Multi-Agent Reinforcement Learning for Adaptive Physical Activity Interventions and Survey-Based State Tracking

Xueqing Liu

Harvard University

February 3, 2026

Outline

ADAPTS study

- Benefits of physical activity: reducing the risk of chronic diseases like heart disease and diabetes, cognitive functions, emotional well-being...
- Mobile health (mHealth) interventions: high accessibility, low cost, support for user autonomy, real-time monitoring, personalization
- **Goal:** Develop an mHealth intervention that are applicable to cardiac-rehabilitation (CR) patients who can safely exercise unsupervised

Affect-based Walking Suggestions

- "Push" notifications: Walking suggestion paired with cute animal GIFs
- Positive Valence: Designed to induce smiling and create a pleasant user experience
- **Goal:** Create associations between positive emotions and the thought of walking

Next: Using affective association as the RL reward function.

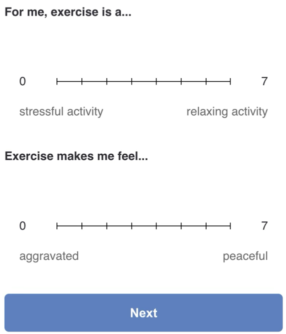
Start your day strong by taking care of yourself. Le
go, woof!

Reward/State measured via Weekly Check-in Survey

- **Affective valuation:** *1 item*; gut reaction when thinking about being active

Reward/State measured via Weekly Check-in Survey

- **Affective valuation:** *1 item*; gut reaction when thinking about being active
- **Core affective experiences (CAE):** *12 items*; how participants felt about being active during the week



Reward/State measured via Weekly Check-in Survey

- **Affective valuation:** *1 item*; gut reaction when thinking about being active
- **Core affective experiences (CAE):** *12 items*; how participants felt about being active during the week
- **Perceived utility:** *2 items*; whether the app is pleasant/helpful to use

Reward/State measured via Weekly Check-in Survey

- **Affective valuation:** *1 item*; gut reaction when thinking about being active
- **Core affective experiences (CAE):** *12 items*; how participants felt about being active during the week
- **Perceived utility:** *2 items*; whether the app is pleasant/helpful to use

Minimize Participant Burden: Utilize a single-item question as a proxy emission for the user's weekly CAE.

RL problem formulation

- The RL algorithm runs independently for each participant in the ADAPTS RCT.
- Decision points are indexed by (w, k):
  - $w \in \{1, \ldots, W\}$ denotes the **week**.
  - $k \in \{0, 1, 2\}$ denotes the **decision index** within the week.
- **Objective:** The agent maximizes the expected cumulative weekly affective association (affective valuation + CAE) over the study duration:

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{w=1}^{W} Y_w \right]$$

.

Action Spaces

- **Query Agent (Weekly):** Let $I_{w,0} \in \{0, 1\}$ denote the intervention at the end of week $w - 1$.
  - $I_{w,0} = 1$: Request **Full Survey** (Higher burden, accurate state/reward).
  - $I_{w,0} = 0$: Request **Short Survey** (Lower burden, proxy emission).
- **Walking Agent (Twice-Daily):** Let $A_{w,k} \in \{0, 1\}$ be the action at decision point (w, k).
  - $A_{w,k} = 1$: Deliver a **Walking Suggestion** (with affective GIF).
  - $A_{w,k} = 0$: No intervention.

Trade-offs of the Survey Action ($I_{w,0}$)

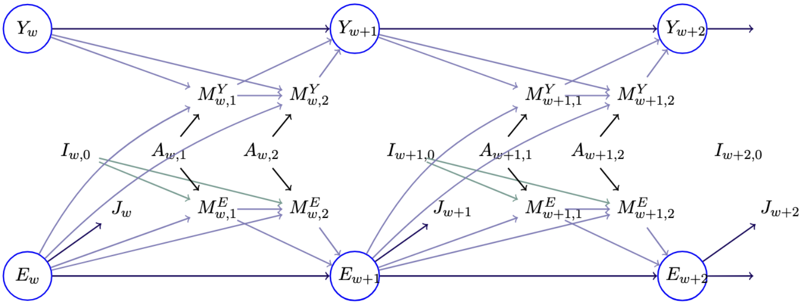
Benefits of Full Survey ($I_{w,0} = 1$)

- **Information Gain:** Resolves uncertainty in state estimation.
- **Improved Control:** Leverages richer state data to personalize the type of walking suggestions.
- **Enhanced Learning:** Provides the ground-truth reward (Y_w) needed to update the policy accurately.

Costs of Full Survey ($I_{w,0} = 1$)

- **Response Quality:** High burden can degrade the quality/honesty of affective valuation and CAE responses.
- **Intervention Fatigue:** Excessive querying may reduce responsiveness to future walking suggestions.

Causal DAG



- Emissions O^Y_w and O^E_w, along with real-time contexts $C_{w,k}$, are omitted
- Vector $M^Y_{w,k}$ ($M^E_{w,k}$) contains mediators on the causal pathway between $A_{w,k}$ and Y_{w+1} (E_{w+1})
- E_w indicates perceived utility
- J_w indicates weekly survey completion

Simplify the complicated environment to let the RL algorithm learn faster! Not an assumption on the real environment.

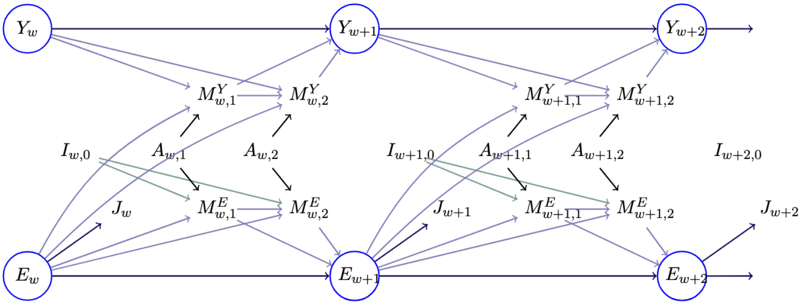
State Construction Based on DAG

Based on Gao et al. (2025a), we construct the “best” state as a *dynamical Bayesian sufficient statistic*:

- The state of $I_{w,0}$ is $S^{I}_{w,0} = [Y_w, E_w, C_{w,0}]$.
- The state of $A_{w,1}$ is $S^{A}_{w,1} = [Y_w, E_w, C_{w,1}]$.
- The state of $A_{w,2}$ is $S^{A}_{w,2} = [Y_w, E_w, M^{Y}_{w,1}, M^{E}_{w,1}, C_{w,2}]$.

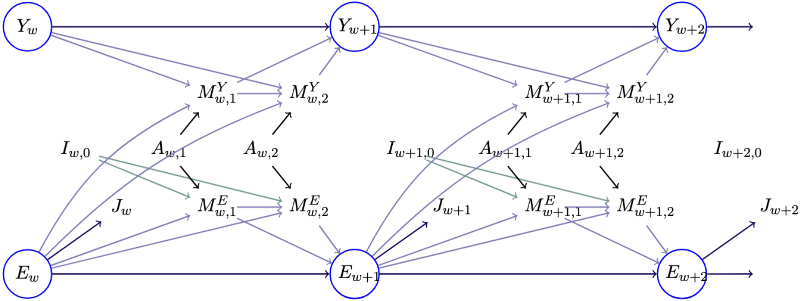
so that the state transition is **Markovian** within and across weeks.

Periodic POMDP formulation (weekly cycle)



- Stationary across week; nonstationary within week
- We can focus on Markov, stationary, and stochastic policies $\pi = \{\pi_{0:K}\}$, where $\pi_k := \mathcal{S}_k \mapsto \Delta(\mathcal{A}_k)$ (*Gaoet al.*, 2025*b*)

Belief updates for latent weekly state Y_w and E_w



Two-stage belief update within week $w+1$:

1. **Pre-query $I_{w+1,0}$ update:** incorporate information available *prior to* (Y_{w+1}, E_{w+1})
2. **Post-query $I_{w+1,0}$ update:** incorporate the emissions

$$O_w^E = \begin{cases} E_w, & \text{if } J_w = 1, \\ \emptyset, & \text{if } J_w = 0. \end{cases} \qquad O_w^Y = \begin{cases} Y_w, & \text{if } I_w = 1 \text{ and } J_w = 1, \\ \widetilde{Y}_w, & \text{if } I_w = 0 \text{ and } J_w = 1, \\ \emptyset, & \text{if } J_w = 0. \end{cases}$$

Y_w: mean score from the full 13-item affective valuation + CAE battery; $\widetilde{Y}_w$: mean score from the 2-item short battery.

Adaptation of RLSVI/randomized value function

Why randomized least square value iteration (RLSVI) (Osband et al., 2016)?

1. Stochastic, we need explicit randomization probability to enable after study data analysis
2. Model-free, robust to misspecified assumptions of the causal DAG

A Single-Agent Algorithm

For each week w:

- **Backward Learning:** Fit Bayesian linear regression $\boldsymbol{y}_{w,k} = [y_{1:(w-1),k}]^\top$ on $\boldsymbol{X}_{w,k} = [X_{1:(w-1),k}]^\top$ for $k = 2, 1, 0$, where

$$X_{w',k} = \begin{cases} \phi_0(S^I_{w',0}, I_{w',0}) & k = 0 \\ \phi_k(S^A_{w',k}, A_{w',k}) & k \in \{1, 2\} \end{cases}$$

$$y_{w',k} = \begin{cases} \max_{a \in \mathcal{A}} \phi_1(S^A_{w',1}, a)^\top \widetilde{\beta}_{w,1} & k = 0 \\ \max_{a \in \mathcal{A}} \phi_2(S^A_{w',2}, a)^\top \widetilde{\beta}_{w,2} & k = 1 \\ Y_{w'} + \max_{i \in \mathcal{I}} \phi_0(S^A_{w'+1,0}, i)^\top \widetilde{\beta}_{w-1,0} & k = 2 \end{cases}$$

  Draw $\widetilde{\beta}_{w,k} \sim N(\mu_{w,k}, \Sigma_{w,k})$.

- **Forward action selection:** Calculate the action sampling probability $\pi_{w,k}(1 \mid H_{w,k}) = \Pr(\phi_k(S_{w,k}, 1)^\top \widetilde{\beta}_{w,k} > \phi_k(S_{w,k}, 0)^\top \widetilde{\beta}_{w,k})$ and deliver action 1 with probability $\pi_{w,k}(1 \mid H_{w,k})$ for $k = 0, 1, 2$

A Dual-Agent Algorithm

- **Issue 1:** Weekly survey responses vary in **quality** (burden $\rightarrow$ low-quality reports), but the single-agent update treats them as equally informative.
- **Issue 2:** The benefit of choosing $I_{w,0}$ (Full vs. Short) is mostly **indirect**—it affects the **post-query belief** and downstream actions—so its value-of-information is a **delayed / weak** signal
- **Separate Query Agent:** *Reward design*

Future work

- Deep exploration with explicit randomization (Osband et al., 2019)
- Optimize reward signal design for both query and walking agents (Ibrahim et al., 2024)
- Develop more stable methods for real-time belief propagation
- Establish regret bounds for two-agent RLSVI
- Empirical evaluation on the ADAPTS simulation testbed

Thanks for listening

Joint work with Daiqi Gao, Soo Ji Serisse Choi, Steven De La Torre, Predrag Klasnja, and Susan A. Murphy

Xueqing Liu
xueqing_liu@fas.harvard.edu
https://xueqingliuu.github.io/

References

Gao, D., Lai, H.-Y., Klasnja, P., and Murphy, S. (2025a). Harnessing causality in reinforcement learning with bagged decision times. In *International Conference on Artificial Intelligence and Statistics*, pages 658–666. PMLR.

Gao, D., Xu, Z., Rawashdeh, A., Klasnja, P., and Murphy, S. A. (2025b). Active measuring in reinforcement learning with delayed negative effects. *arXiv preprint arXiv:2510.14315*.

Ibrahim, S., Mostafa, M., Jnadi, A., Salloum, H., and Osinenko, P. (2024). Comprehensive overview of reward engineering and shaping in advancing reinforcement learning applications. *IEEE Access*.

Osband, I., Van Roy, B., Russo, D. J., and Wen, Z. (2019). Deep exploration via randomized value functions. *Journal of Machine Learning Research*, 20(124):1–62.

Osband, I., Van Roy, B., and Wen, Z. (2016). Generalization and exploration via randomized value functions. In *International Conference on Machine Learning*, pages 2377–2386. PMLR.

iMSi Institute for Mathematical and Statistical Innovation

Curbing the Opioid Crisis:
Optimal Dynamic Policies for Preventive and Mitigating Interventions

Sina Ansari
Driehaus College of Business, DePaul University

In collaboration with:
Shakiba Enayati, College of Business Administration, University of Missouri St. Louis
Raha Akhavan-Tabatabaei, Sabanci Business School, Sabanci University
Julie M. Kapp, College of Health Sciences, University of Missouri

February 2026

Opioid Crisis Background

What is Opioid? Natural or synthetic chemicals that reduce the intensity of pain signals.

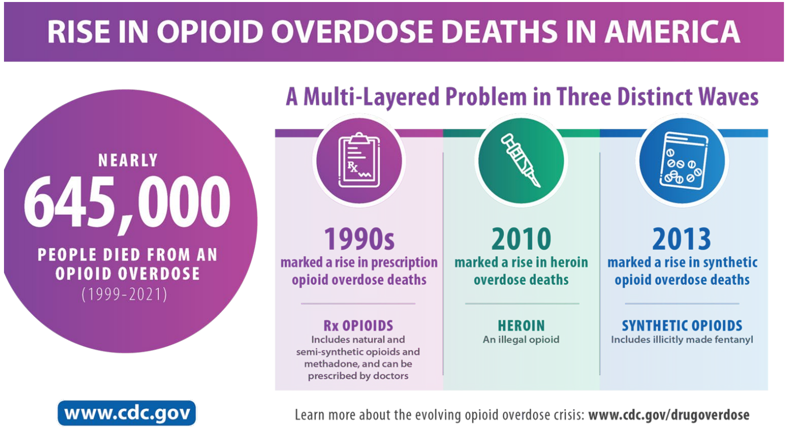
How can it be obtained? Legally by prescription and illegally through pill diversion or illicit networks.

Opioid Misuse? Use of prescription drugs without a prescription or not as directed by a doctor => causes Opioid Use Disorder (OUD) => excessive use may result in overdose

When did it become a crisis?

- HHS declared a **public health emergency** in 2017.
- In 2022, nearly 108,000 people died from drug overdoses (296 overdoses each day).

Economic Burden of Prescription Misuse?
$78.5 billion a year (2016)



How to respond to the opioid crisis?

Limit the Supply of Prescriptions (Preventive Interventions)

Decreasing the number of individuals who misuse opioids & reducing fatal overdose

How?

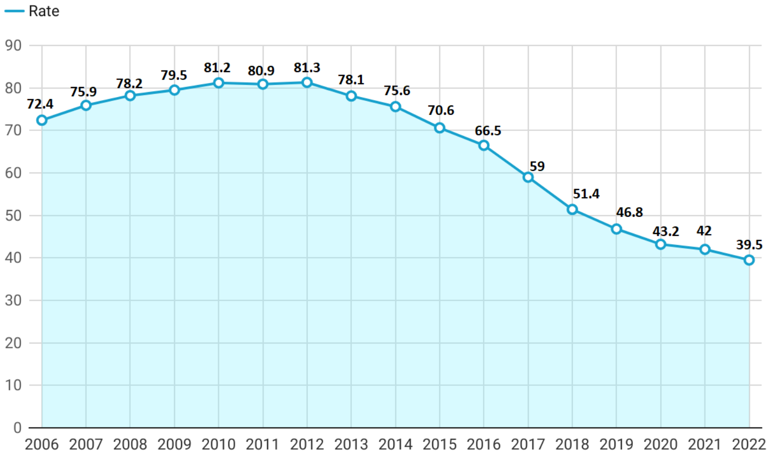
- Patient-level interventions (education, safe disposal, not sharing meds)
  - e.g., Albert et al. (2011), Fass et al. (2011), Steward et al. (2015).
- Prescriber-level interventions (pain contracts, risk assessment, evidence-based prescribing)
  - e.g., Baehren et al. (2010), Kolodny et al. (2015), Haffajee et al. (2015), Chang et al. (2017), **Andereck et al. (2019)**.
- Regulator-level interventions (research, public education, increase access, data)
  - e.g., Franklin et al. (2015), Rutkow et al. (2016).

Question: Is limiting the prescription supply enough?

Prescribing Rates vs. Overdose Deaths

Rate of prescription opioids dispensed in the United States (per 100 persons)

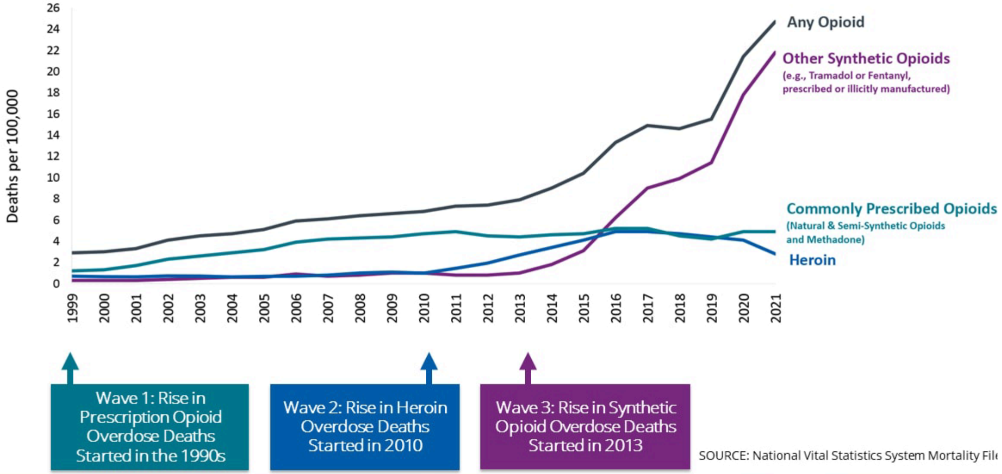
Prescription Rate in Numbers



(Rate in Numbers)

Source: Market.us Media

Three Waves of Opioid Overdose Deaths



Primary Preventive Interventions are not enough!

Interventions to Control the Crisis

A "**Wise**" selection of interventions to invest in is very important!

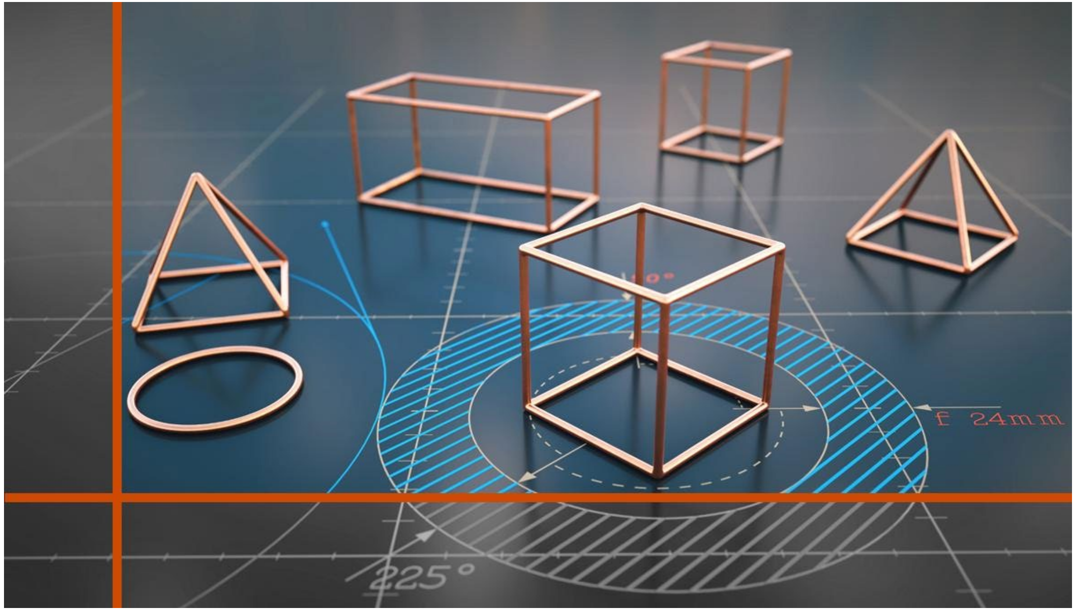
Preventive: limiting prescription supply,
e.g., prescription drug monitoring programs, drug rescheduling, drug reformulation, etc.

Mitigating: reducing overdose deaths,
e.g., increase availability of Naloxone, needle exchange, medication assisted treatment, etc.

Trade-off: intervention effect sizes and health costs (misuse and overdose)

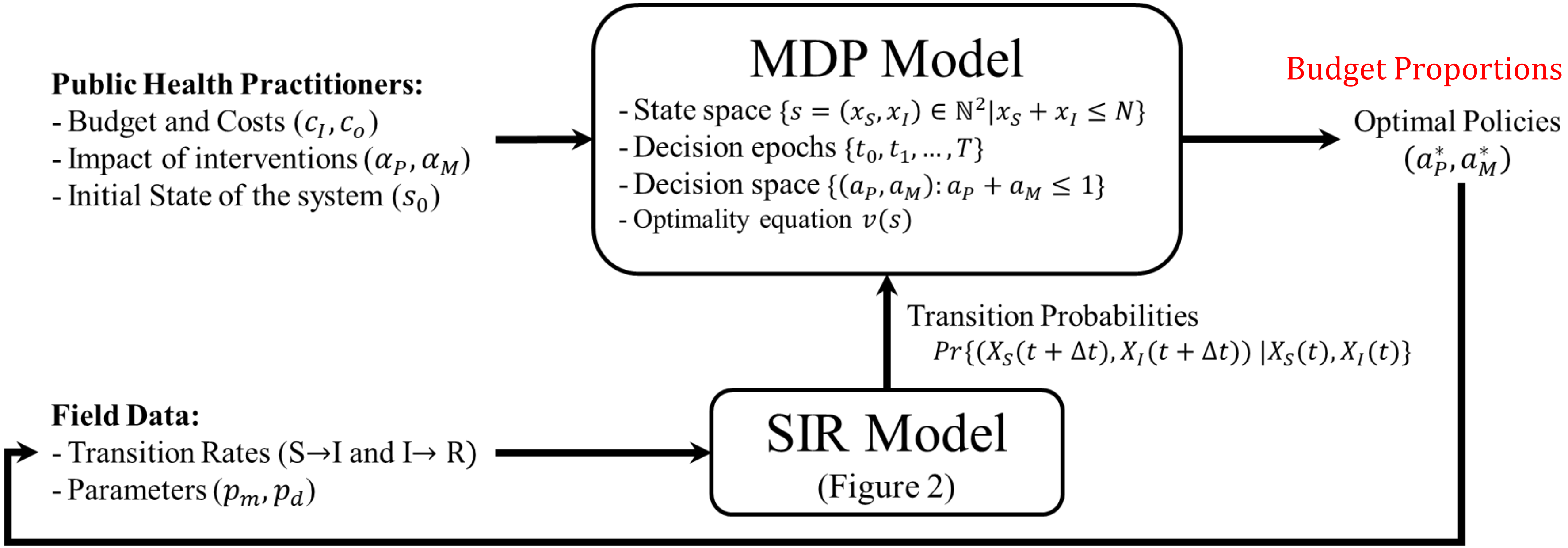
Goal: To develop a state-dependent **Decision Support Tool** to recommend the optimal **budget** allocation.

Model



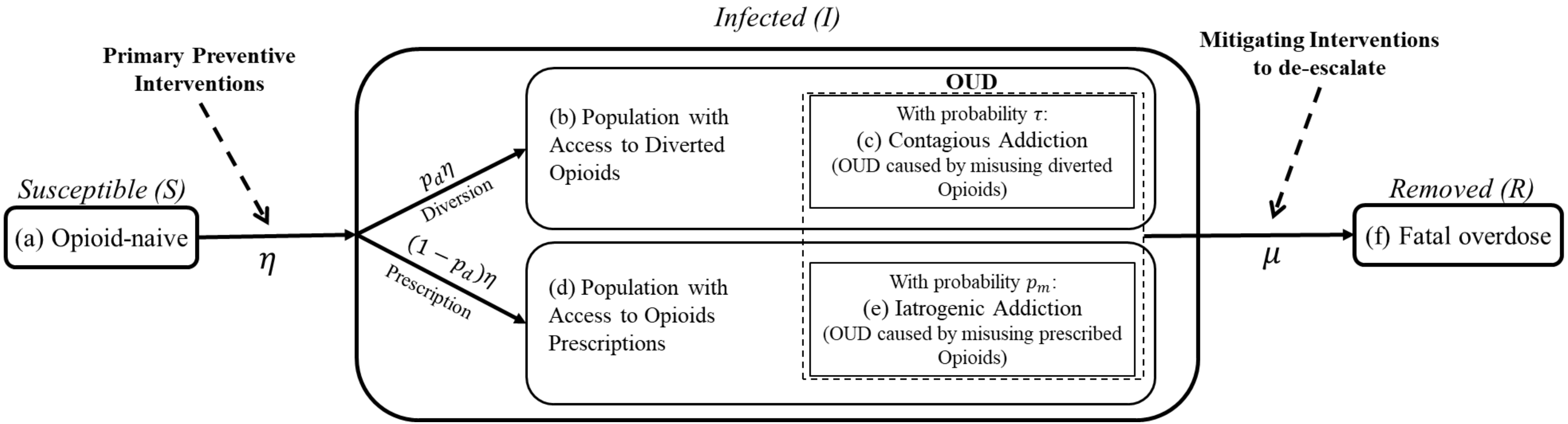
Dynamics of the Decision Framework

We model the crisis dynamics as a discrete-time Markov process.



SIR Compartmental Model

We replicate access to prescription opioids and the ensuing pill diversion dynamic as a transmissible condition.



Optimality Equation

The optimality equation to minimize expected total discounted cost:

$$v(s) = \min_{a \in A}\{r(s) + \gamma \sum_{s'} P(s'|s, \mathbf{a})v(s')\}, \text{for } s = (X_s(t), X_I(t)) \in \text{S and } a_p + a_M \leq 1,$$

where discount factor $\gamma \in [0,1]$ and

$$r(s) = \underbrace{c_I(p_d\tau + (1 - p_d)p_m)X_I(t)}_{Infection\ Cost} + \underbrace{c_o(N - X_s(t) - X_I(t))}_{Fatal\ Overdose\ Cost}$$

p_d: Probability of an individual transitioning from S to I through pill diversion
p_m: Probability of misuse and subsequent development of OUD
τ: Probability of contagious addiction for an opioid-naive patient who receives diverted opioids

Solving the MDP

Approximating the Markov chain using the proportion of populations in each compartment and Policy Improvement Algorithm.

Algorithm 1: To calculate the transition probability $Pr\{(\Theta_S(t+\Delta t), \Theta_I(t+\Delta t)) = (\theta_S, \theta_I) | \Theta_S(t), \Theta_I(t)\}$

Result: $Pr\{(\Theta_S(t+\Delta t), \Theta_I(t+\Delta t)) = (\theta^S_{j_1}, \theta^I_{j_2}) | (\Theta_S(t), \Theta_I(t)) = (\theta^S_{i_1}, \theta^I_{i_2})\}$

for a given (i_1, i_2) and (j_1, j_2)

1. Form $\mathcal{S}_{X(t)} = \{(x_S, x_I) \in \mathbb{N}^2 | 0 \leq x_S \leq \lfloor N\theta^S_{i_1} \rfloor, x_S + x_I \leq N\}$
2. Form $\mathcal{S}_{\Theta(t)} = \{(x_S, x_I) \in \mathbb{N}^2 | \lceil Nb^S_{j_1-1} \rceil \leq x_S \leq \lfloor Nb^S_{j_1} \rfloor, \lceil Nb^I_{j_2-1} \rceil \leq x_I \leq \lfloor Nb^I_{j_2} \rfloor\}$

if $(\theta^S_{i_1} + \theta^I_{i_2}) - (\theta^S_{j_1} + \theta^I_{j_2}) \geq 0$ *and* $\theta^S_{i_1} - \theta^S_{j_1} \geq 0$ *and* $(\theta^S_{j_1} + \theta^I_{j_2}) - (\theta^S_{i_1} + (1-p_d)(1-p_m)\theta^I_{i_2}) \geq 0$ **then**

$$\begin{aligned} P((i_1, i_2), (j_1, j_2)) &= Pr\{(\theta^S_{j_1}, \theta^I_{j_2}) | (\theta^S_{i_1}, \theta^I_{i_2})\} \\ &= \sum_{(x_S, x_I) \in \mathcal{S}_{\Theta(t)}} Pr\{(x_S, x_I) | (\lfloor N\theta^S_{i_1} \rfloor, \lfloor N\theta^I_{i_2} \rfloor)\} \end{aligned}$$

;

else

$P((i_1, i_2), (j_1, j_2)) = Pr\{(\theta^S_{j_1}, \theta^I_{j_2}) | (\theta^S_{i_1}, \theta^I_{i_2})\} = 0$

if $\mathcal{S}_{X(t)} \cap \mathcal{S}_{\Theta(t)} \neq \emptyset$ **then**

$$\begin{aligned} P((i_1, i_2), (i_1, i_2)) &= P((i_1, i_2), (i_1, i_2)) + Pr\{(\theta^S_{i_1}, \theta^I_{i_2}) | (\theta^S_{i_1}, \theta^I_{i_2})\} \\ &= P((i_1, i_2), (i_1, i_2)) \\ &\quad + \sum_{(x_S, x_I) \in \mathcal{S}_{X(t)} \cap \mathcal{S}_{\Theta(t)}} Pr\{(x_S, x_I) | (\lfloor N\theta^S_{i_1} \rfloor, \lfloor N\theta^I_{i_2} \rfloor)\} \end{aligned}$$

end

end

Figure 11 Algorithm for calculating the transition probabilities for Markov chain $\{(\Theta_S(t), \Theta_I(t)), t \geq 0\}$

Summary of the Policy Improvement Algorithm with Discounted Costs:

Initialization: Select an arbitrary policy R_1. Set $k = 1$.

Step 1 (Value determination at iteration k): For policy R_k, solve the following system of equations:

$$V_{\mathbf{s}}(R_k) = r(\mathbf{s}) + \gamma \sum_{\mathbf{s}' \in \mathcal{S}} Pr(\mathbf{s}' | \mathbf{s}, \mathbf{a}) V_{\mathbf{s}'}(R_{k-1}), \ \mathbf{s} \in \mathcal{S},$$

for all the unknown values of $V_{\mathbf{s}}(R_k)$, for $\mathbf{s} \in \mathcal{S}$.

Step 2 (Policy Improvement at iteration k): Using the values of the $V_{\mathbf{s}}(R_k)$, for $\mathbf{s} \in \mathcal{S}$, find the improved policy R_{k+1} such that, for each state $\mathbf{s} \in \mathcal{S}$,

$$\min_{\mathbf{a} \in \mathcal{A}} r(\mathbf{s}) + \gamma \sum_{\mathbf{s}' \in \mathcal{S}} Pr(\mathbf{s}' | \mathbf{s}, \mathbf{a}) V_{\mathbf{s}'}(R_k),$$

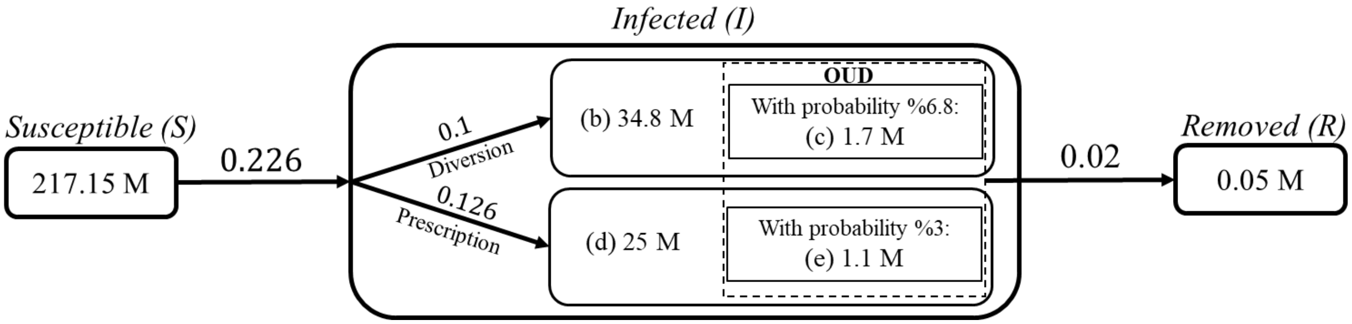
and set $\mathbf{a}_{\mathbf{s}}(R_{k+1}) = \mathbf{a}$, the decision that minimizes the expected discounted cost under policy R_k. In case of ties (i.e., two decisions that lead to the same expected discounted cost), we choose the policy with higher health benefits (e.g., we choose "both" policy over individual policies and "preventive policy only" over "mitigating policy only."

Optimality Test: The policy R_{k+1} is optimal if it is identical to policy R_k. If this condition holds, stop. Otherwise, reset $k = k + 1$ and repeat steps 1 and 2.

Numerical Results

Data Sources and Parameter Estimation

Using the data from the 2020 U.S. population census and opioid-related statistics

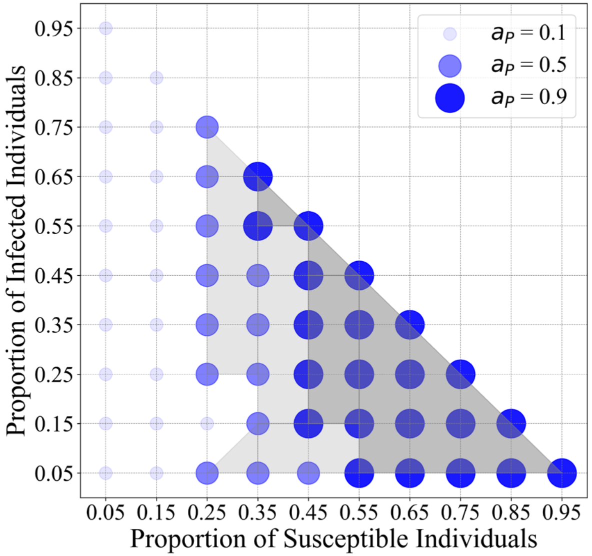


- Calibrated our MDP model for a population size normalized to 1,000 people.
- Assumed an annual decision epoch and discount factor of 0.9.
- Chose $d_S = d_I = 10$ in our MDP approximation procedure.

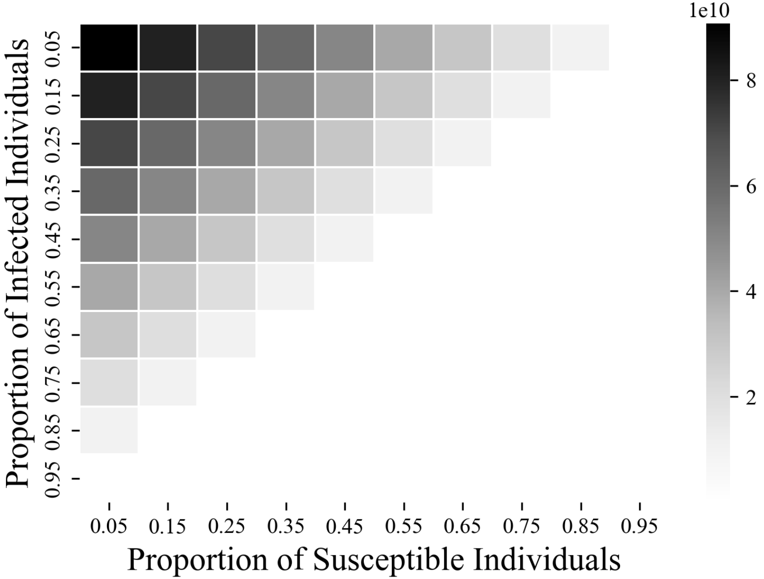
Base case

We estimated $\frac{c_O}{c_I} = 55$ and show the graphs for $a_P \in \{0.1, 0.5, 0.9\}$.

Optimal Interventions Recommended by the Model



Optimal Objective Function Values



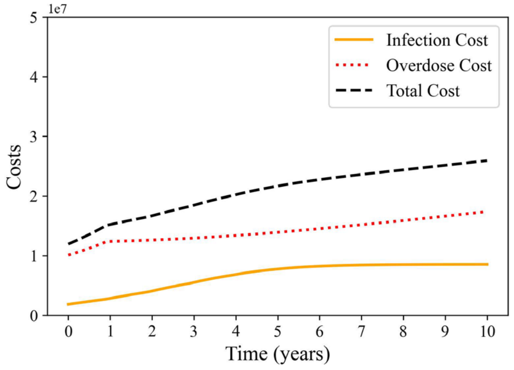
Sensitivity Analysis performed on $\frac{c_O}{c_I}$, nonlinear Impact magnitudes, and parameters.

Dynamic Implementation of Policies

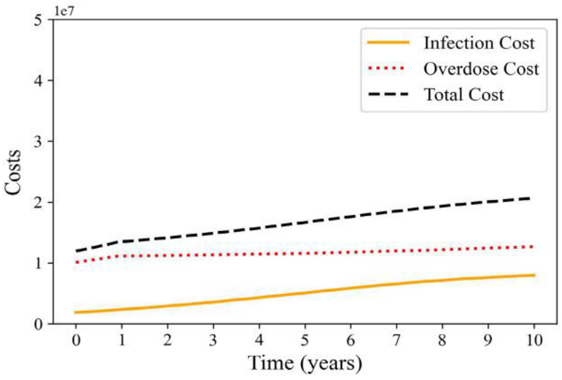
Goal: To investigate the costs associated with the unfolding crisis.

- Computed costs monthly, over 10 years.
- Employed the **MODSIM** package in Python.
- Compared the optimal with no intervention, 50-50, 90-10, and 10-90 allocation policies.
- Sensitivity analysis on the initial state.

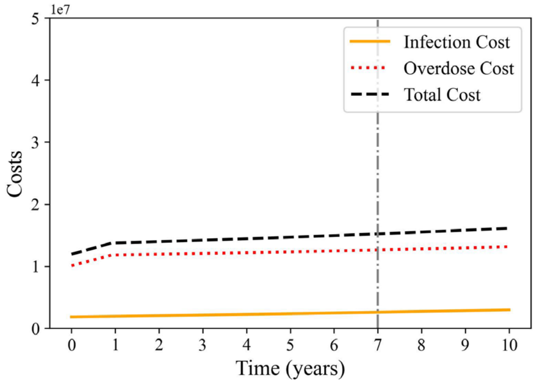
Simulation Outcomes-Cost Trends over 10 Years



Under No Intervention



Under the 50-50 policy



Under the Optimal Policy

Implementing our proposed optimal policy yields an average improvement of **29% (12%)** in total costs compared with the no intervention (50-50 policy), over 10 years.

Note. The vertical line shows the year that costs start to stabilize. Base case initial state is used.

Conclusion

- Even though the prescribing rates are controlled but overdose death rates are still rising => need for a **better budget allocation** for eventual mitigation.
- We developed a **state-dependent decision-support tool** to recommend the optimal budget allocation between interventions.
- *Prevention still is better than the cure!* But we need to invest in the cure, too!

Let's combat the epidemic!

Thank you

Sina Ansari
sina.ansari@depaul.edu

Scan this for more info about me!

Backup Slides

Table 2. Cost Savings Comparison: Optimal vs. No Allocation and Optimal vs. Fixed Allocation

	Optimal vs. no allocation			Optimal vs. fixed allocation		
Year	Infection cost	Overdose cost	Total cost	Infection cost	Overdose cost	Total cost
1	31%	5%	10%	17%	−6%	−2%
2	50%	5%	16%	30%	−6%	1%
3	61%	7%	23%	40%	−7%	5%
4	67%	9%	29%	48%	−7%	8%
5	70%	11%	32%	53%	−6%	12%
6	70%	14%	34%	58%	−6%	15%
7	69%	17%	36%	60%	−6%	18%
8	68%	19%	36%	62%	−5%	20%
9	67%	22%	37%	63%	−4%	21%
10	65%	24%	38%	62%	−4%	22%
Average	62%	13%	29%	49%	−6%	12%

FROM PREDICTION TO POLICY: AI-ENABLED DECISION-MAKING IN FERTILITY TREATMENT

BOGDANA JELIĆ, UNIVERSITY OF OXFORD
JOINT WORK WITH HOLLY WIBERG, AGNI ORFANOUDAKI , GURPREET KALRA

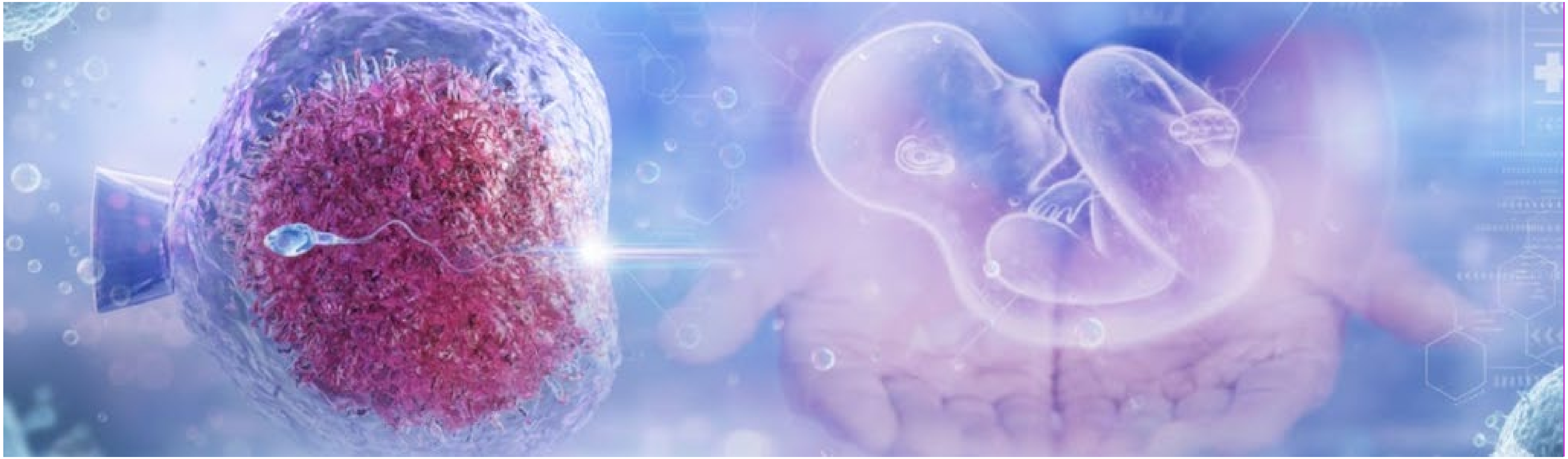
MEDICUS

UNIVERSITY OF OXFORD

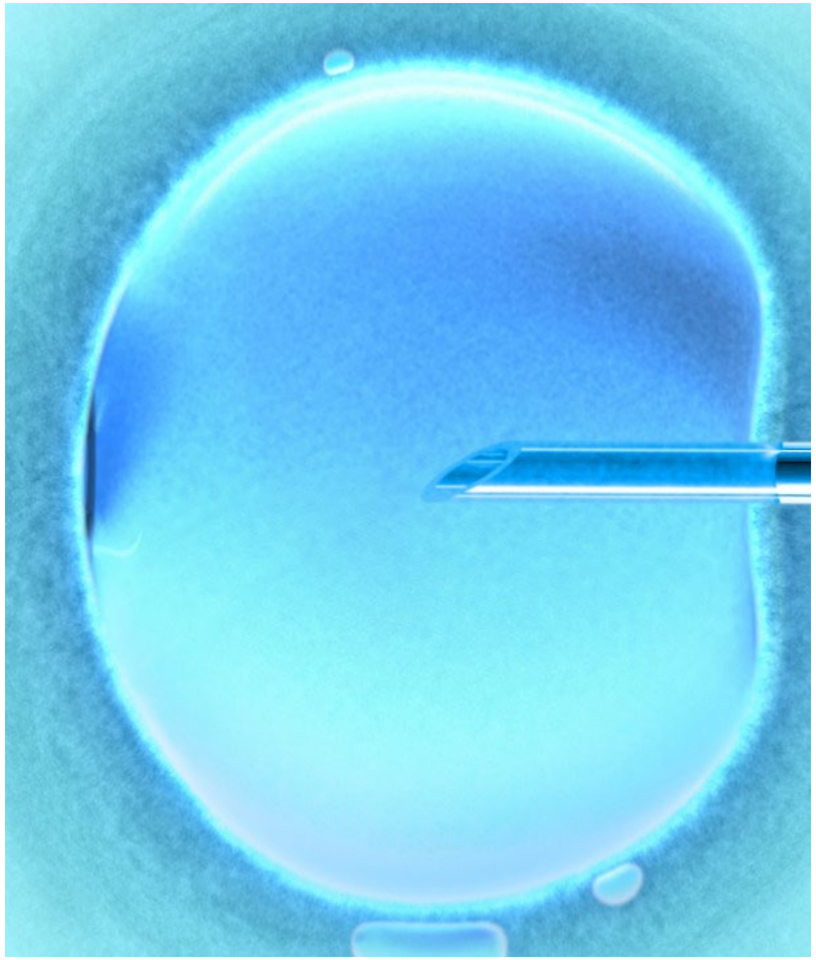
Saïd Business School

Carnegie Mellon University Heinz College

INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT



AGENDA



IVF overview

Role of AI in IVF

Impact of estimation uncertainty

- In planning
- In decision-making

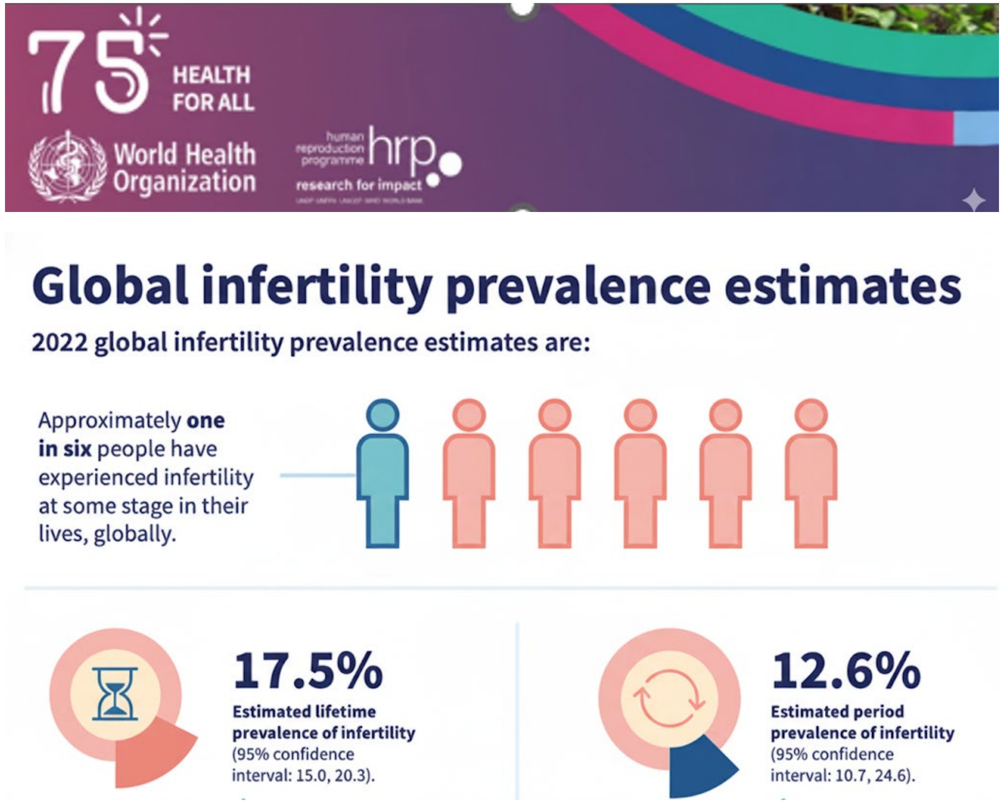
Stakeholder insights

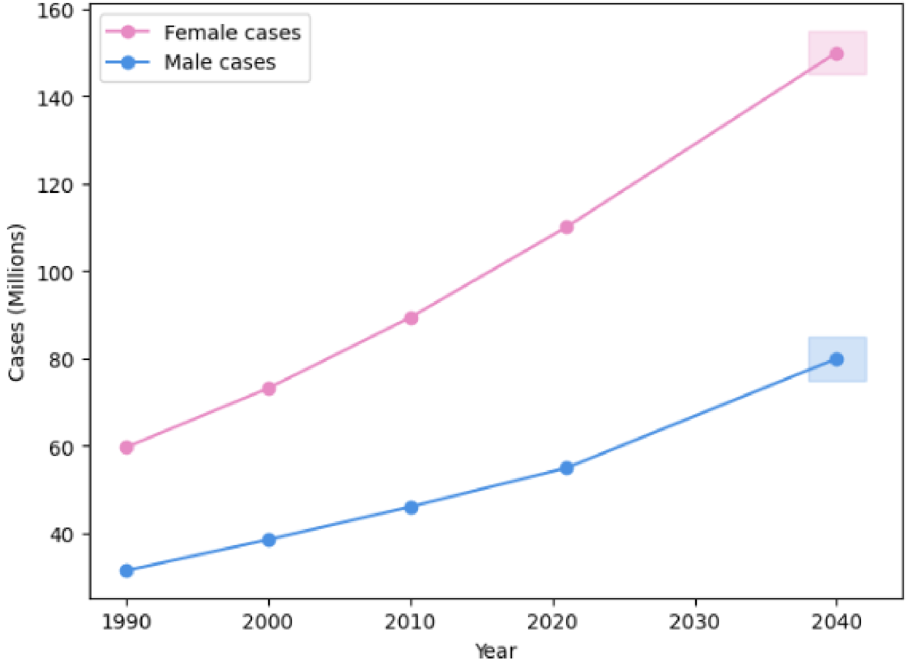
Goals:

- **Improve care navigation for patients**
- **Inform AI model development**
- **Guide AI use in IVF clinics**

THE BURDEN OF INFERTILITY

Infertility is prevalent **globally** and across **diverse populations.**





Infertility is **costly**:

- Financially (Katz et al 2012, Bögl et al 2025)
- Psychologically (Boivin and Lancastle 2010, Volgsten and L. Ekselius 2008)
- Socially (Wang and Fu 2022, Wang 2022)

Source: WHO Infertility Prevalence Estimates, 1990-2021

ASSISTED REPRODUCTIVE TECHNOLOGIES (ART) OFFER A SOLUTION

In-vitro fertilization (IVF) was introduced in 1978

Treatment is getting better over time!

IVF in 2025:

- **500K** children are born with IVF *per year*
- **10M** children have been born with IVF
- IVF comprises **2% of all live births**

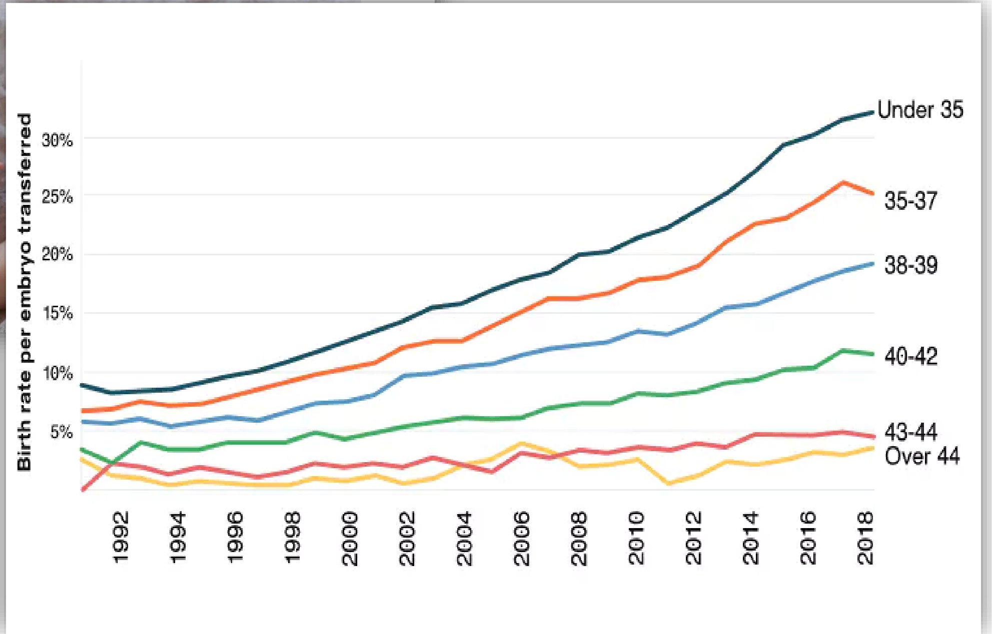

BBC Sign in Home News Sport Weather iPlayer Sounds

NEWS

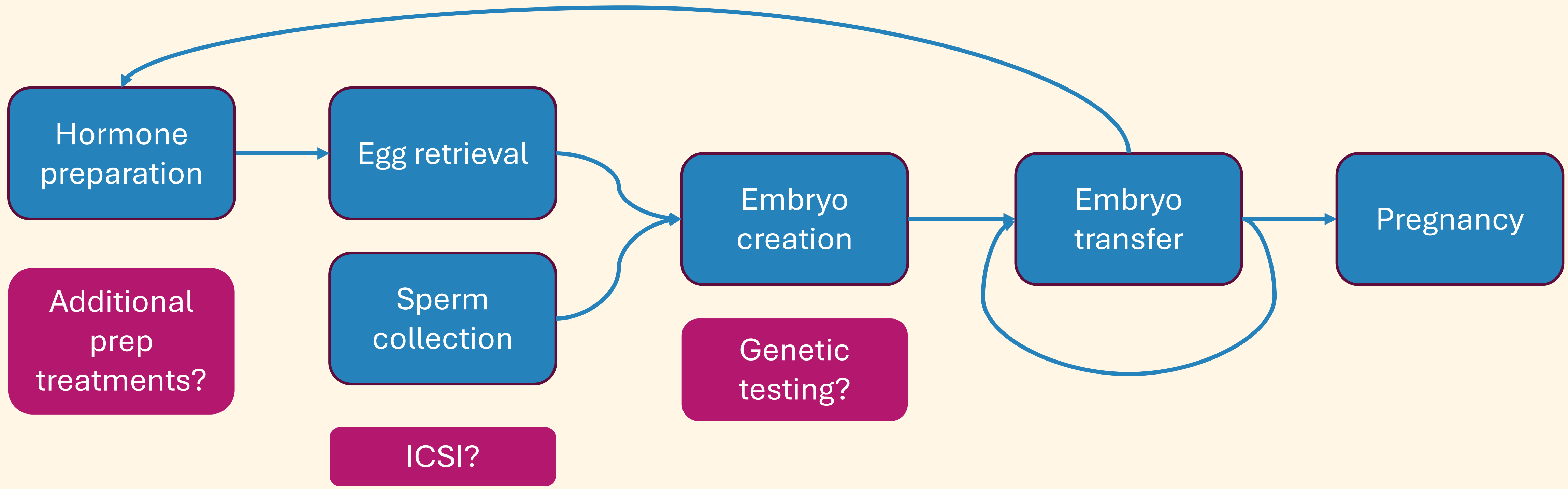
Home | InDepth | Israel-Gaza war | War in Ukraine | Climate | UK | World | Business | Politics | Culture

Health

IVF births now represent one child in every classroom, data suggests



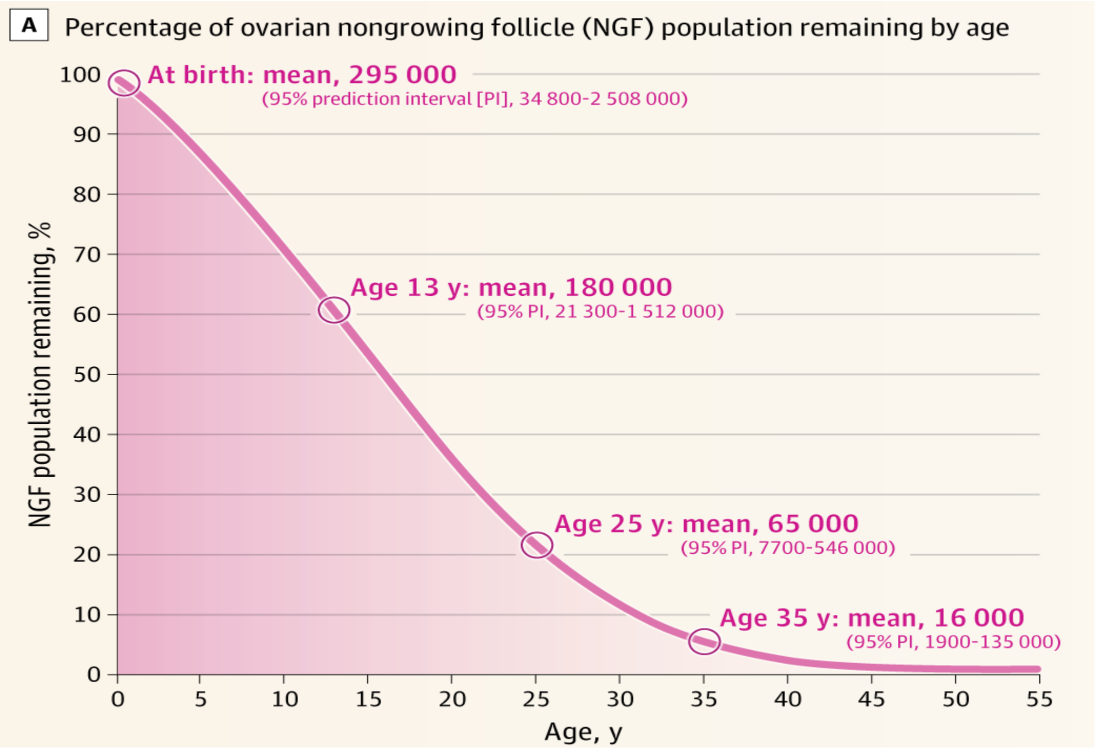
IVF: A DYNAMIC AND PERSONALIZED DECISION-MAKING PROBLEM

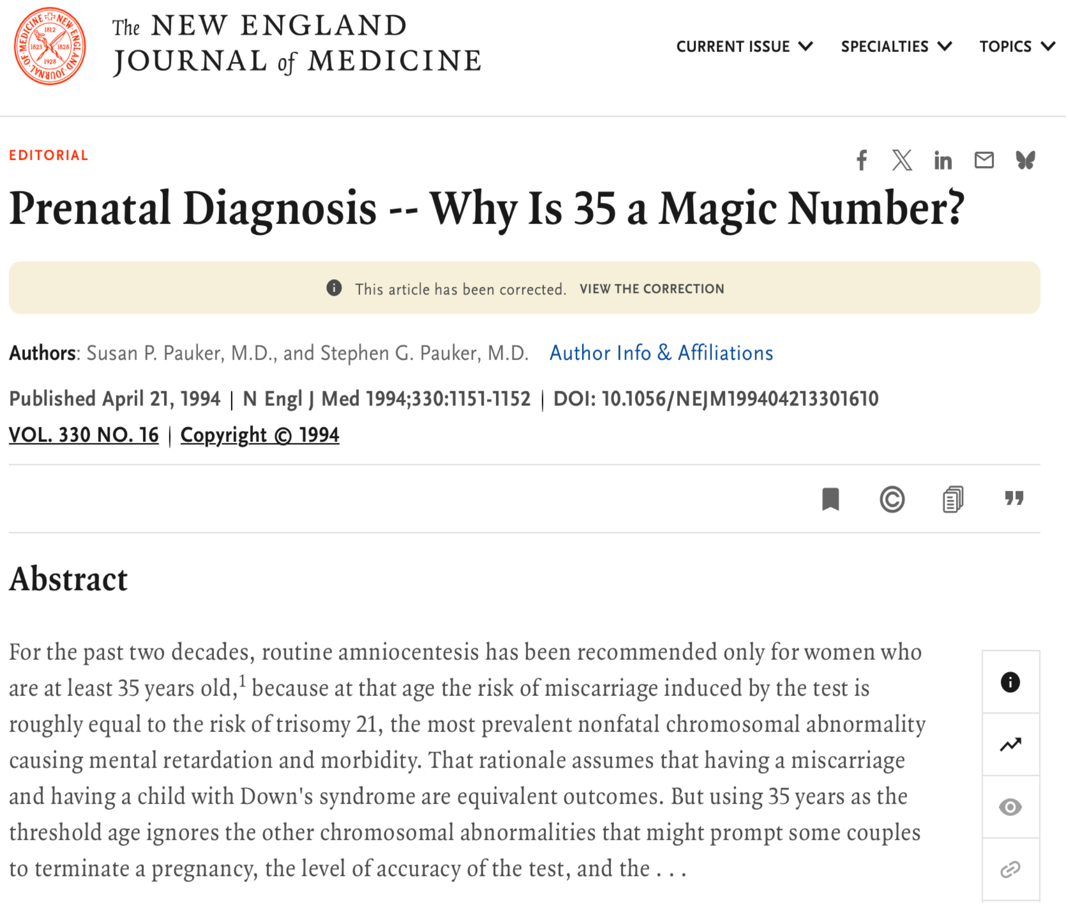


CURRENT STATE OF IVF

Average cost of an IVF cycle in the US: **$15,000 to $20,000**

Average cost of an IVF cycle in the UK: **£4,500 to £7,500**





The NEW ENGLAND JOURNAL of MEDICINE

CURRENT ISSUE SPECIALTIES TOPICS

EDITORIAL

Prenatal Diagnosis -- Why Is 35 a Magic Number?

This article has been corrected. VIEW THE CORRECTION

Authors: Susan P. Pauker, M.D., and Stephen G. Pauker, M.D. Author Info & Affiliations

Published April 21, 1994 | N Engl J Med 1994;330:1151-1152 | DOI: 10.1056/NEJM199404213301610

VOL. 330 NO. 16 | Copyright © 1994

Abstract

For the past two decades, routine amniocentesis has been recommended only for women who are at least 35 years old,[1] because at that age the risk of miscarriage induced by the test is roughly equal to the risk of trisomy 21, the most prevalent nonfatal chromosomal abnormality causing mental retardation and morbidity. That rationale assumes that having a miscarriage and having a child with Down's syndrome are equivalent outcomes. But using 35 years as the threshold age ignores the other chromosomal abnormalities that might prompt some couples to terminate a pregnancy, the level of accuracy of the test, and the . . .

AI FOR PERSONALIZED SUCCESS ESTIMATION IN IVF

Background and History

How old are you?

32

How much do you weigh?

145

How tall are you?

5 8

How many times have you used IVF in the past (include all cycles, even those not resulting in pregnancy)?

I've never used IVF

How many prior pregnancies have you had?

None

Diagnosis and Plan

What is the reason you are using IVF? (select all that apply)

- [ ] Male factor infertility
- [ ] Endometriosis
- [ ] Tubal factor
- [ ] Ovulatory disorder (including PCOS)
- [ ] Diminished ovarian reserve
- [ ] Uterine factor
- [ ] Other reason

(Or)

- [ ] Unexplained (Idiopathic) infertility

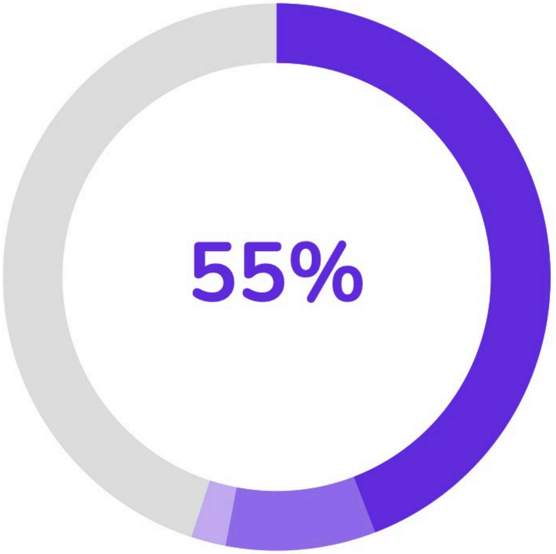
(Or)

- [x] I don't know/no reason

Do you plan to use your own eggs or donor eggs?

My own eggs

Calculate Success   Start Over



Cumulative chance of live birth (%)

After 1 retrieval and 1 transfer	44%
After 1 retrieval and 2 transfers	53%
After 1 retrieval and 3+ transfers	55%

Source: https://www.cdc.gov/art/ivf-success-estimator/index.html

CONSEQUENCES OF IGNORING UNCERTAINTY

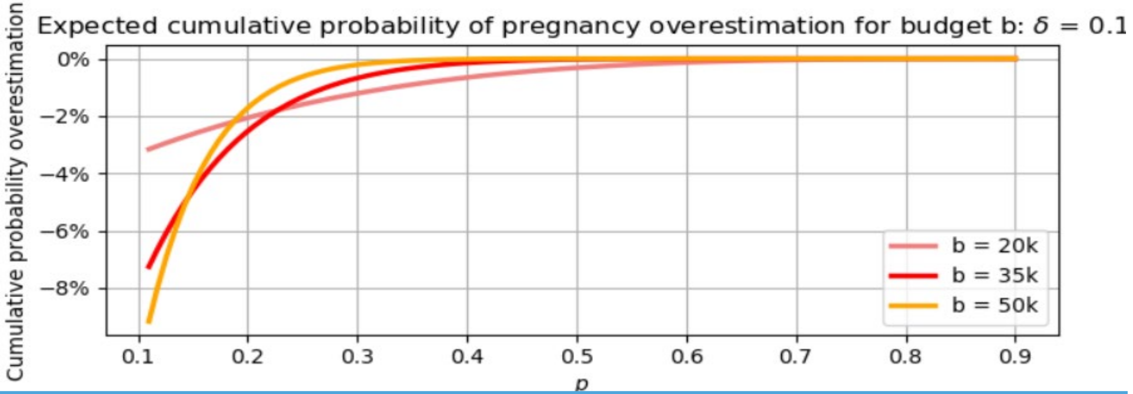
Assume that individual per-embryo **probability of success $\boldsymbol{p}$** is available subject to uncertainty: $\tilde{\boldsymbol{p}} \sim \boldsymbol{U}[\boldsymbol{p}-\boldsymbol{\delta}, \boldsymbol{p}+\boldsymbol{\delta}]$

Proposition: Overestimation of cumulative success

For a fixed budget b, we have that the expected cumulative probability of pregnancy satisfies:

$$\mathbb{E}[C(b,\tilde{p})] < C(b;p) \qquad \forall p \in (0,1), b \in \mathbb{R}^+$$

Insight: Ignoring uncertainty leads to overestimation of success and to budget underestimation



Proposition: Underestimation of the budget

If the individual probability of per-embryo pregnancy satisfies $\tilde{p} \sim U[p-\delta, p+\delta]$, under the continuous approximation, then:

$$\mathbb{E}_{q\sim\tilde{p}}[B(\alpha;q)] > B(\alpha;p) \quad \forall p \in (0, 1-e^{-2}], \alpha \in [0,1]$$

CONSEQUENCES OF IGNORING UNCERTAINTY

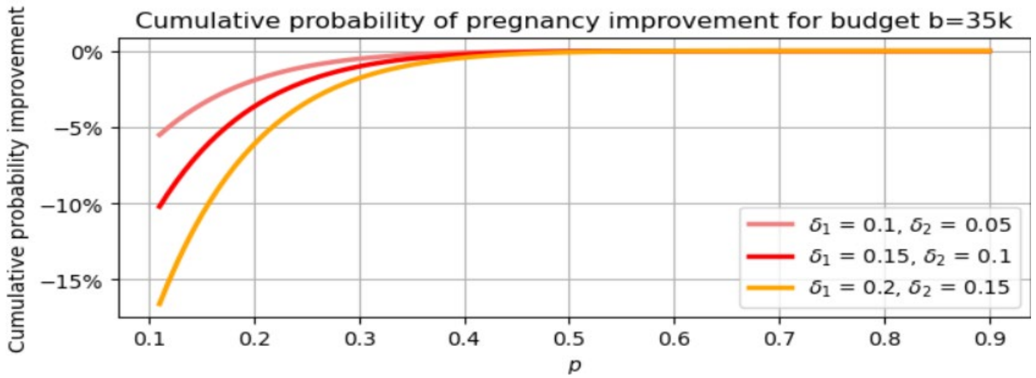
Insight: The error in overestimation of success and budget underestimation is being most significant for patients with a low probability of success.

Proposition: For any fixed budget b and uncertainty improvement, $\delta_1 \to \delta_2$, where $\delta_1 > \delta_2$, define the improvement in cumulative success probability as:

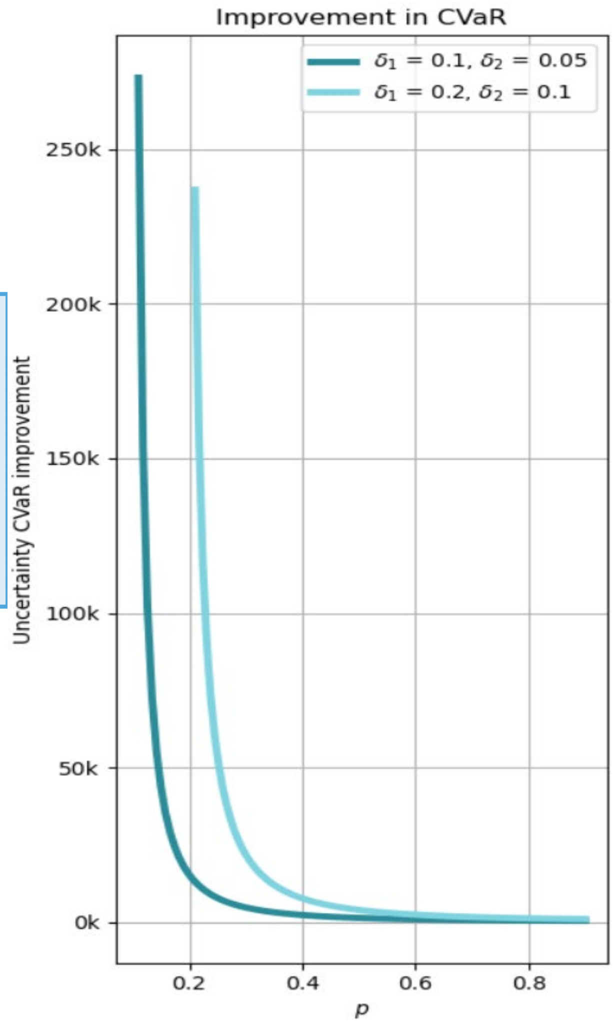
$$\Delta_b(\delta_1, \delta_2)(p) = \mathbb{E}_{\{p \sim U[p-\delta_1, p+\delta_1]\}}[C(p;b)] - \mathbb{E}_{\{p \sim U[p-\delta_2, p+\delta_2]\}}[C(p;b)]$$

Then we have that $\Delta_b(\delta_1, \delta_2)(p) > 0 \ \ \forall p$ and that

$$\Delta_b(\delta_1, \delta_2)(p_1) > \Delta_b(\delta_1, \delta_2)(p_2) \ \ \forall p_1 < p_2.$$



Insight: Uncertainty reduction improve relevant measures and the reduction is greatest for low p.



DECISION MAKING – PERSONALIZED RISKS AND DECISIONS

Patients have heterogeneous **priorities, constraints, and risk preferences**.

Time to pregnancy

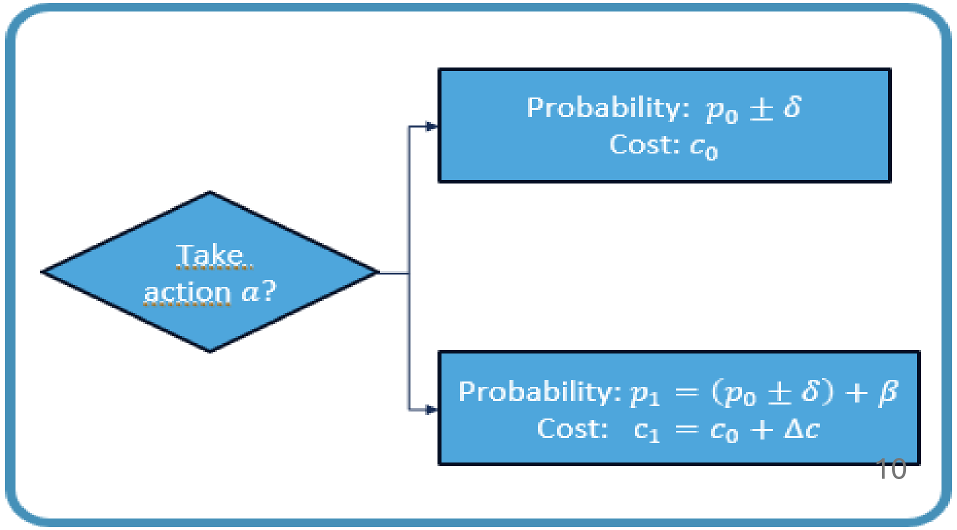
Cost of treatment

Risk aversion

Even for similar medical profiles, the optimal decision pathway can look different for each patient.

Patients can choose to take **costly** actions to **increase probability** of success.

- PGT-A (genetic screening)
- Endometrial receptivity assay (ERA)
- Progesterone support
- ICSI



VALUE FUNCTION

$$V(b) = \max_{\{i=0,1\}} \left\{ -\alpha\, E^{i}_{exp} + p_i \frac{1 - \left((1 - p_i)\gamma\right)^N}{1 - (1 - p_i)\gamma} + \left((1 - p_i)\gamma\right)^N V(b - E^{i}) \right\}$$

Financial cost — Time — Future discounted value

Where:

- $\boldsymbol{\alpha} \in [0,1]$: large value = high **budget sensitivity**
- $\boldsymbol{\gamma} \in [0,1]$: small value = large **time penalty**
- E^{i}_{exp} - expected cost of one cycle with decision i (E^{i} = maximum cost of cycle with decision i)

11

WHAT IS THE OPTIMAL CHOICE?

Theorem: the limiting optimal choice as budget grows $b \to \infty$ is

$$\text{Opt-choice} = \text{argmax}_{\{i=0,1\}} \frac{C_i}{1-r_i}$$

Value-weighted cost of decision

Probability of pregnancy for the decision

where $C_i = -\alpha\, E_{\text{exp}}^{i} + \frac{p_i\left(1-((1-p_i)\gamma)^N\right)}{1-(1-p_i)\gamma}$ and $r_i = ((1-p_i)\gamma)^N$

With limited budget, the patient's optimal decision can differ, leading to unstable decision paradigms.

DECISION AS A FUNCTION OF α

Stabilization point: Budget at which the decision stabilizes to the long-term optimum (“final switching point”)

$$D(\cdot) = \sup\{\, b : \exists p' \in [p_0 - \delta, p_0 + \delta]\ s.t.\ O(p', p' + \beta;, \alpha, \gamma, a)(b) \neq opt - choice\}$$

Ambiguity introduction point: Minimum budget at which the decision is ambiguous

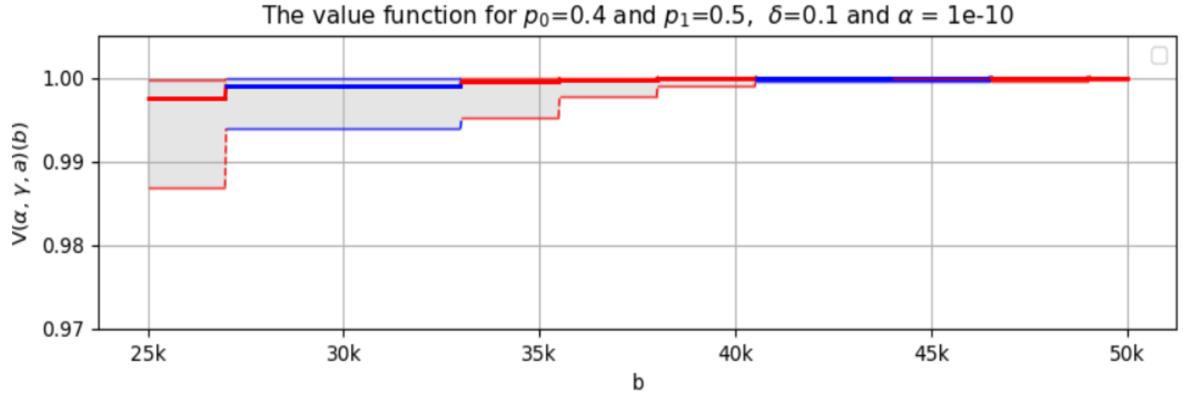
$$S(\cdot) = \inf\{\, b : \exists p', p'' \in [p_0 - \delta, p_0 + \delta]\ s.t.\ O(p', p' + \beta; \alpha, \gamma, a)(b) \neq O(p'', p'' + \beta;\ \alpha, \gamma, a)(b)\}$$

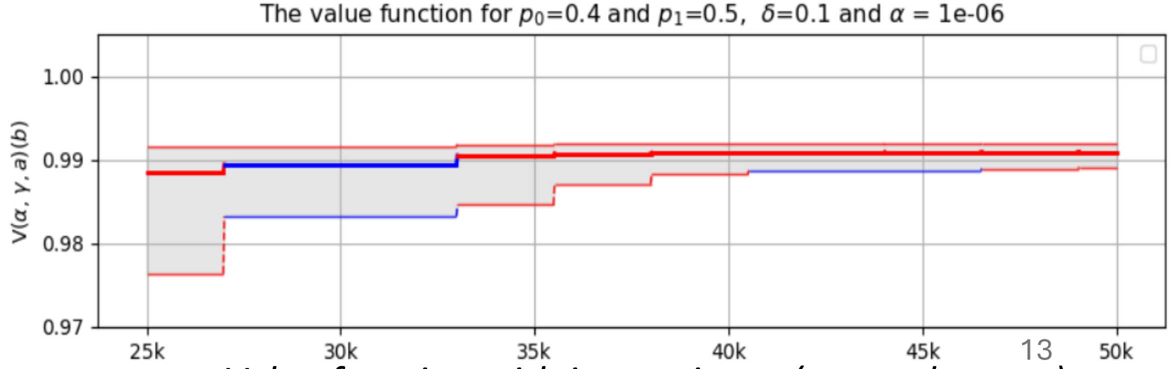
Proposition [informal]: as a patient becomes more budget sensitive (α increases), their decisions stabilize at the optimal choice **earlier** (at a lower budget).

Formally, $\mathrm{D}(\alpha_1) \geq \mathrm{D}(\alpha_2)\ \forall\ \alpha_1 < \alpha_2$.

Proposition [informal]: as a patient becomes more budget sensitive (α increases), they face **more decision ambiguity**; namely, ambiguity at a lower budget.

Formally, $\mathrm{S}(\alpha_1) \geq\ \mathrm{S}(\alpha_2)\ \forall\ \alpha_1 < \alpha_2$.





Value function with increasing α (top vs. bottom)

KEY TAKAWAYS

Consequences of ignoring uncertainty	Ignoring uncertainty results in systematically over-estimating success and under-estimating budget, with the greatest burden on low probability patients.
Benefit of uncertainty reduction	Uncertainty reduction improves success and budget estimates, with greatest benefit for low probability patients.
Impact on decision-making	The optimal decision pathway depends on patient preferences, with uncertainty having a greater impact on budget constrained patients with more cost-sensitivity, with similar results for time sensitivity.