R]$, $E[\Delta|R, D]$ $E[N|R]$: Expected delay (months) given risk, Expected delay given risk and disease, expected number of screenings given risk,

Risk	Metrics	DPSCREEN with self-exam	DPSCREEN w/o self-exam	Annual
Low	$E[N\|R]$	0.32,	0.55	1
	$E[\Delta\|R]$	0.23	0.23	0.24
	$E[\Delta\|R, D]$	9.2	9.2	9.4
High	$E[N\|R]$	0.43	0.72	1
	$E[\Delta\|R]$	0.50	0.52	0.52
	$E[\Delta\|R, D]$	6.7	7.07	7.07



CONCLUSION

- Developed a general framework for screening
- Extended PBVI to address the challenges imposed by screening
- Gains $> 30\%$ on breast cancer screening dataset
- Potential impact beyond screening in stopping time problems