Seeing into the Future:
Machine learning for Personalized Screening

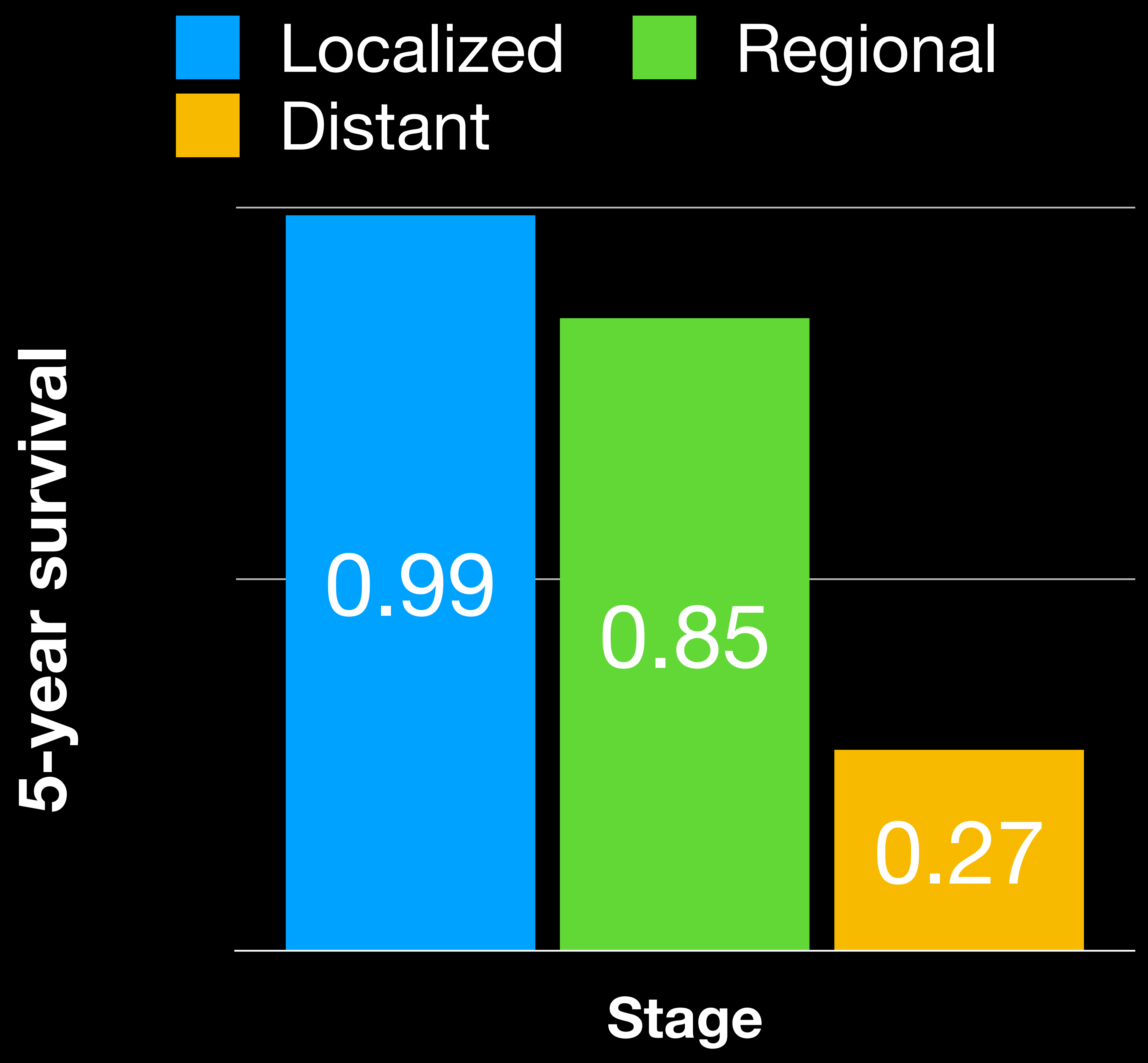
Adam Yala, PhD

Assistant Professor

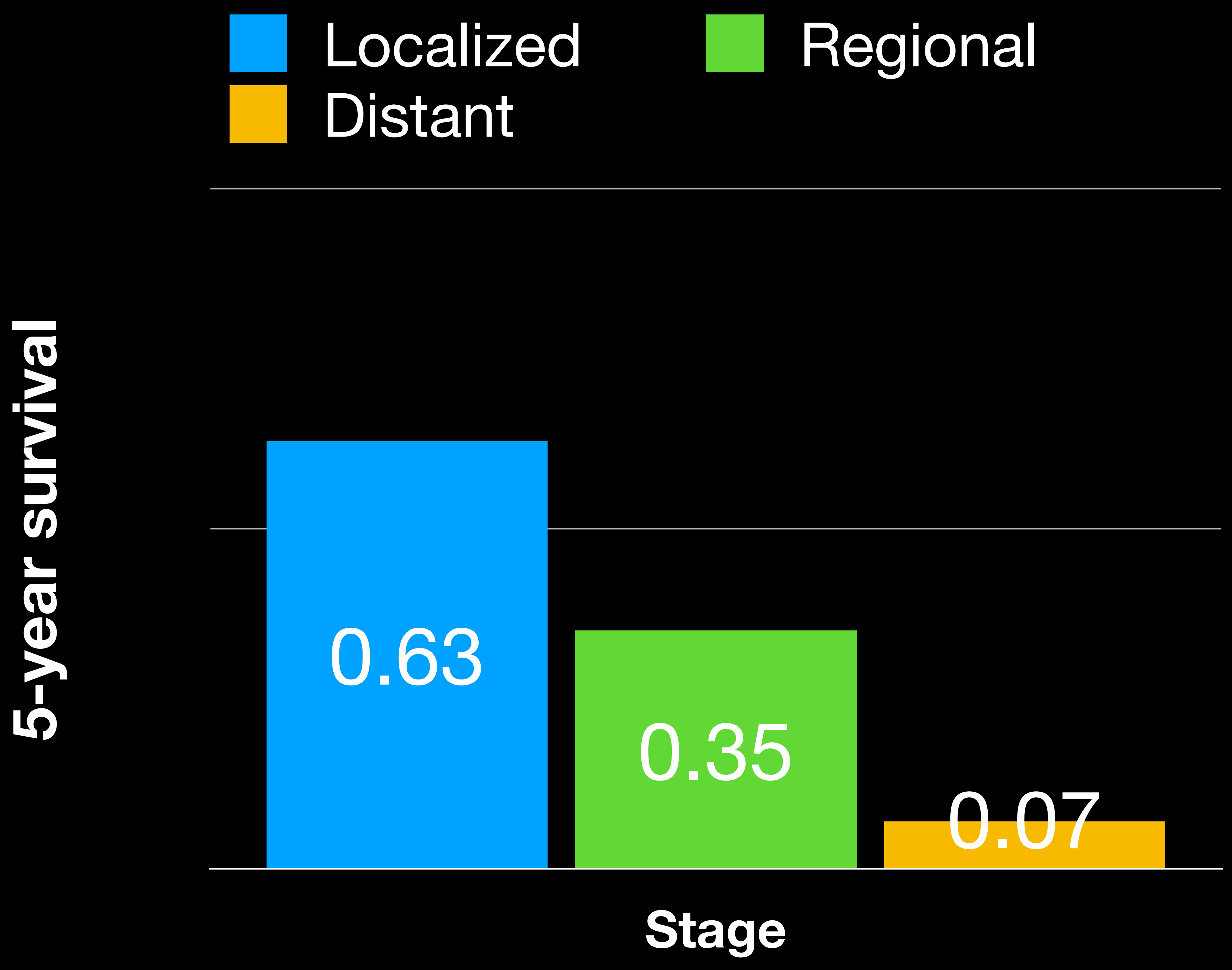
Computational
PRECISION HEALTH

Berkeley | UCSF

Early Detection is critical



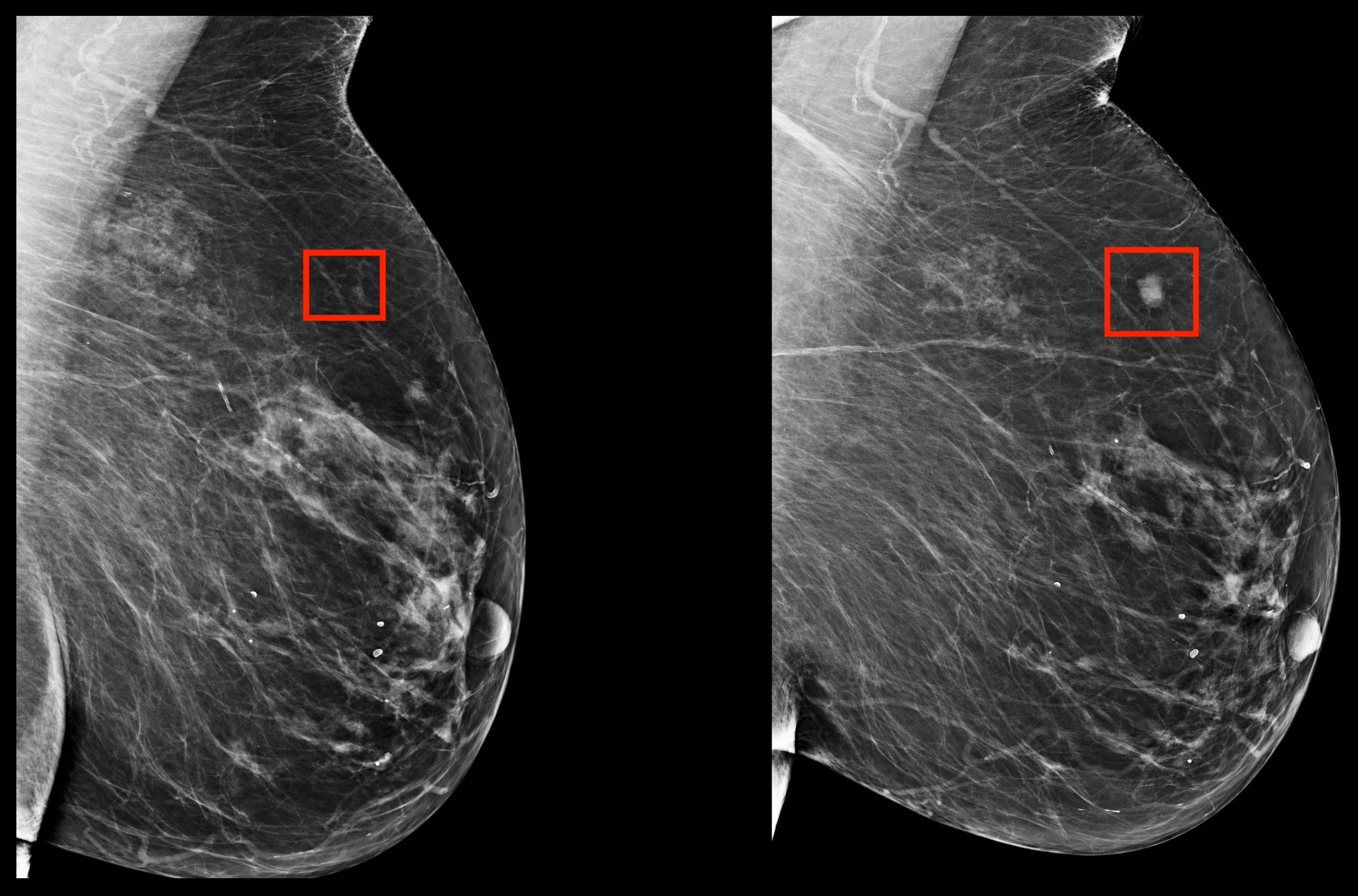
Breast Cancer



Lung Cancer

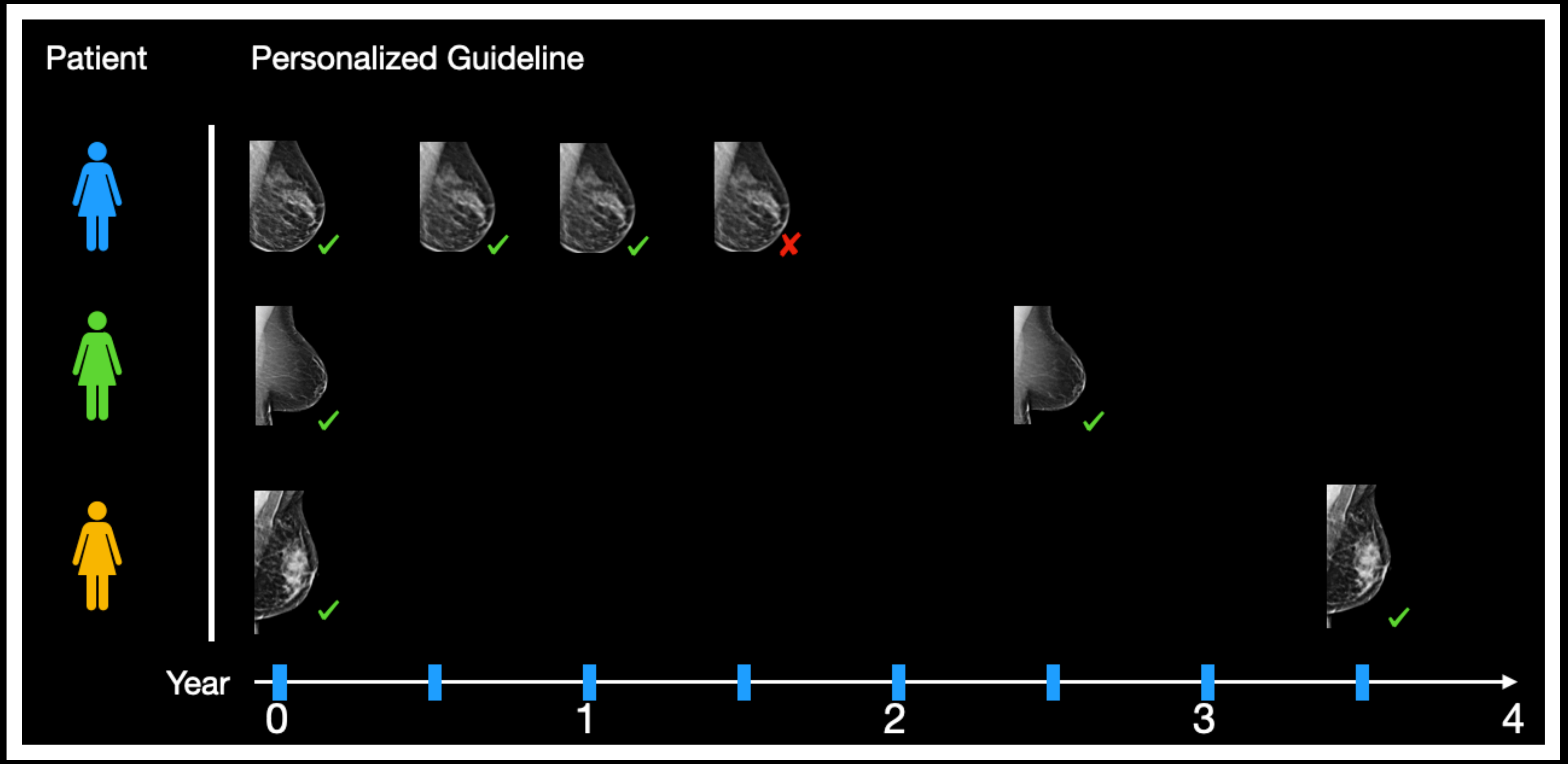
How to catch cancer earlier

Predict Cancer Risk



Year 0 Year 5

Create personalized screening policy

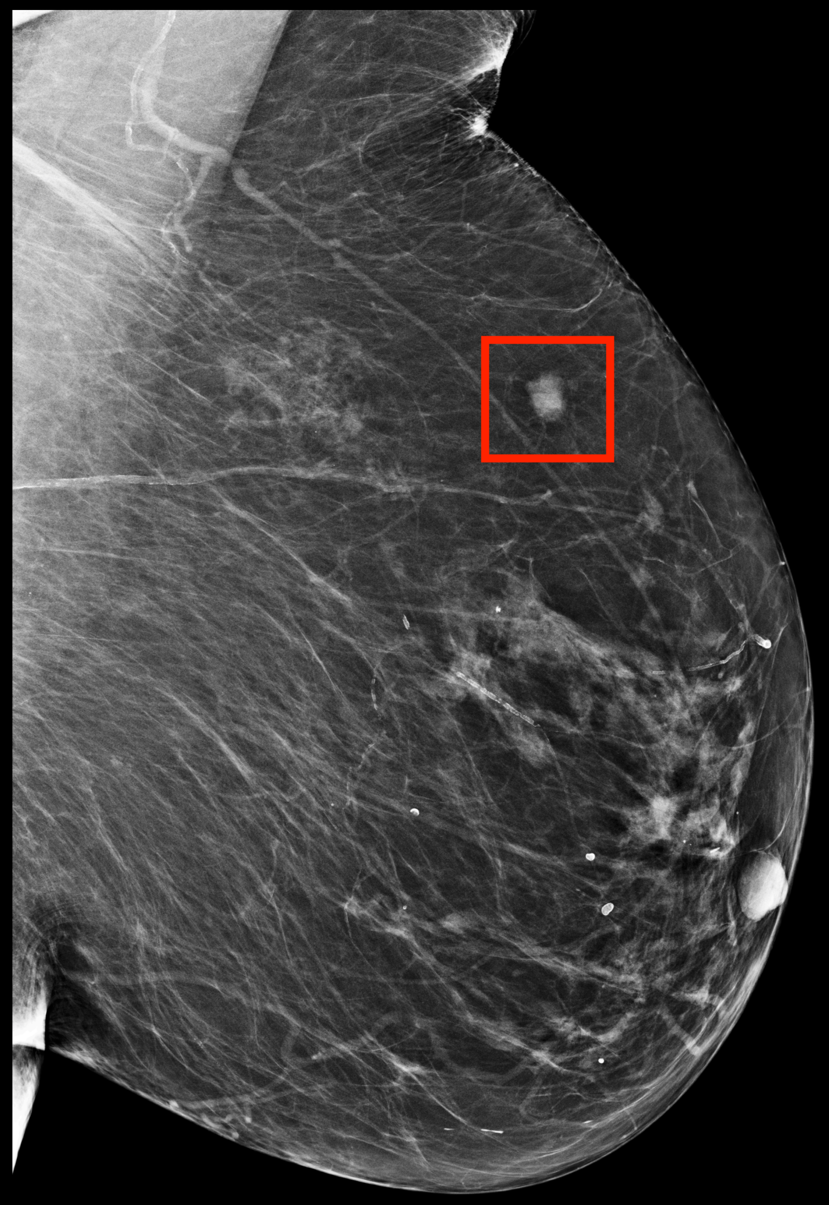


How to catch cancer earlier

Predict Cancer Risk

- Identify which population is at risk of developing cancer

Year 0



Year 5

Journal of Clinical Oncology®
An American Society of Clinical Oncology Journal

Multi-Institutional Validation of a Mammography-Based Breast Cancer Risk Model

Adam Yala, MEng[1,2]; Peter G. Mikhael, BS[1,2]; Fredrik Strand, MD, PhD[3,4]; Gigin Lin, MD, PhD[5]; Siddharth Satuluru, BS[6];

SCIENCE TRANSLATIONAL MEDICINE

Toward robust mammography-based models for breast cancer risk

Adam Yala[1,2]*, Peter G. Mikhael[1,2], Fredrik Strand[3,4], Gigin Lin[5], Kevin Smith[6,7], Yung-Liang Leslie Lamb[8], Kevin Hughes[9], Constance Lehman[8†], Regina Barzilay[1,2†]

Radiology

A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction

Adam Yala, MEng • Constance Lehman, MD, PhD • Tal Schuster, MS • Tally Portnoi, BS • Regina Barzilay, PhD

Traditional approach: use expert knowledge

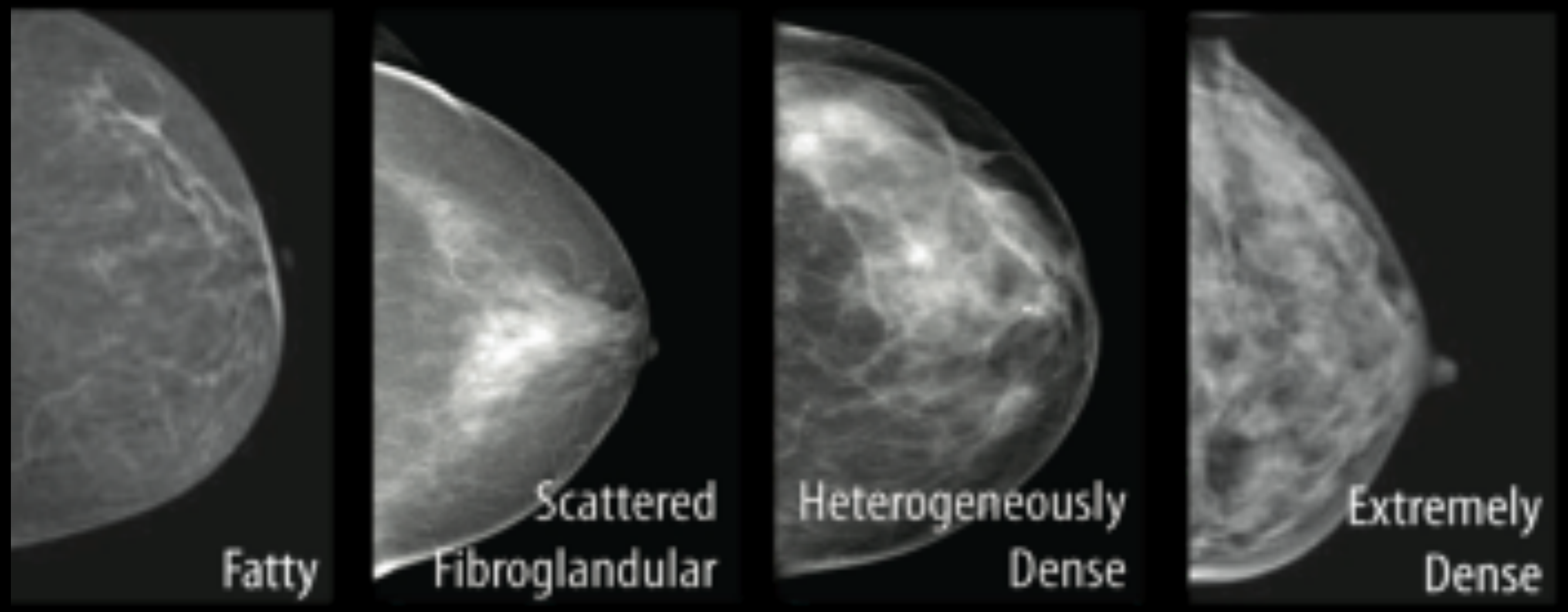
Age

Family History

Prior Breast Procedure

Breast Density

→ Model → Risk



AUC: 0.63

AUC: 0.61 without Density

J Natl Cancer Inst. 2006 Sep 6;98(17):1204-14.

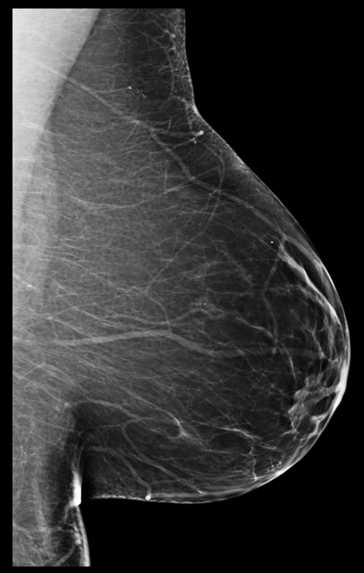
Prospective breast cancer risk prediction model for women undergoing screening mammography.

Barlow WE[1], White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, Tice JA, Buist DS, Geller BM, Rosenberg R, Yankaskas BC, Kerlikowske K.

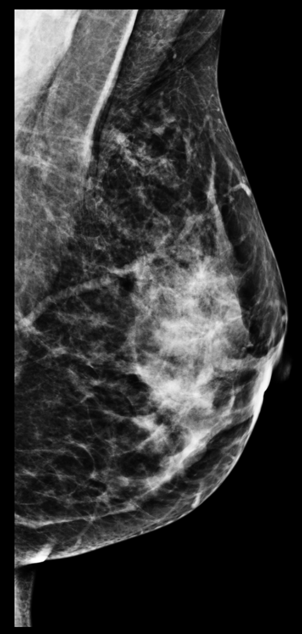
Learning to predict future cancer from imaging

Patient | **Future Outcome**

3 year cancer

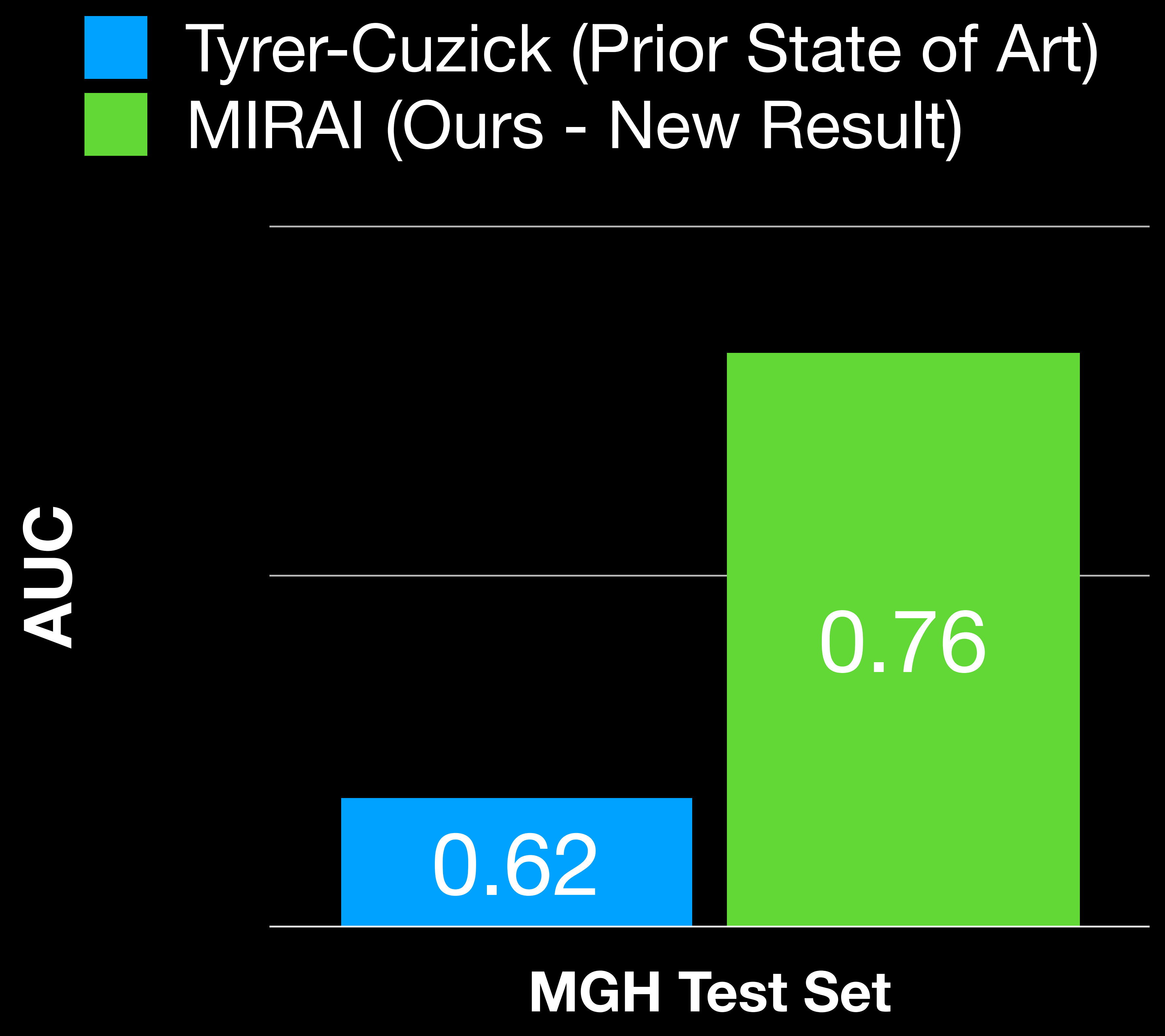


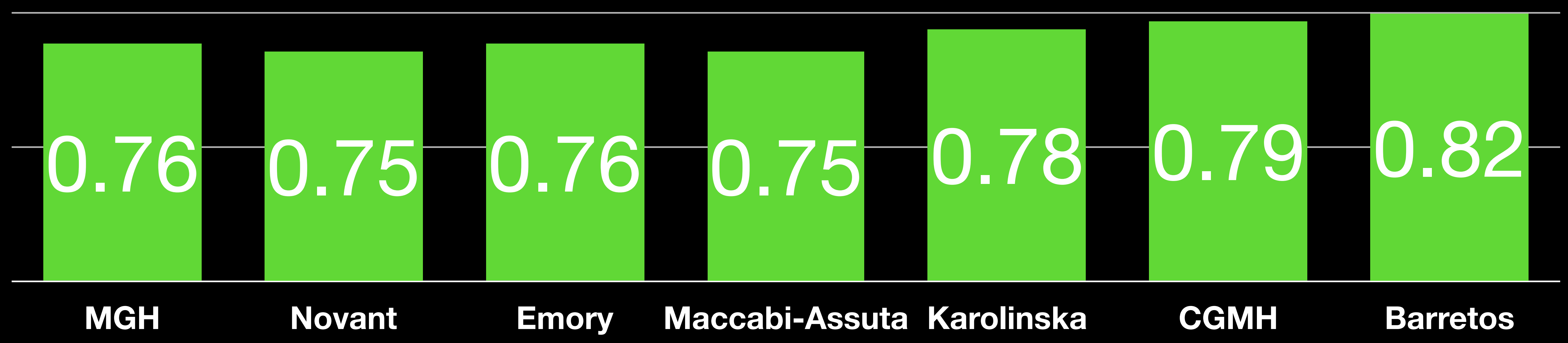
No cancer



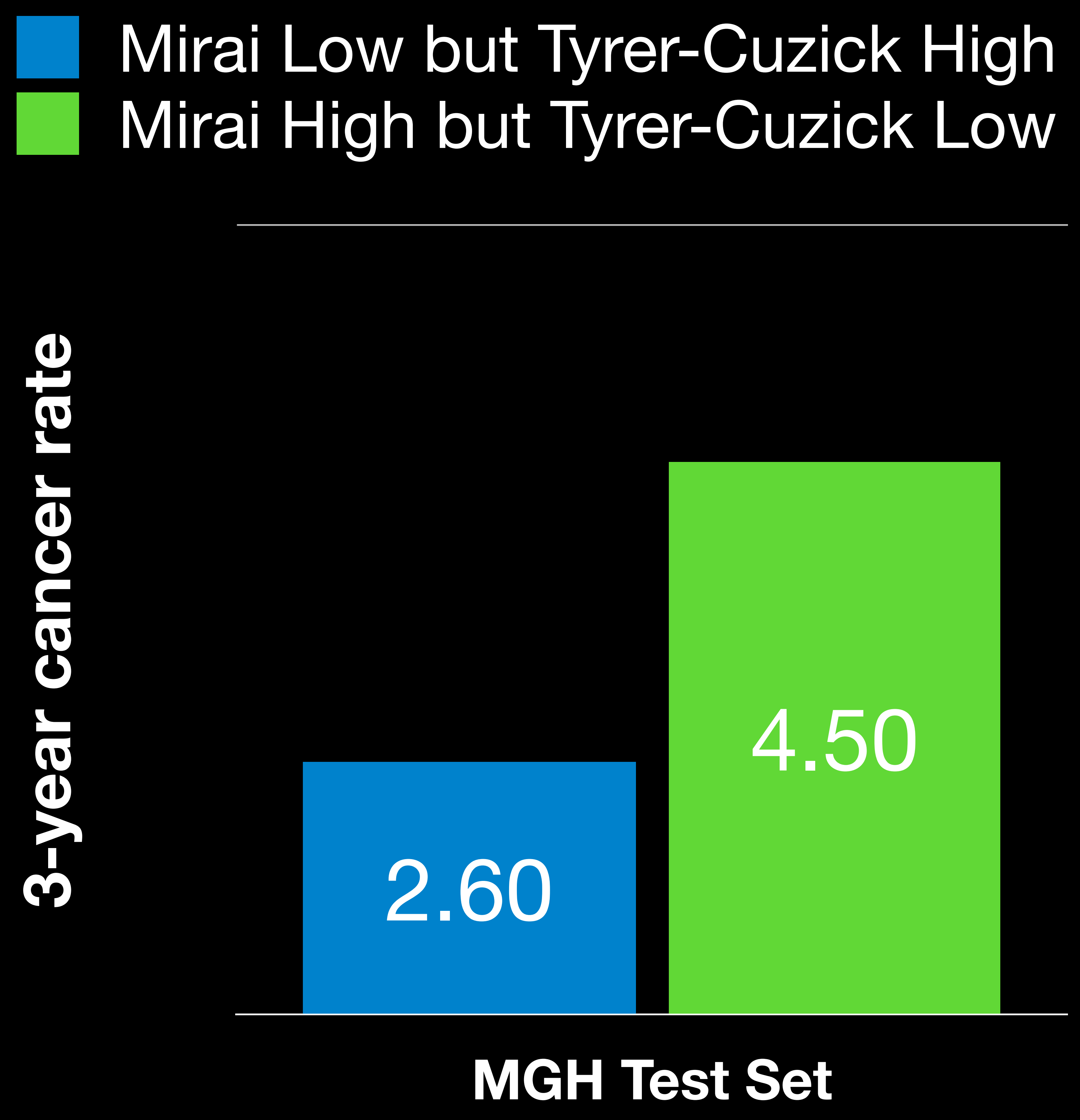
5 year cancer

Maintains accuracy across diverse populations





Selecting patients for supplemental imaging



Retrospective analysis

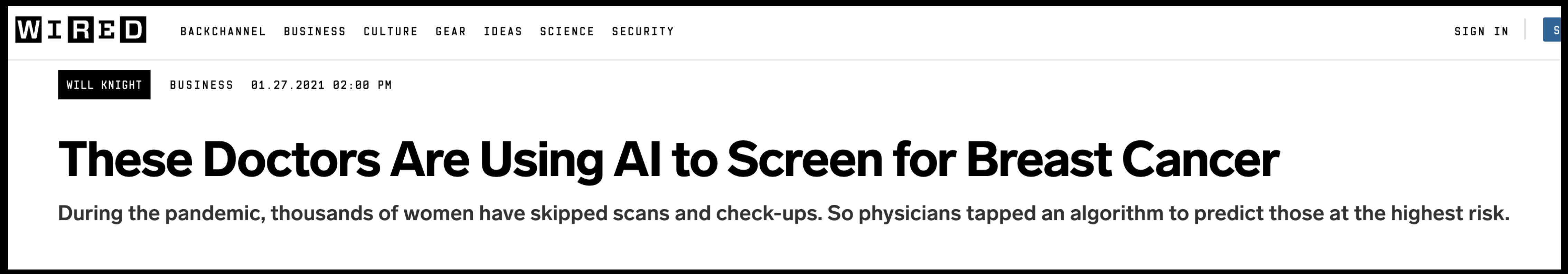
1.7x higher cancer yield, same MRI volume as current care.

Better early detection, same cost.

Mirai Use Cases

Organizing prospective trials for multiple use cases

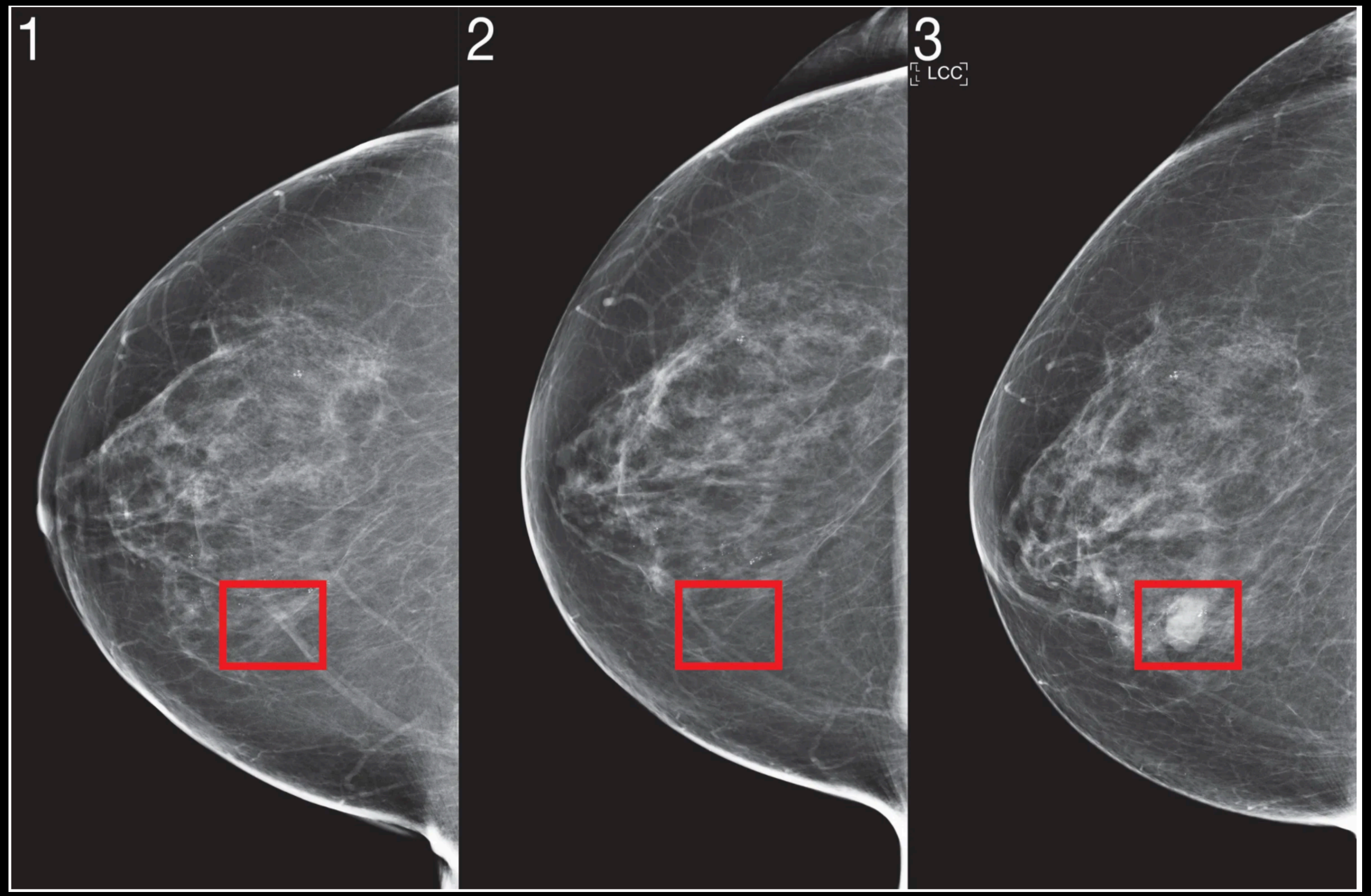
Highlight: Prioritize screening from covid backlog at MGH


WIRED BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY SIGN IN

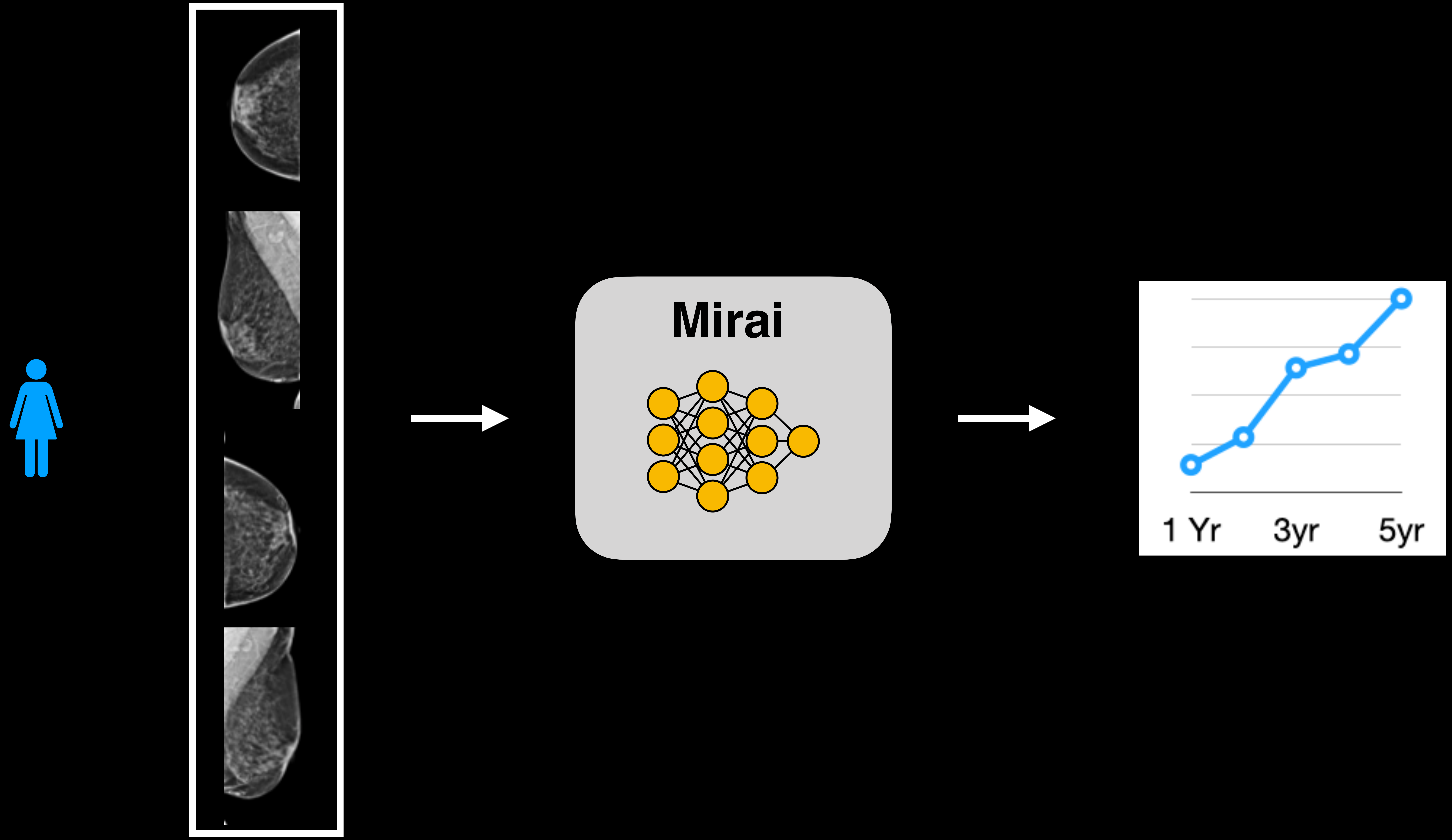
WILL KNIGHT BUSINESS 01.27.2021 02:00 PM

These Doctors Are Using AI to Screen for Breast Cancer

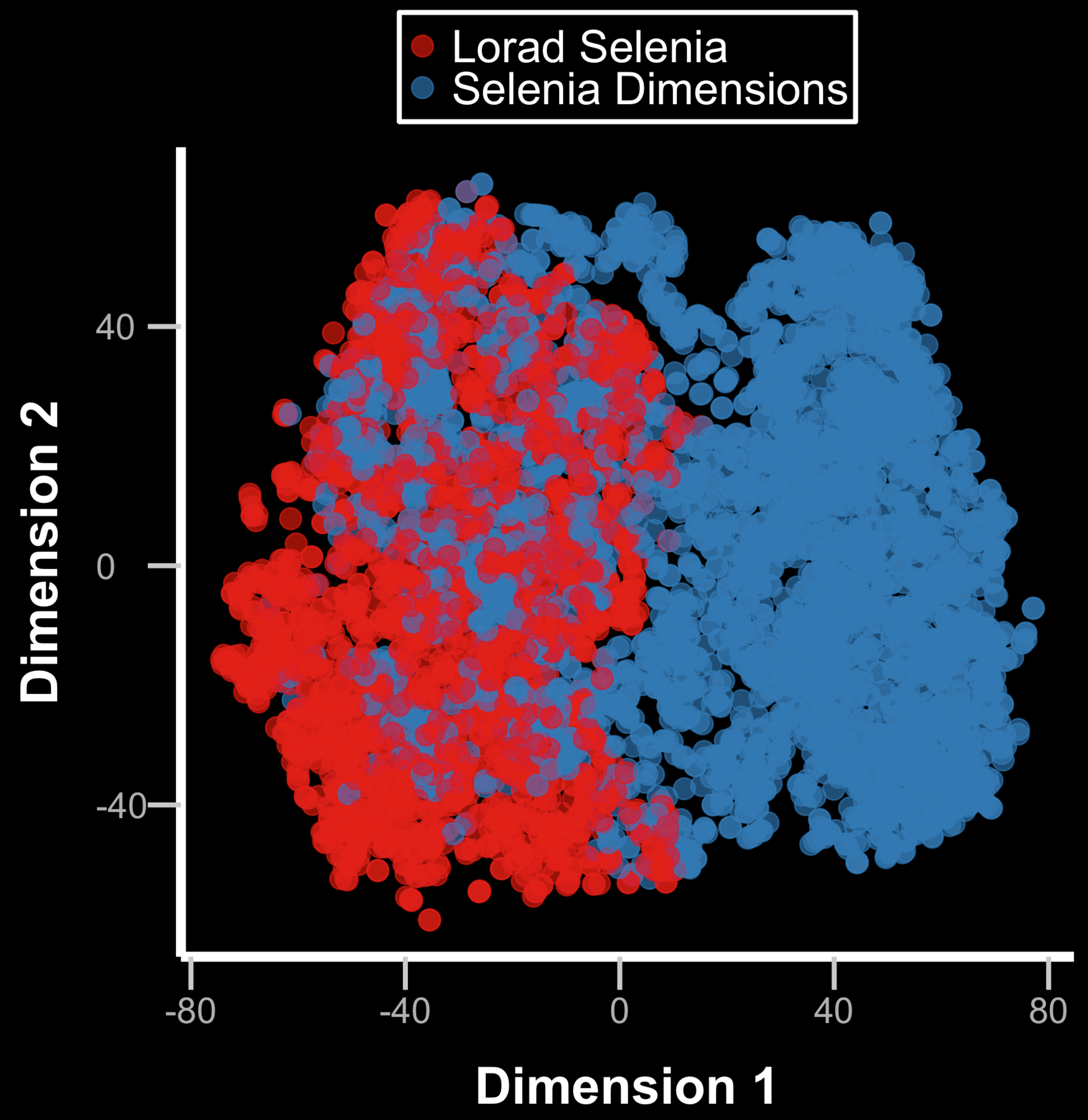
During the pandemic, thousands of women have skipped scans and check-ups. So physicians tapped an algorithm to predict those at the highest risk.



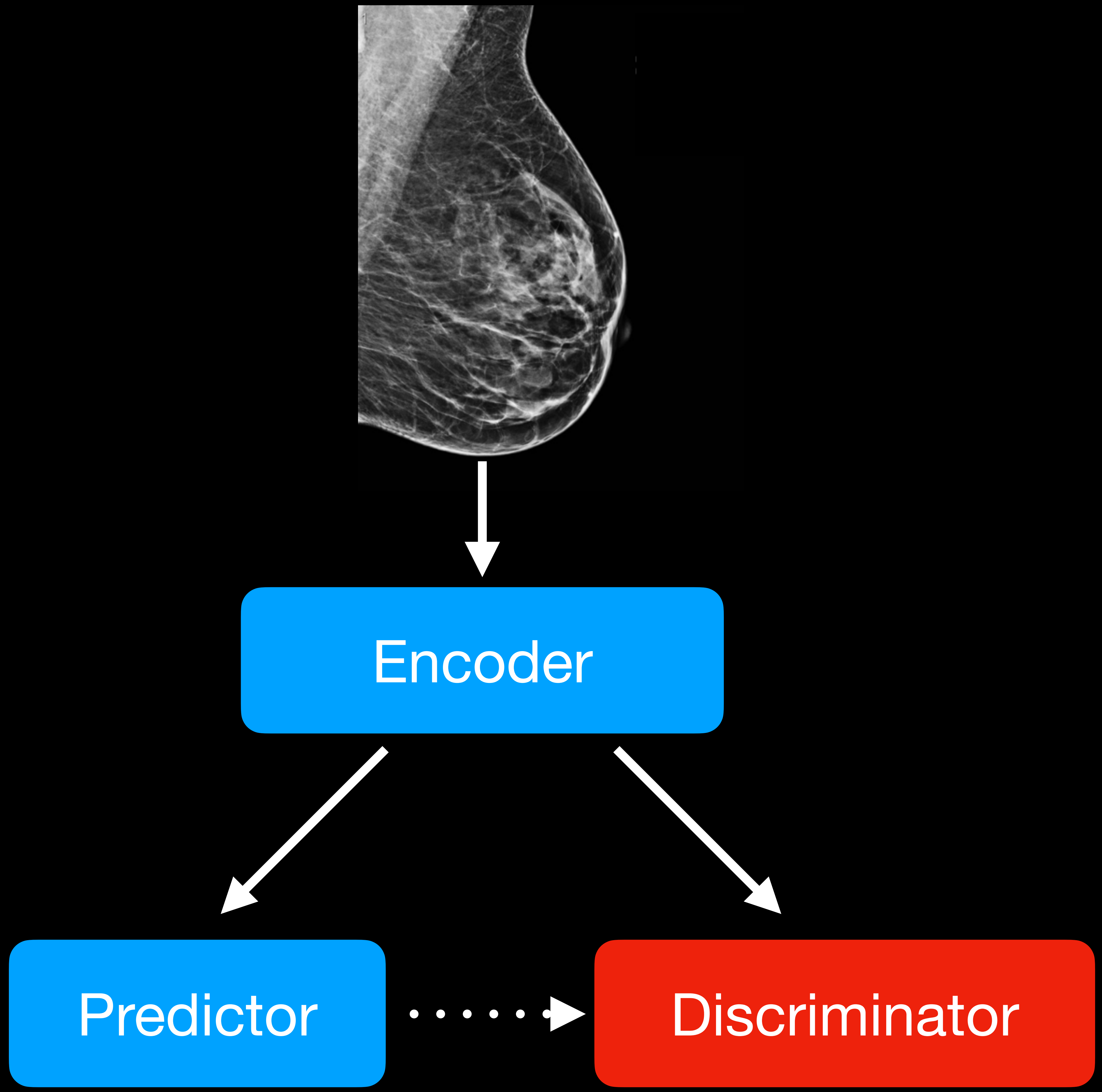
Mirai: Image-based Risk model



Problem 1: Device Invariance



Problem 1: Device Invariance

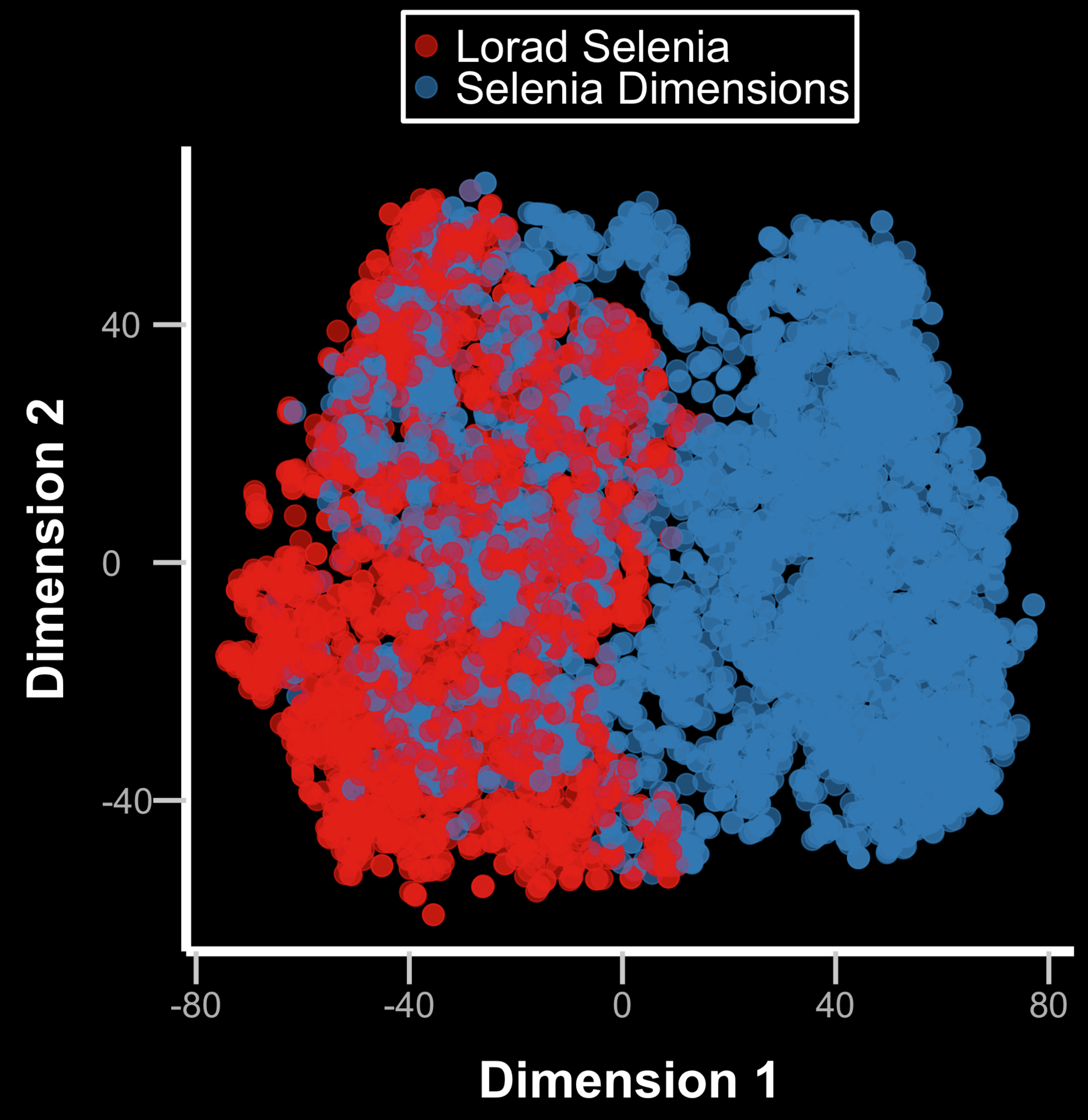


Objective:

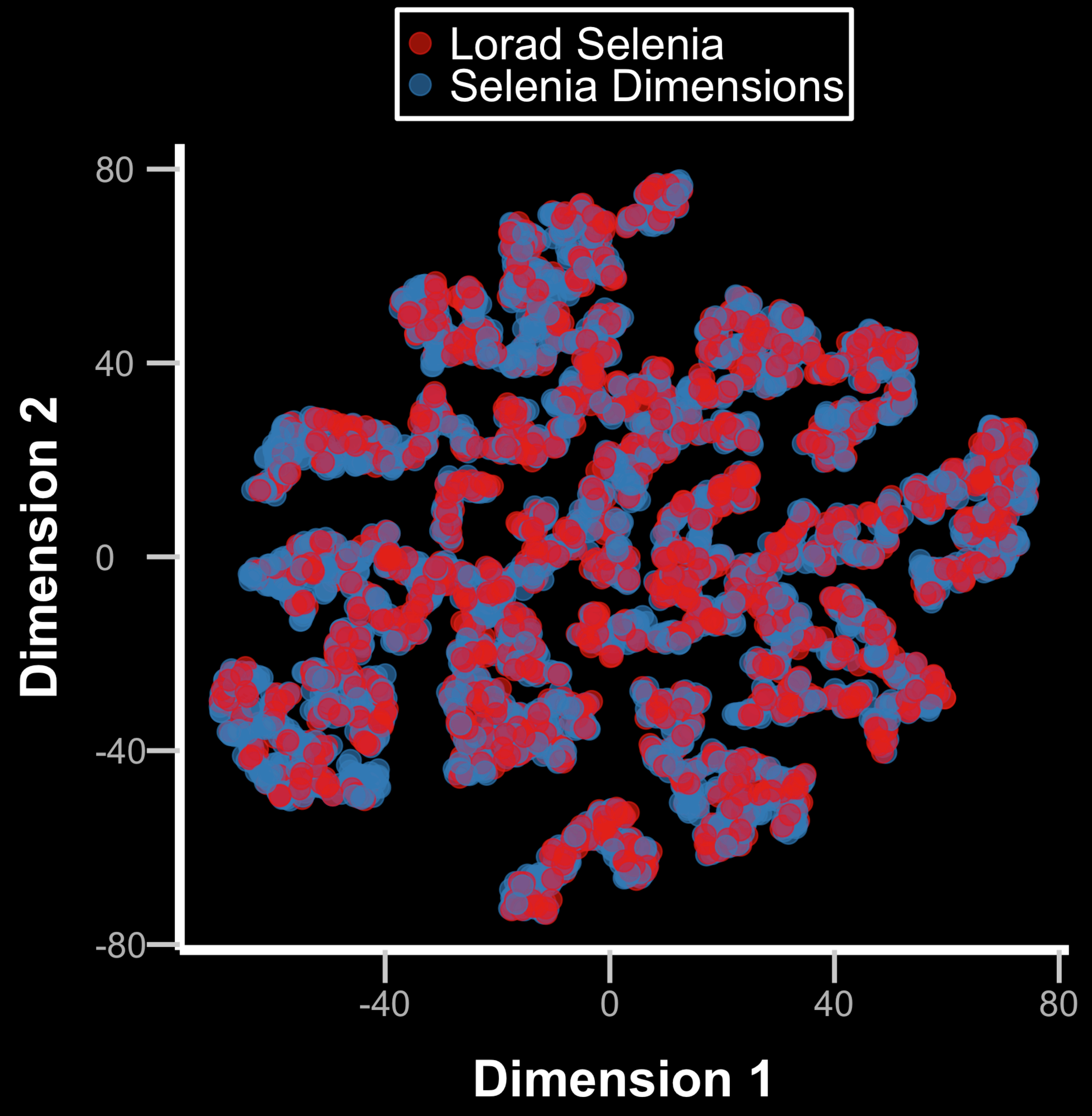
Max accuracy **Predictor**

Min accuracy **Discriminator**

Problem 1: Device Invariance

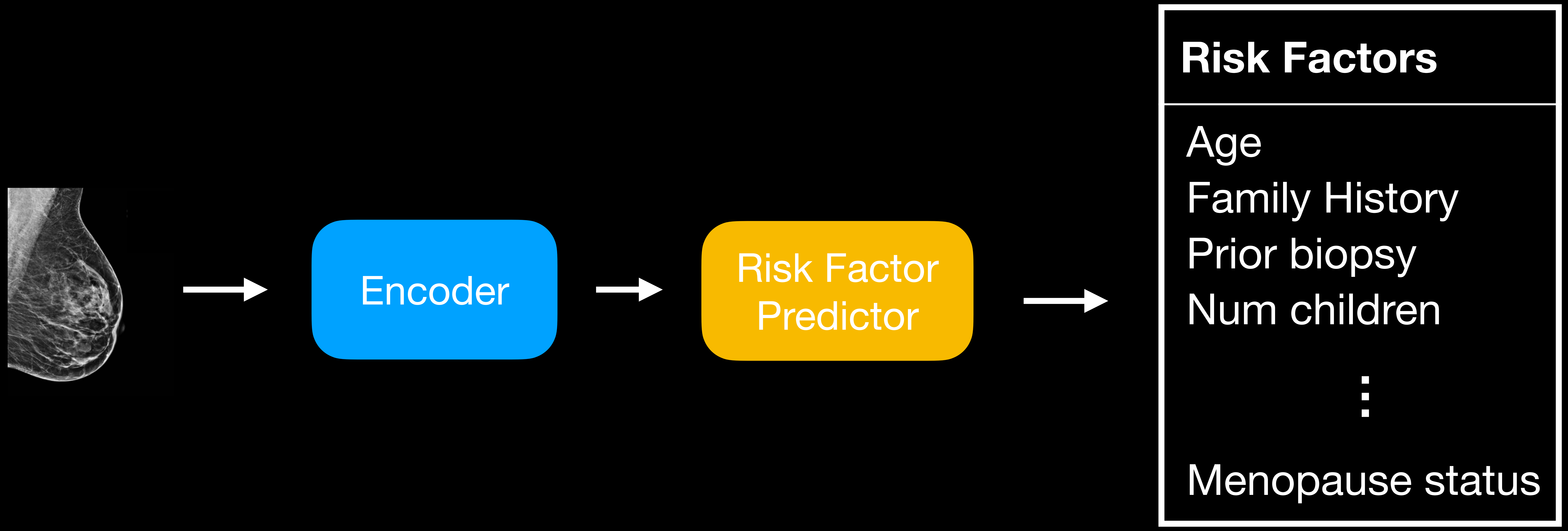


Without Adversary

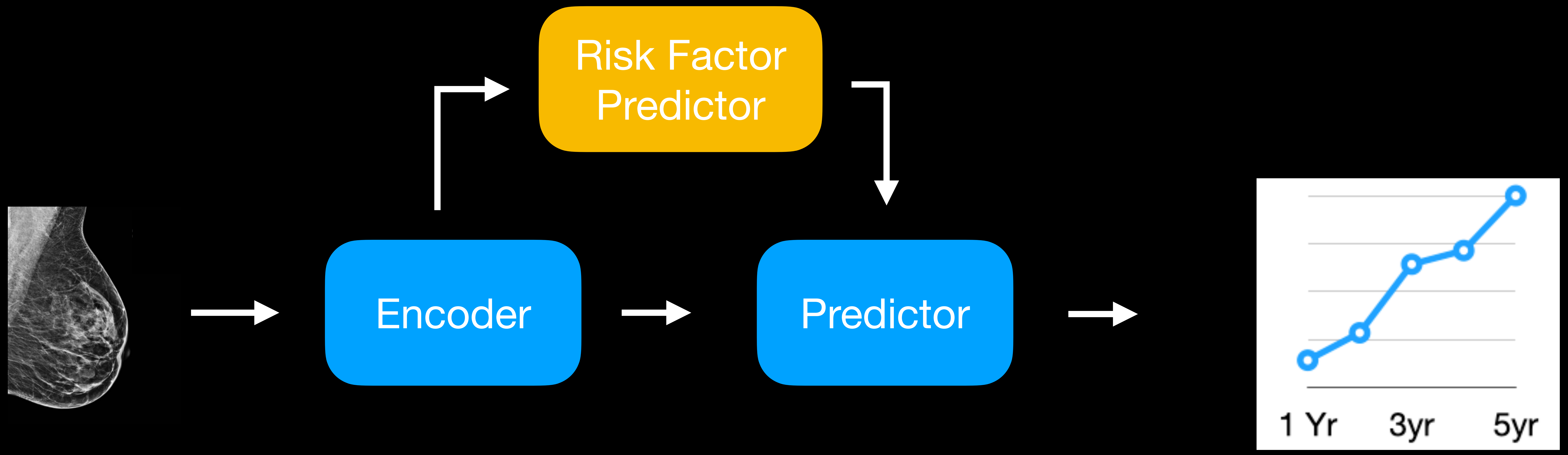


With Adversary

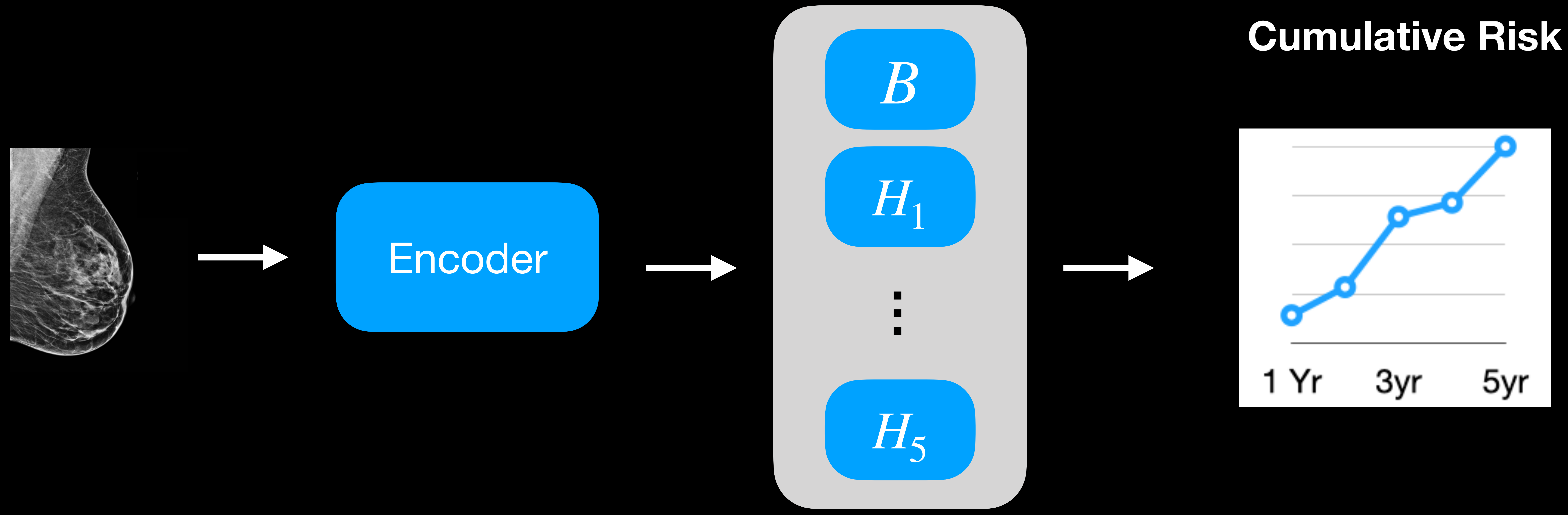
Problem 2: Missing risk factor data



Problem 2: Missing risk factor data

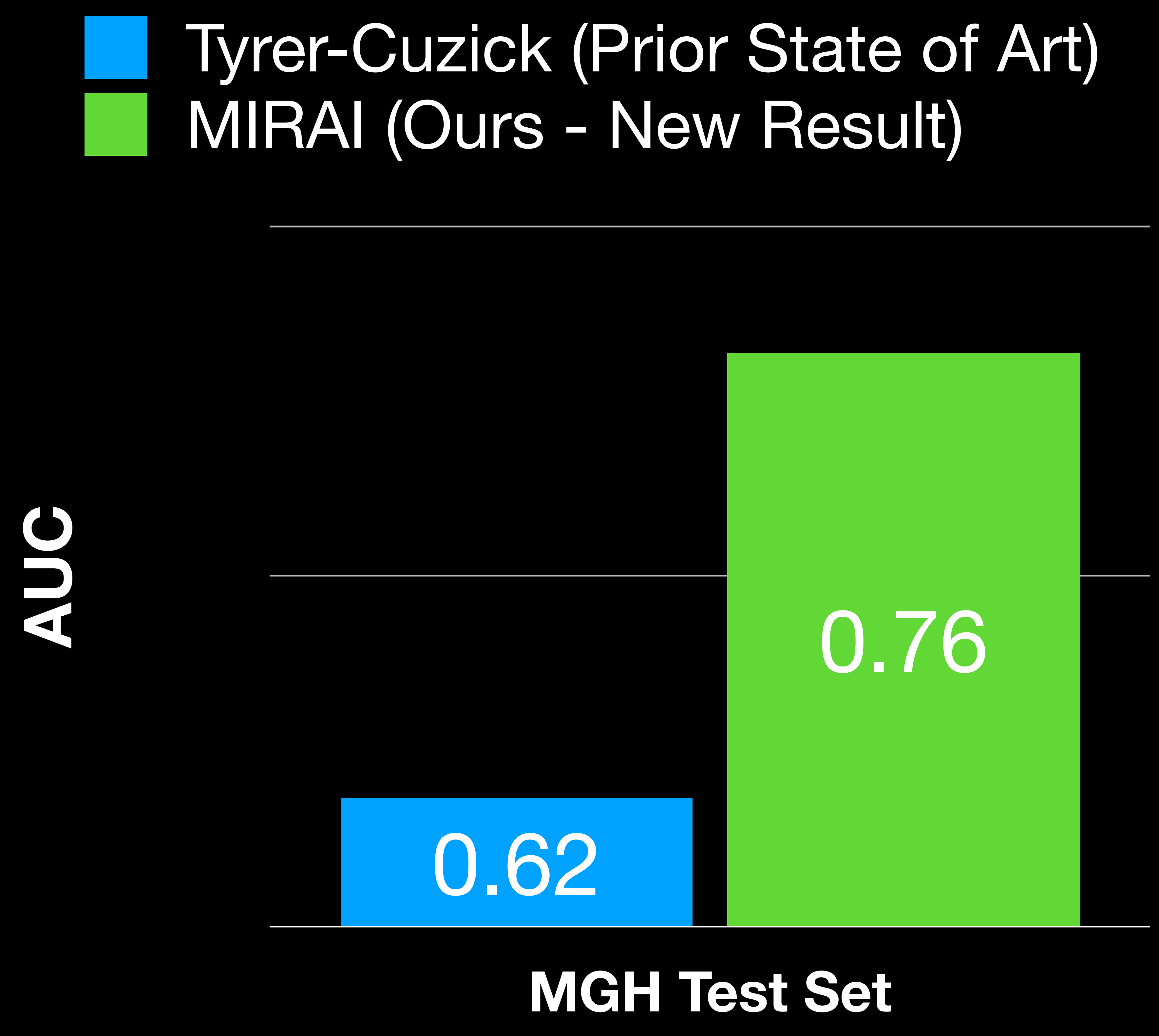


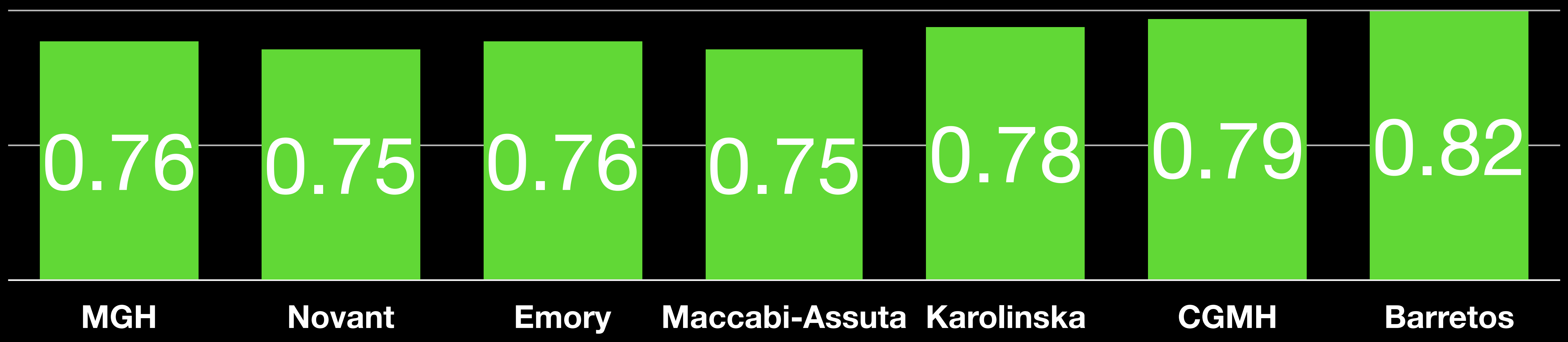
Problem 3: Modeling risk over time



$$P(t_{cancer} = k \mid x) = B(E(x)) + \sum H_i(E(x))$$

Maintains accuracy across diverse populations



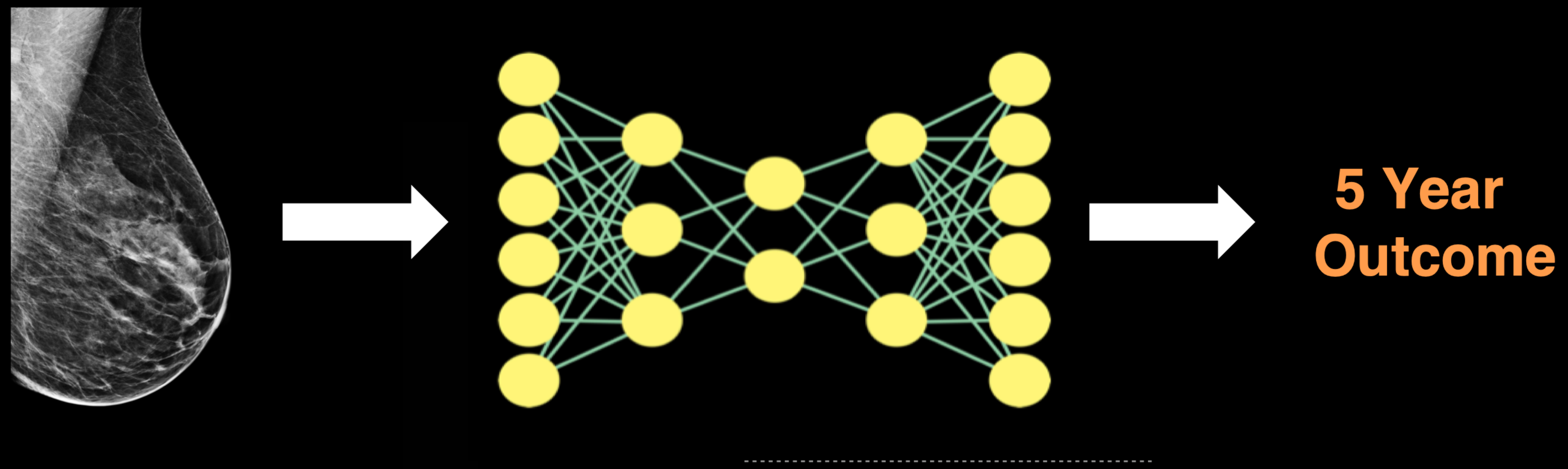


SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

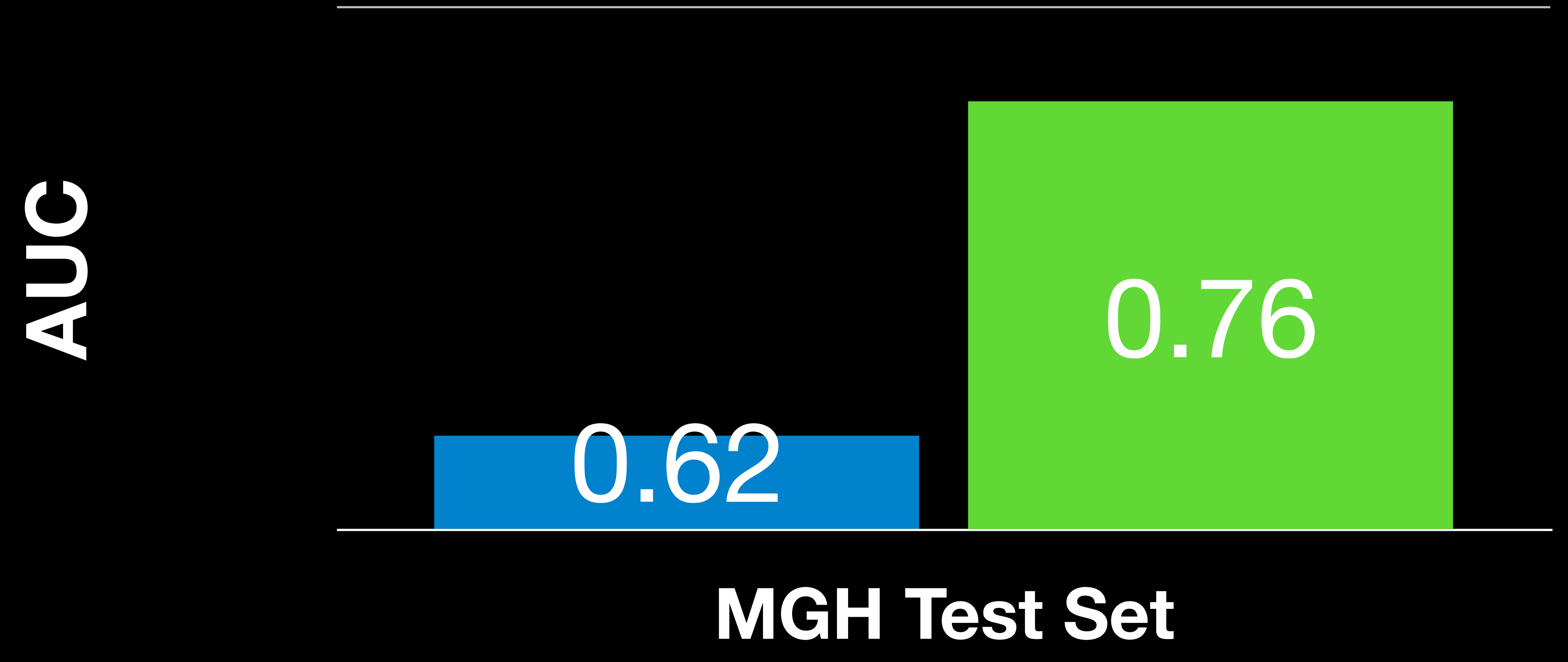
CANCER

Toward robust mammography-based models for breast cancer risk

Adam Yala[1,2*], Peter G. Mikhael[1,2], Fredrik Strand[3,4], Gigin Lin[5], Kevin Smith[6,7], Yung-Liang Wan[5], Leslie Lamb[8], Kevin Hughes[9], Constance Lehman[8†], Regina Barzilay[1,2†]



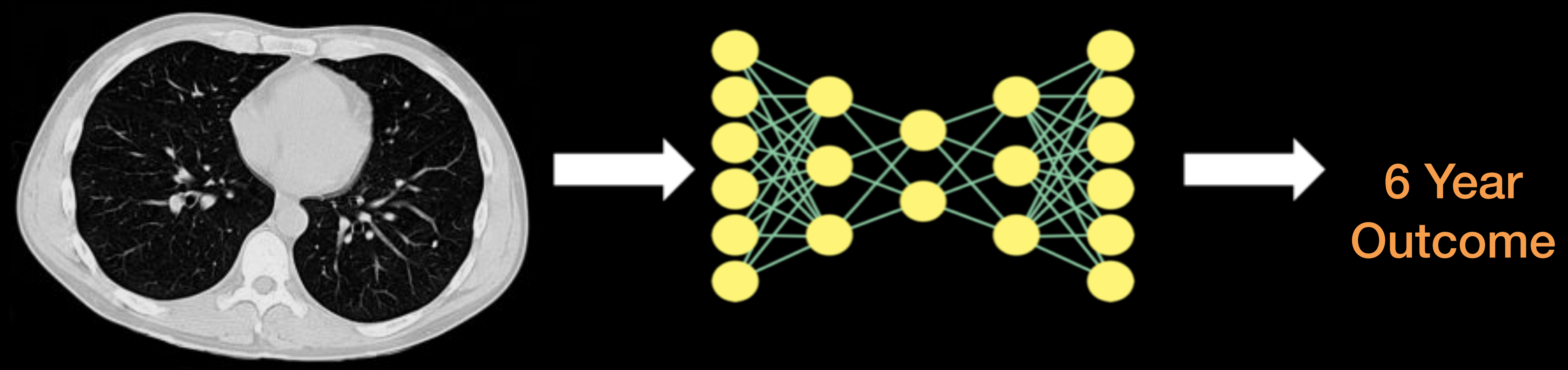
Tyrer-Cuzick (Prior State of Art)
MIRAI (Ours - New Result)



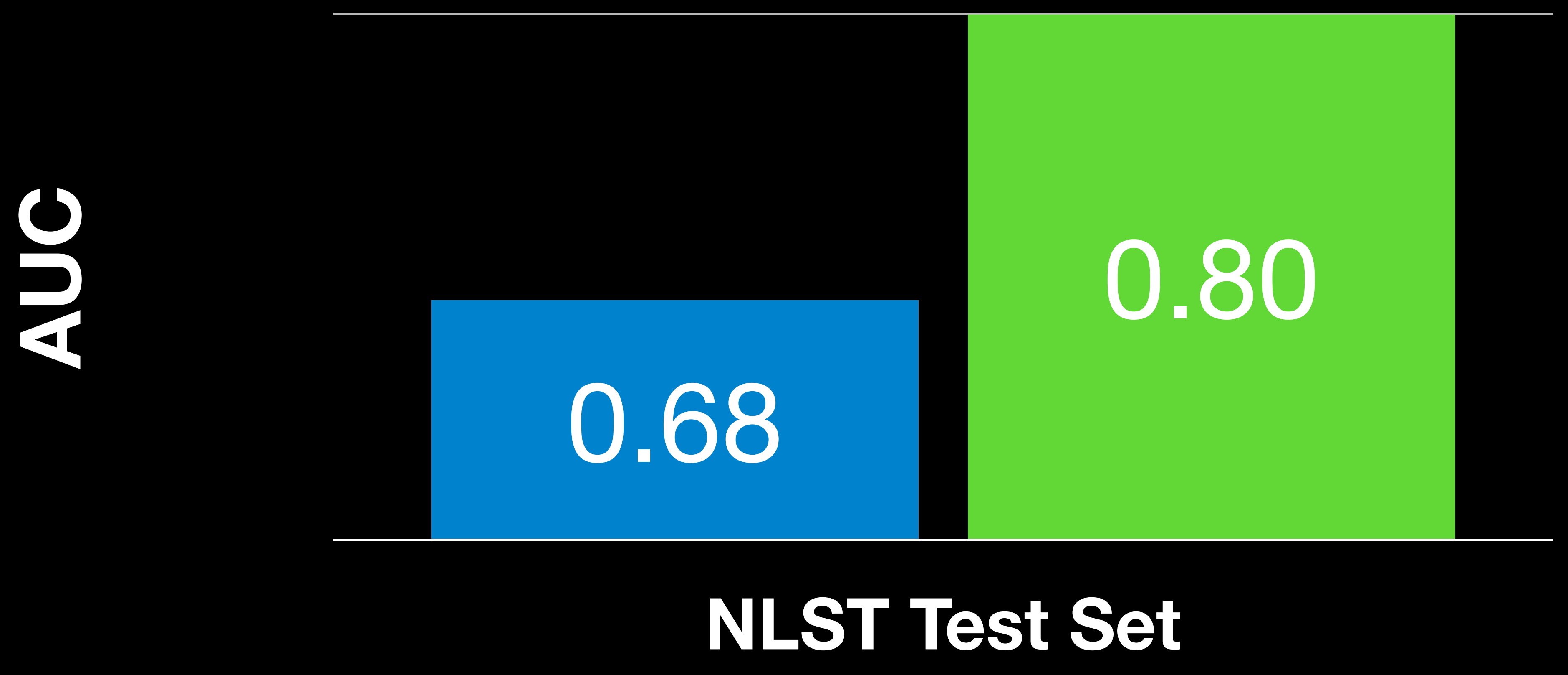
Under Review

Ask Sybil: Predicting Lung Cancer Risk with Low-dose Chest Computed Tomography

Peter G. Mikhael[1,2,†,*], Jeremy Wohlwend[1,2,†], Adam Yala[1,2], Justin Xiang[1,2], Angelo K. Takigami[3,4], Patrick P. Bourgouin[3,4], PuiYee Chan[5], Sofiane Mrah[4], Lecia V. Sequist[3,5], Florian J. Fintelmann[3,4,‡], Regina Barzilay



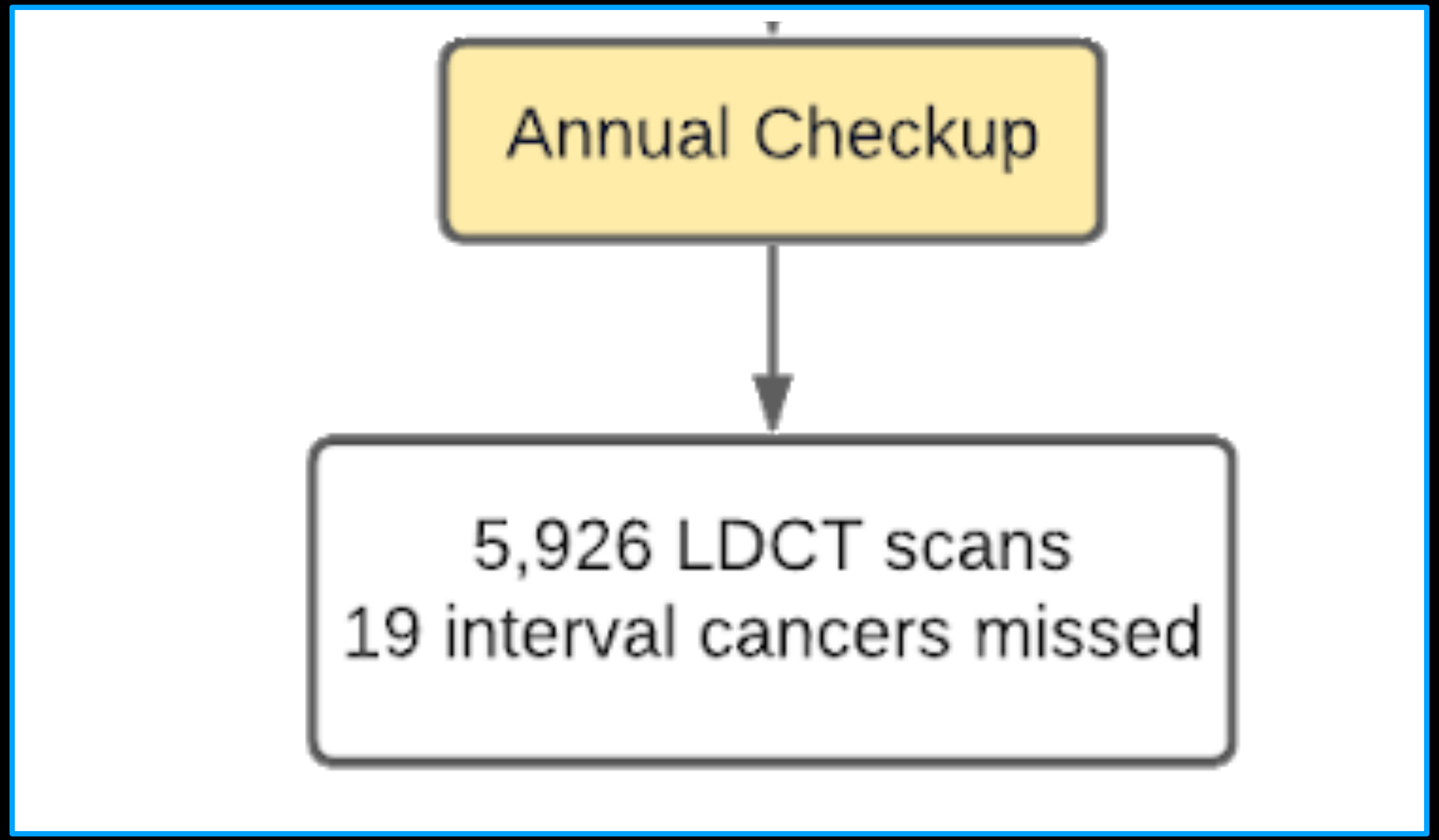
PLCOm2012 (Prior State of Art)
Sybil (Ours - New Result)



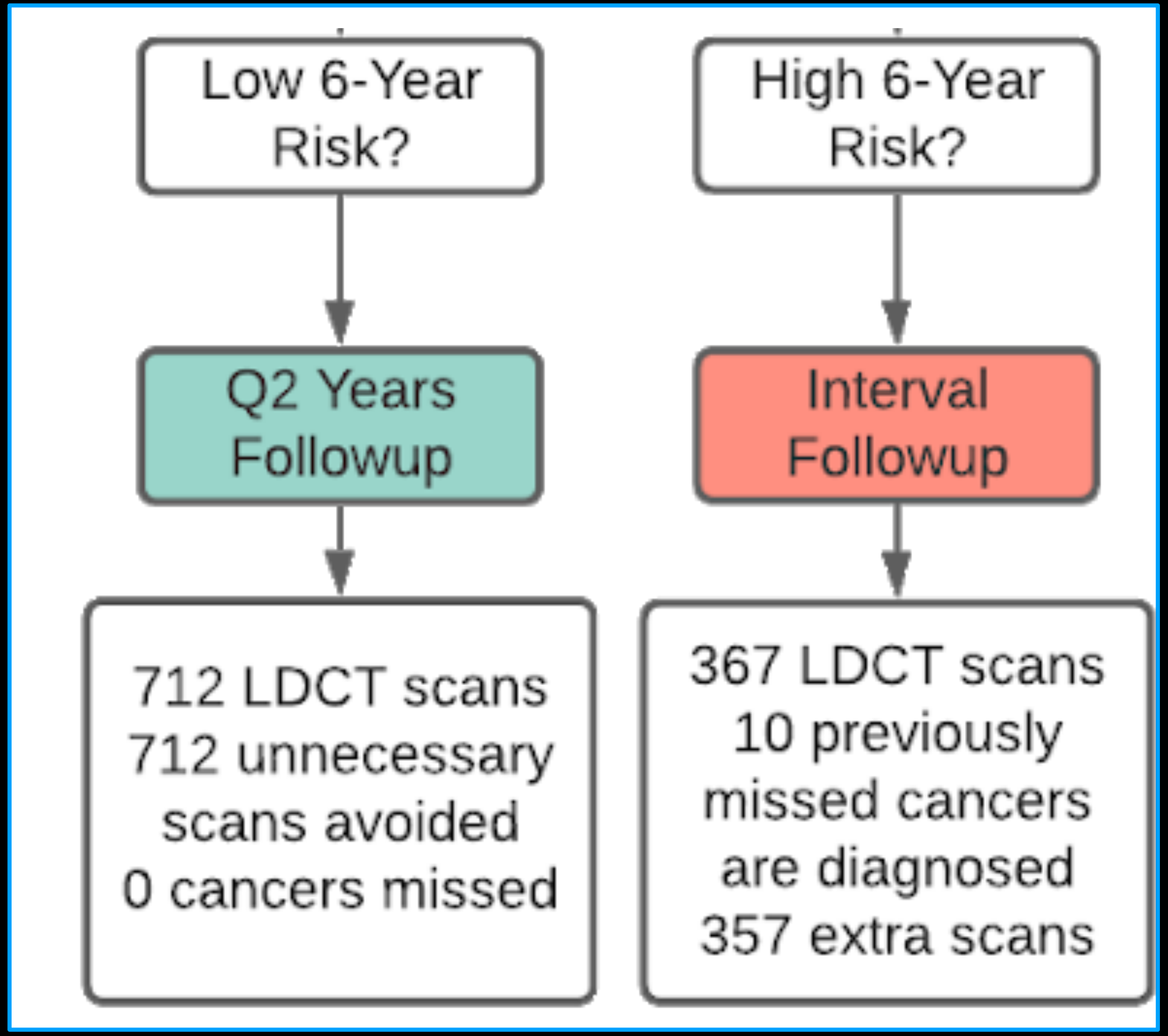
Sybil Clinical Impact : Workflow

Improve early detection and lower screening cost with risk-based followup.

Standard of Care

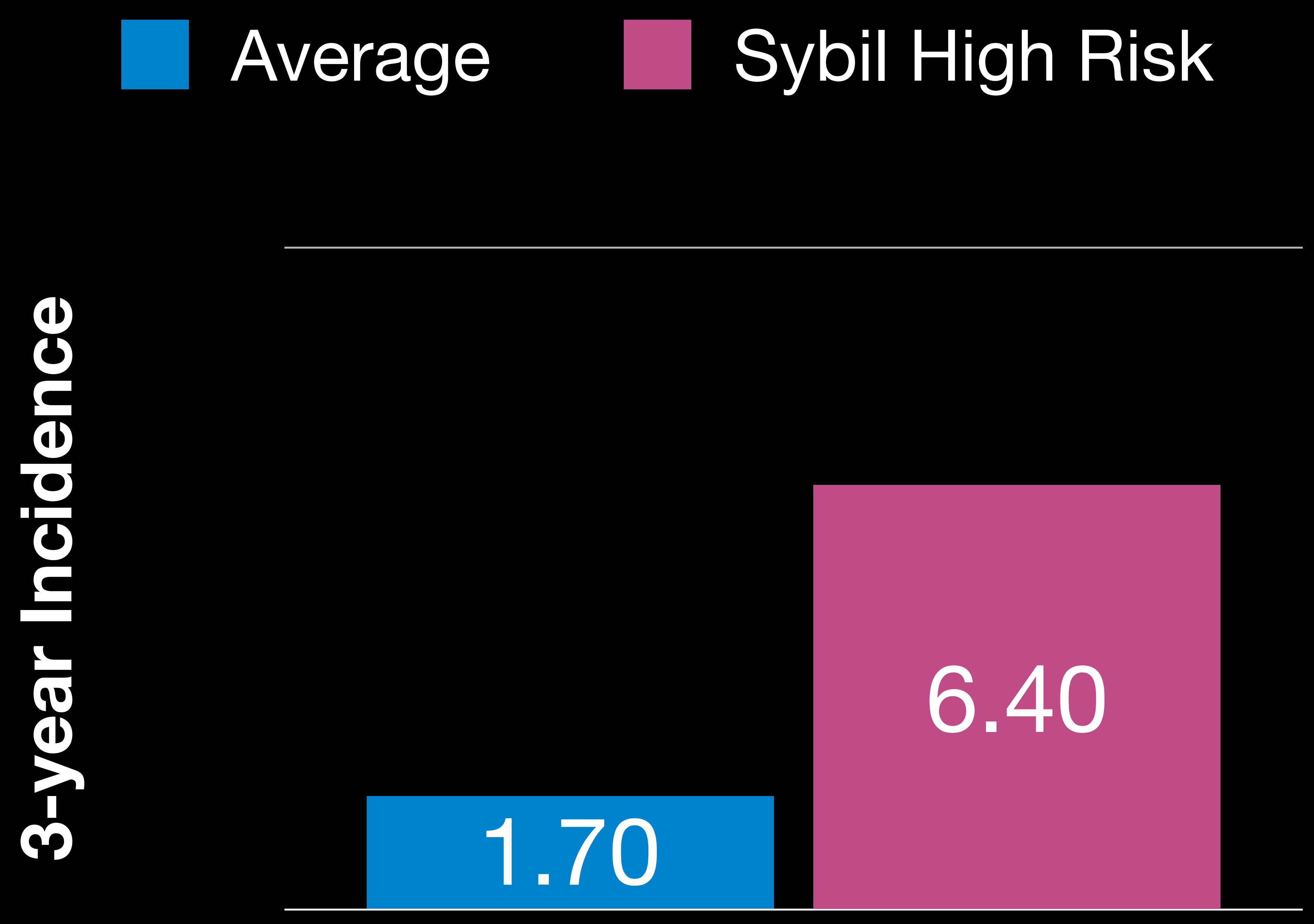


Sybil risk based screening



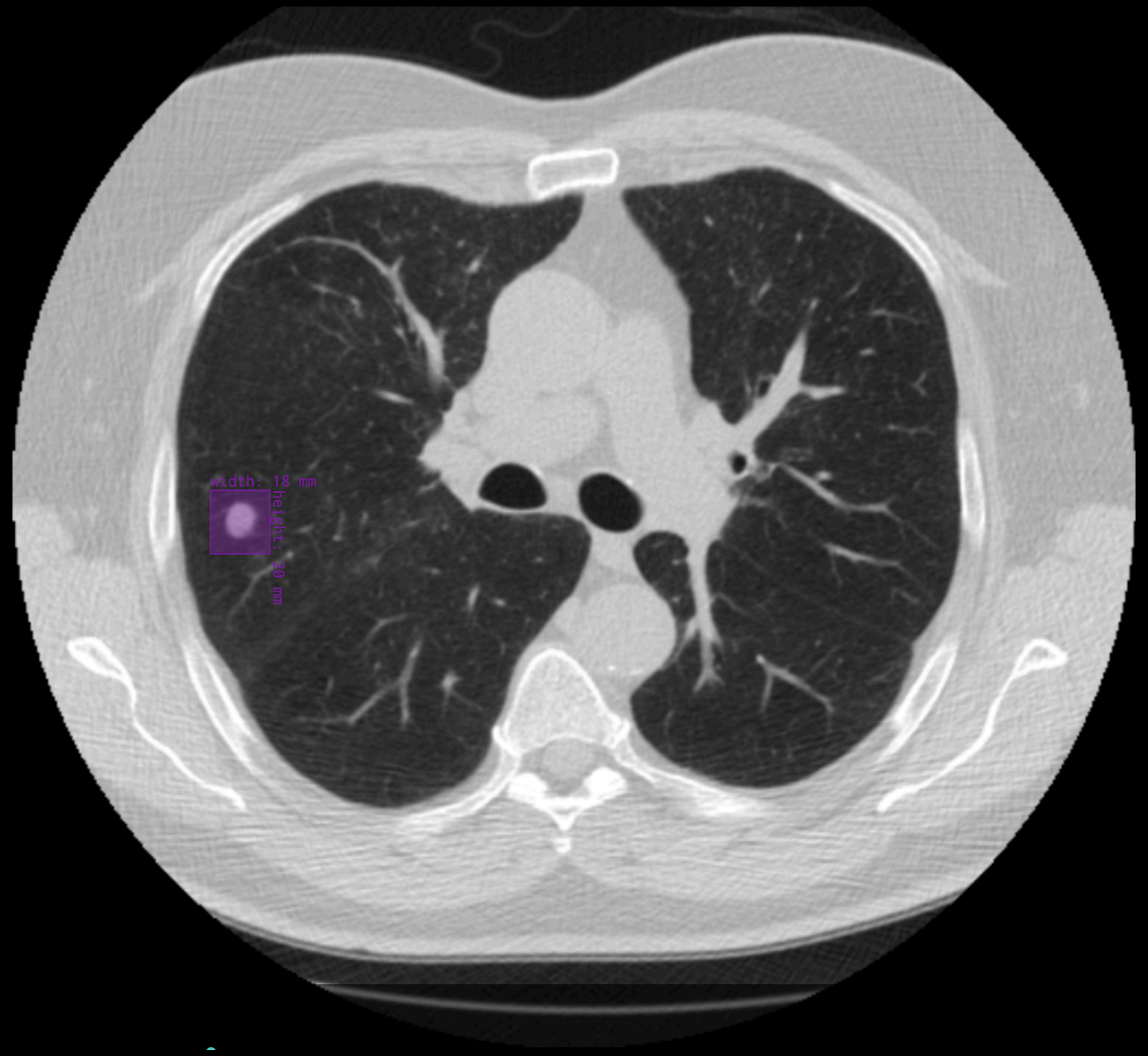
Sybil Clinical Impact : Prevention

Identifying high risk cohorts for clinical trails



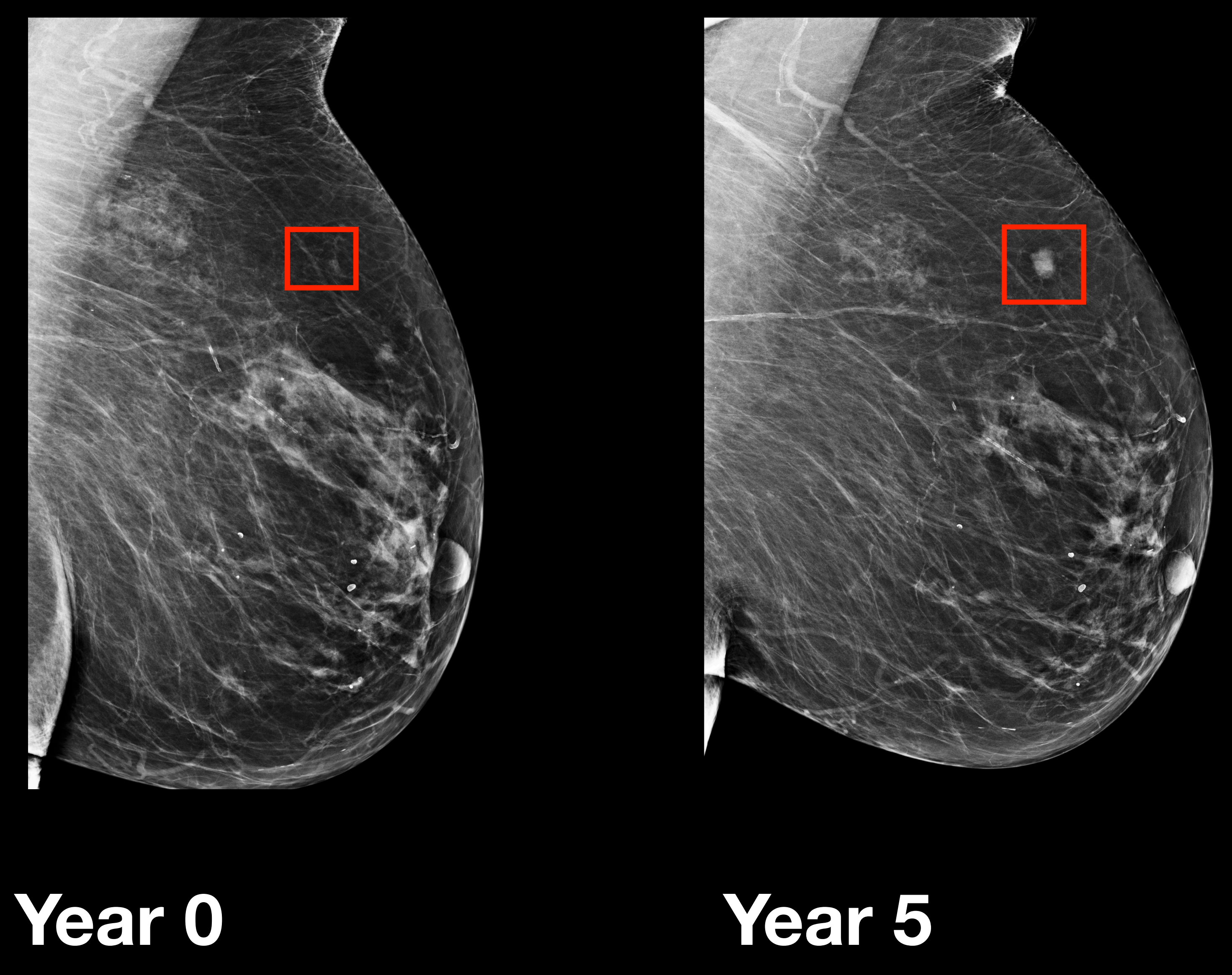
Identifying future cancer location

83% accurate predicting future cancer side

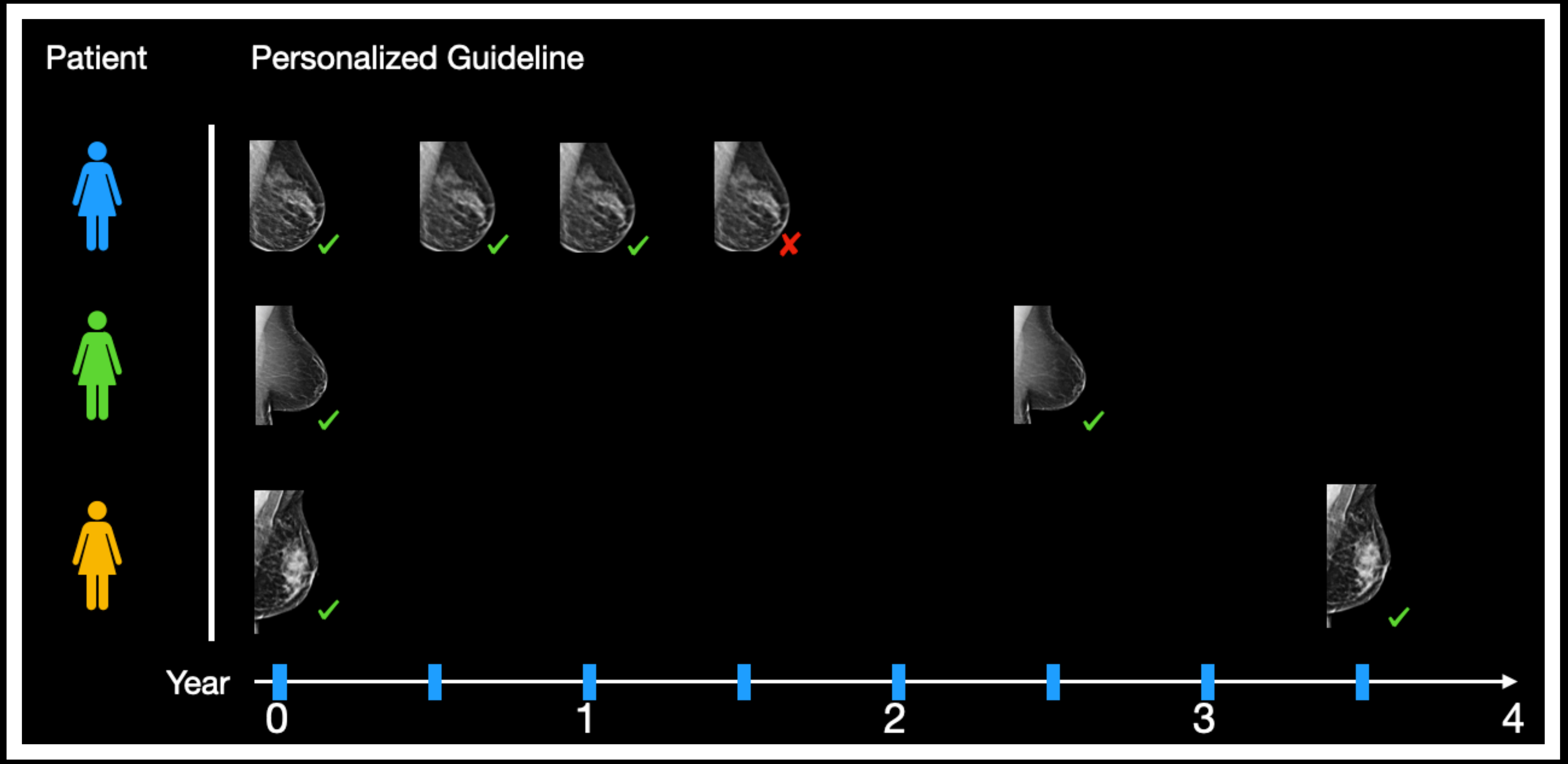


How to catch cancer earlier

Predict Cancer Risk



Create personalized screening policy



How to catch cancer earlier

nature medicine

ARTICLES

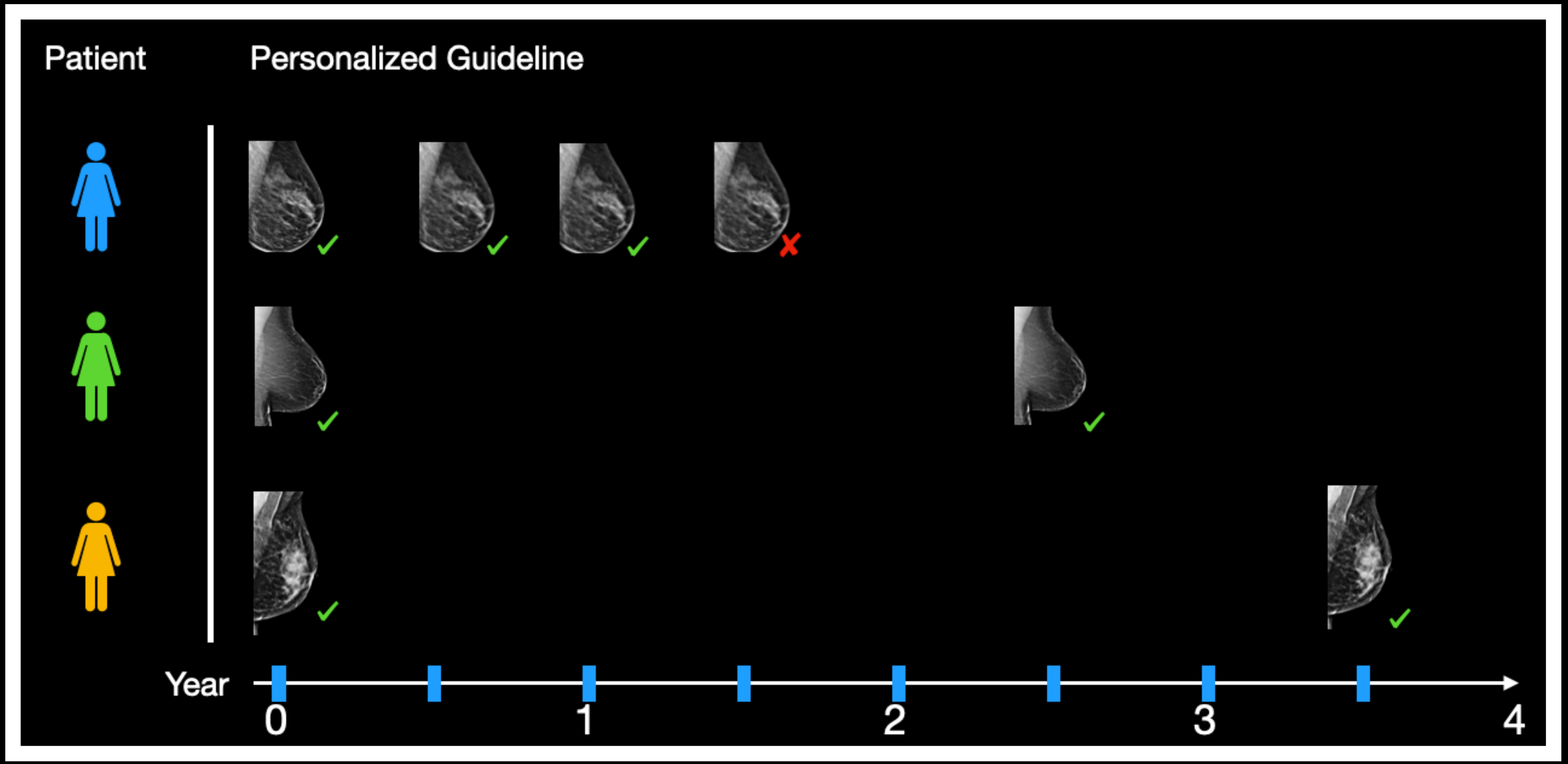
https://doi.org/10.1038/s41591-021-01599-w

Check for updates

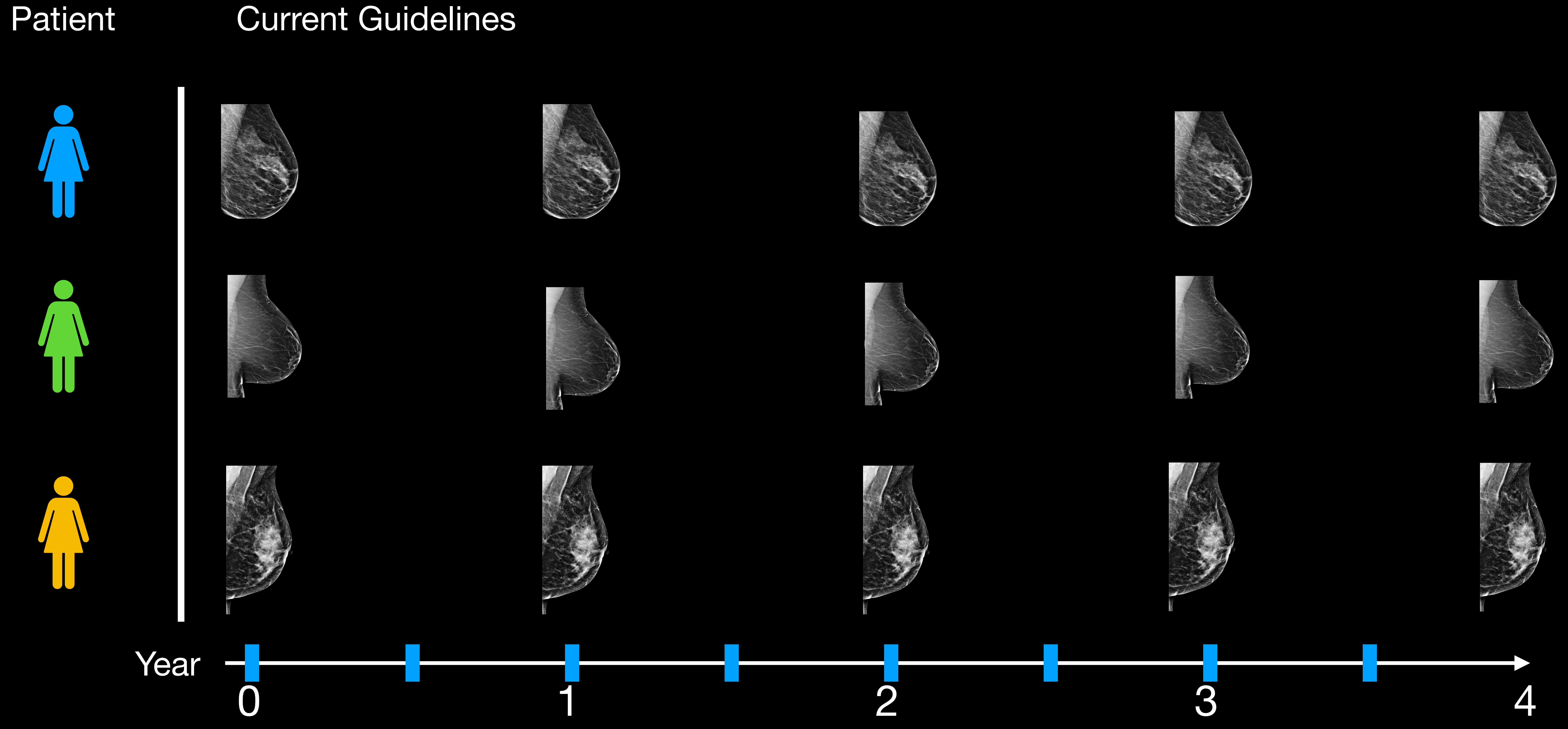
Optimizing risk-based breast cancer screening policies with reinforcement learning

Adam Yala[1,2], Peter G. Mikhael[1,2], Constance Lehman[3], Gigin Lin[4,5], Fredrik Strand[6,7], Yung-Liang Wan[4,5], Kevin Hughes[8], Siddharth Satuluru[9], Thomas Kim[10], Imon Banerjee[11], Judy Gichoya[9], Hari Trivedi[9] and Regina Barzilay[1,2]

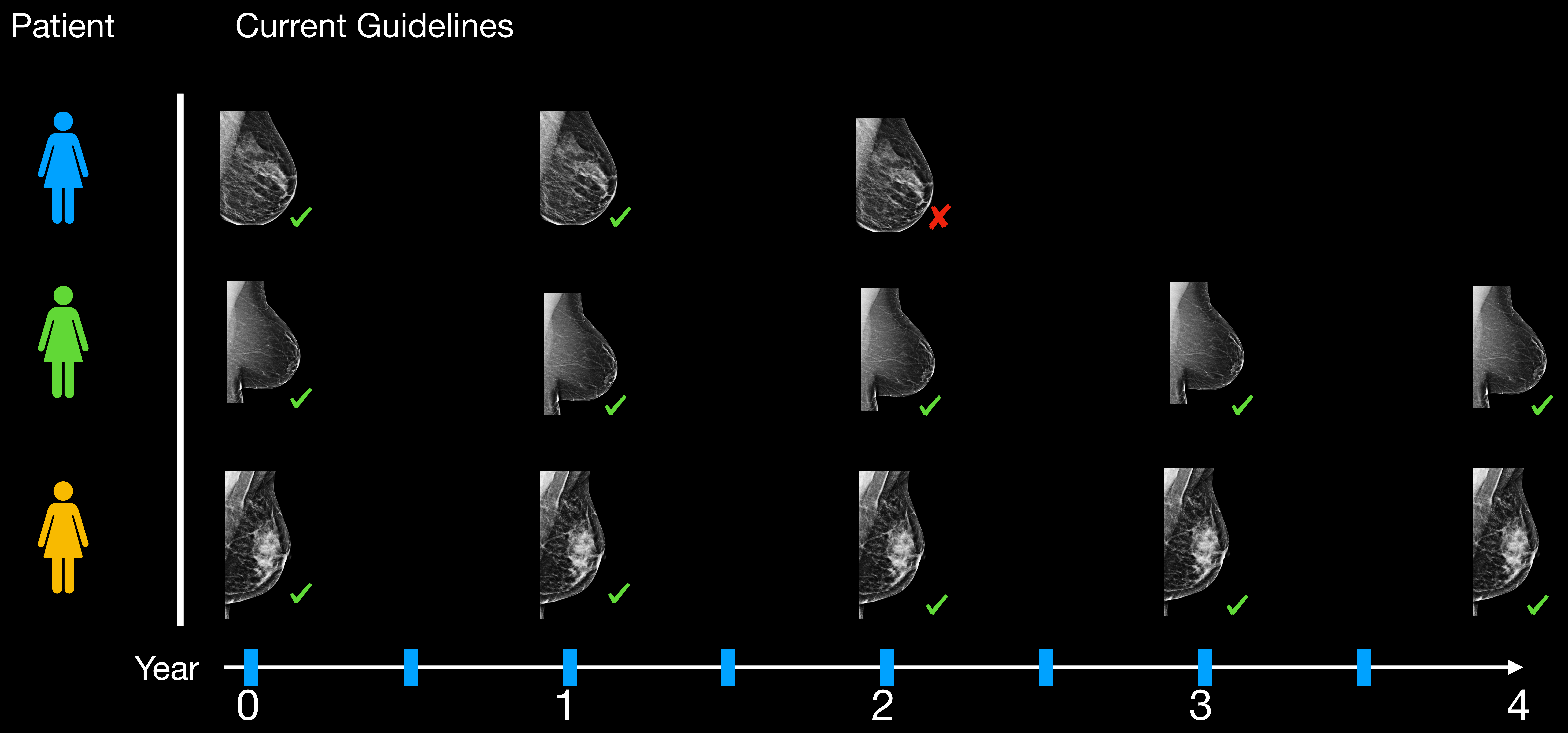
Create personalized screening policy



Screening today

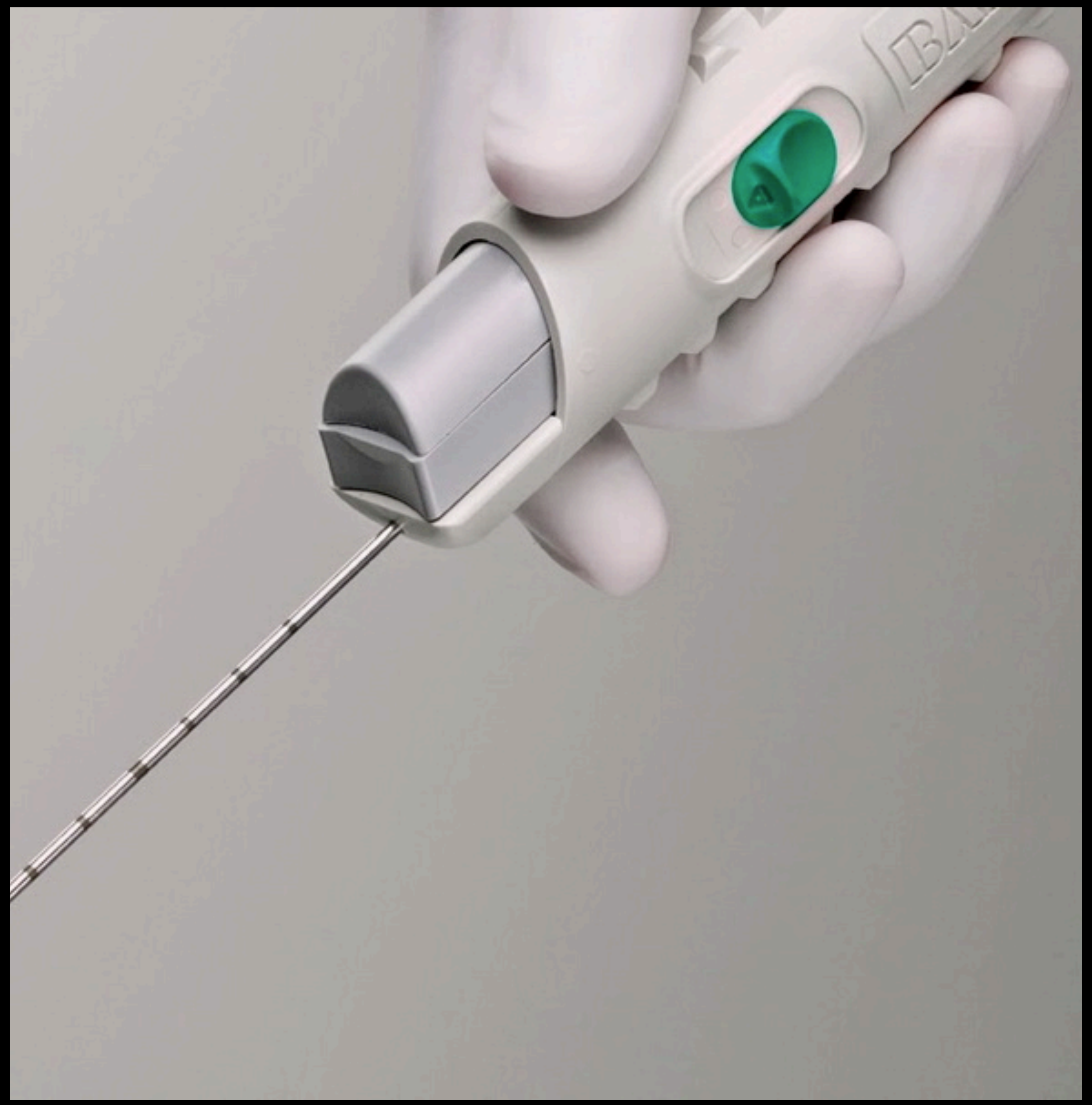


Same screening, different outcomes



Challenges in current screening

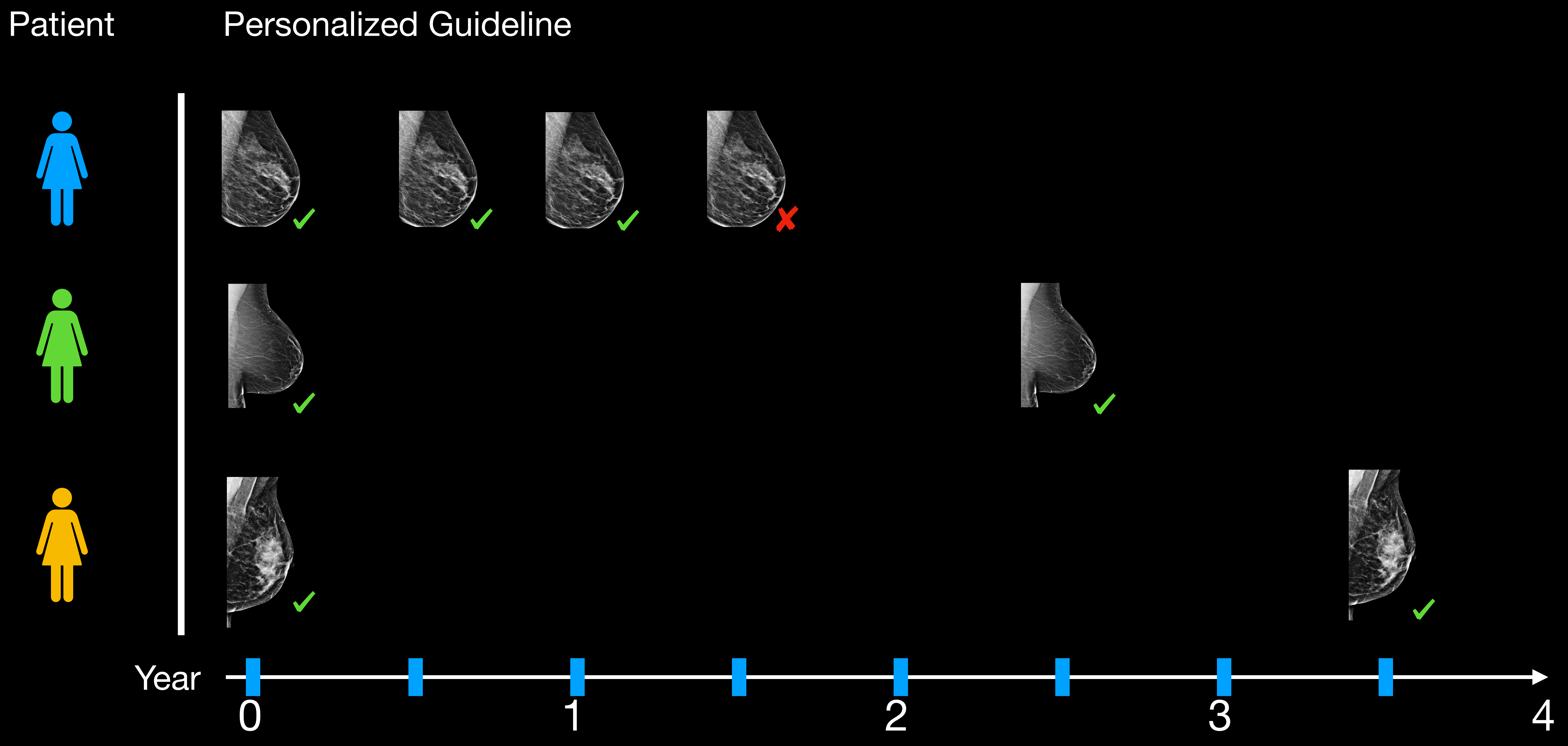
Late Detection



Over Screening

Health Disparities

Tailor screening regime to patient need



Current processes for policy design

National Comprehensive Cancer Network®

U.S. Preventive Services TASK FORCE

ACR® AMERICAN COLLEGE OF RADIOLOGY

…

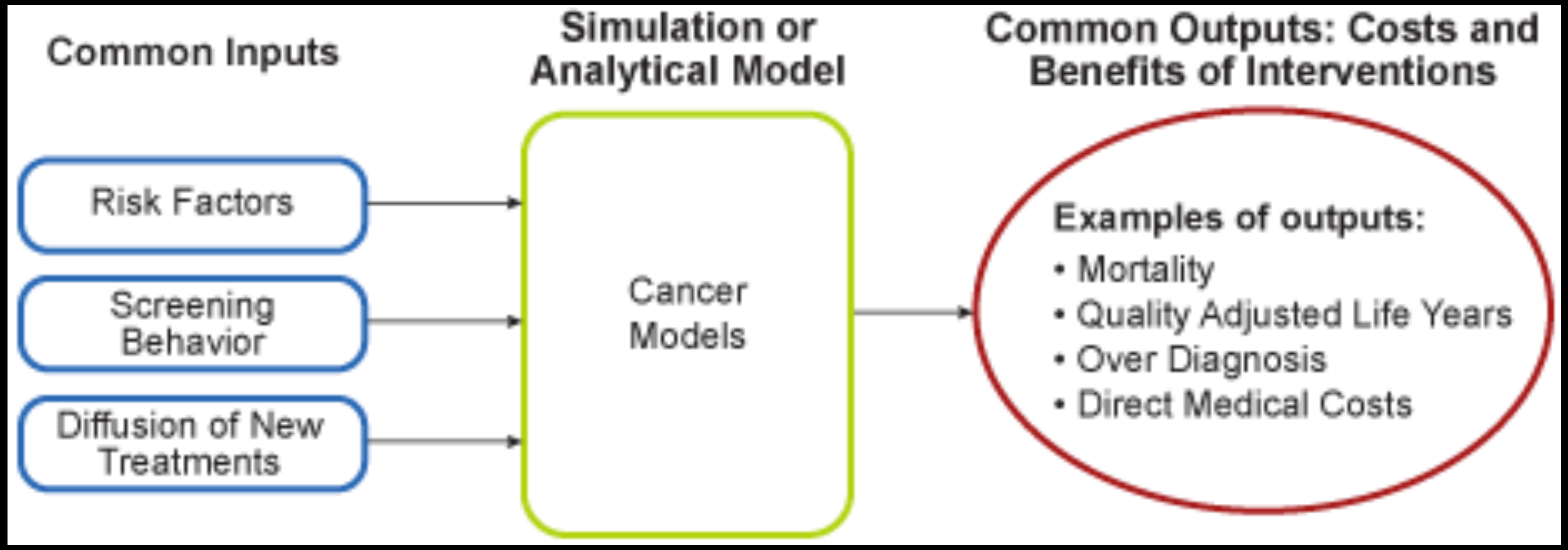
Expert panel meetings by physician organizations

Most meet every five years

Multiple conflicting one-size fits all guidelines

No explicit validation across populations. Health disparities grow

Modeling for clinical guidelines



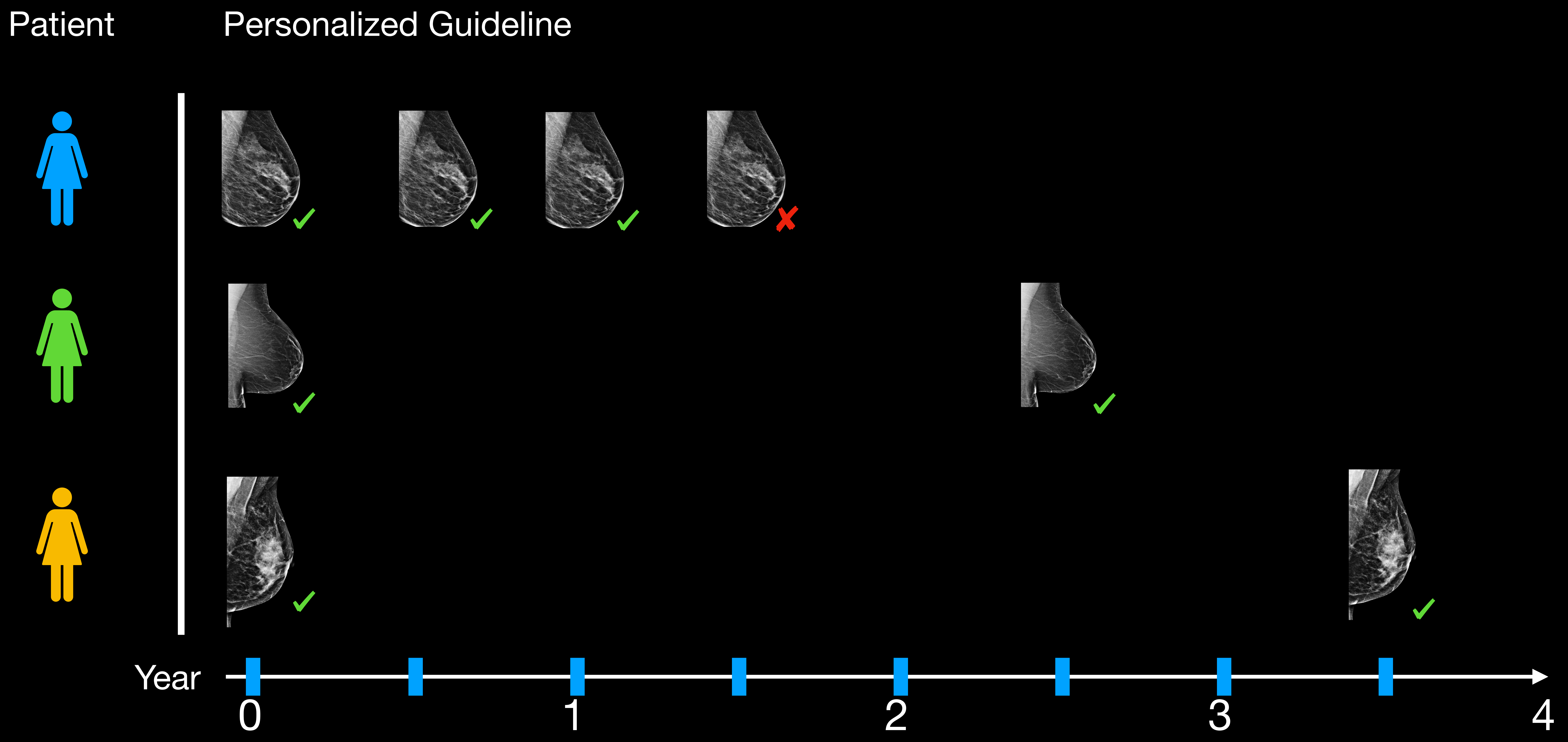
Assume full probabilistic models of disease

Simulate hypothetical patients under screening strategies

Suggest “correct” trade-off

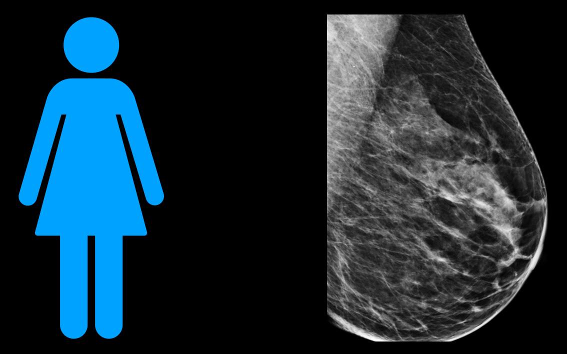
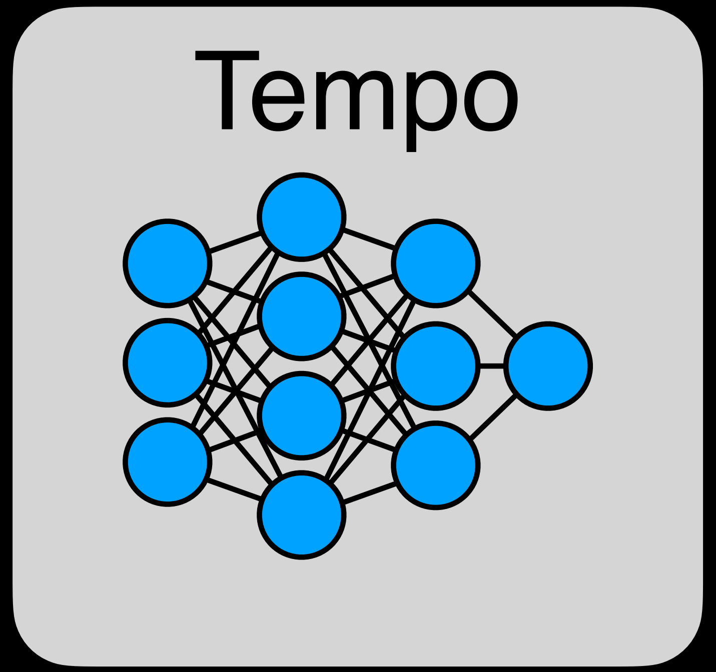
Cannot incorporate new risk models or evaluated on real patients

Tailor screening regime to patient need



Policy Design as a Learning Task

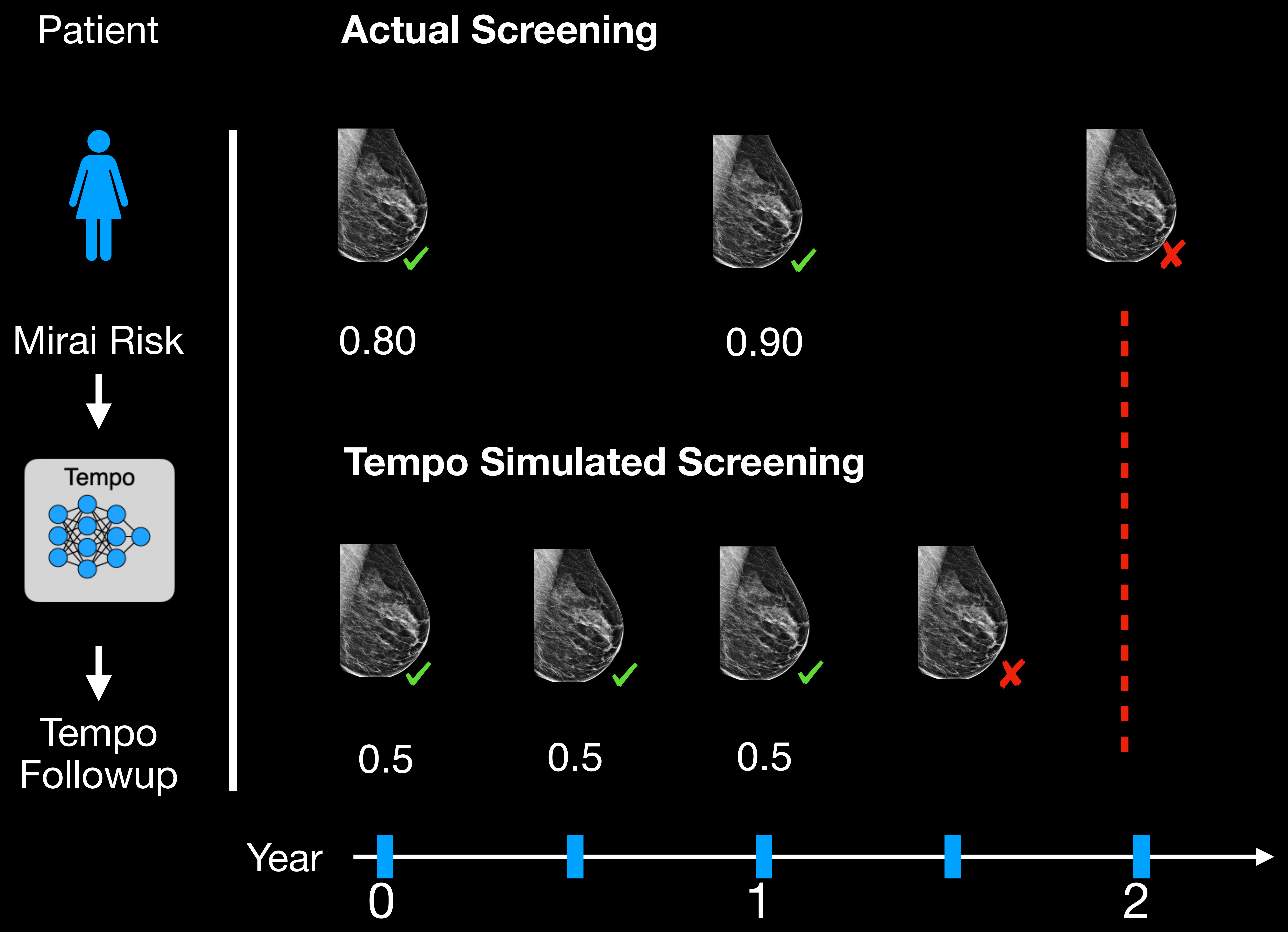
Patient **Risk**

 → 0.80 →  → **0.5 year followup**

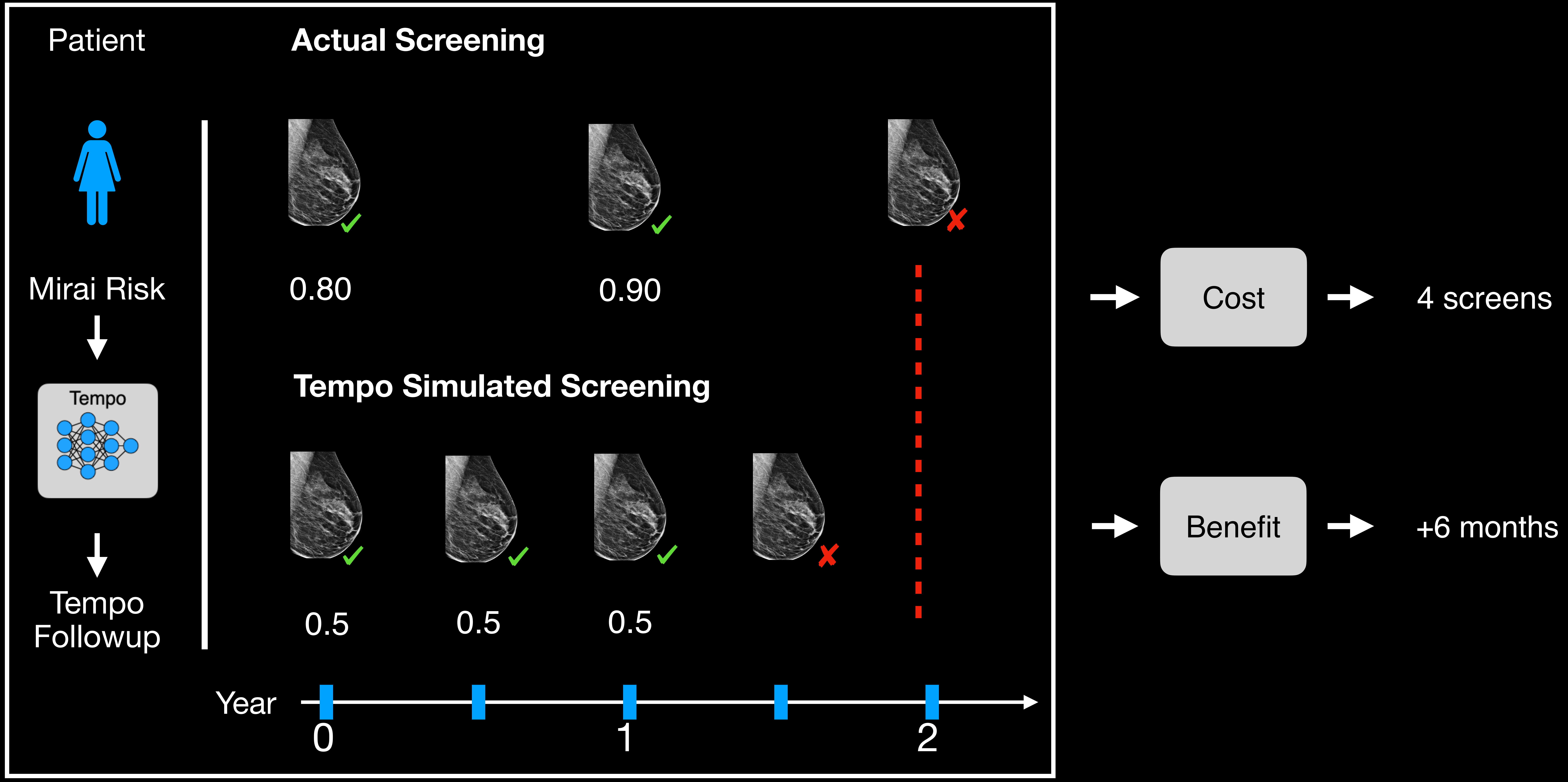
Reward = λ_1 Early Detection Benefit - λ_2Screening Cost

Desiderata: *Testable, Adaptive, Flexible*

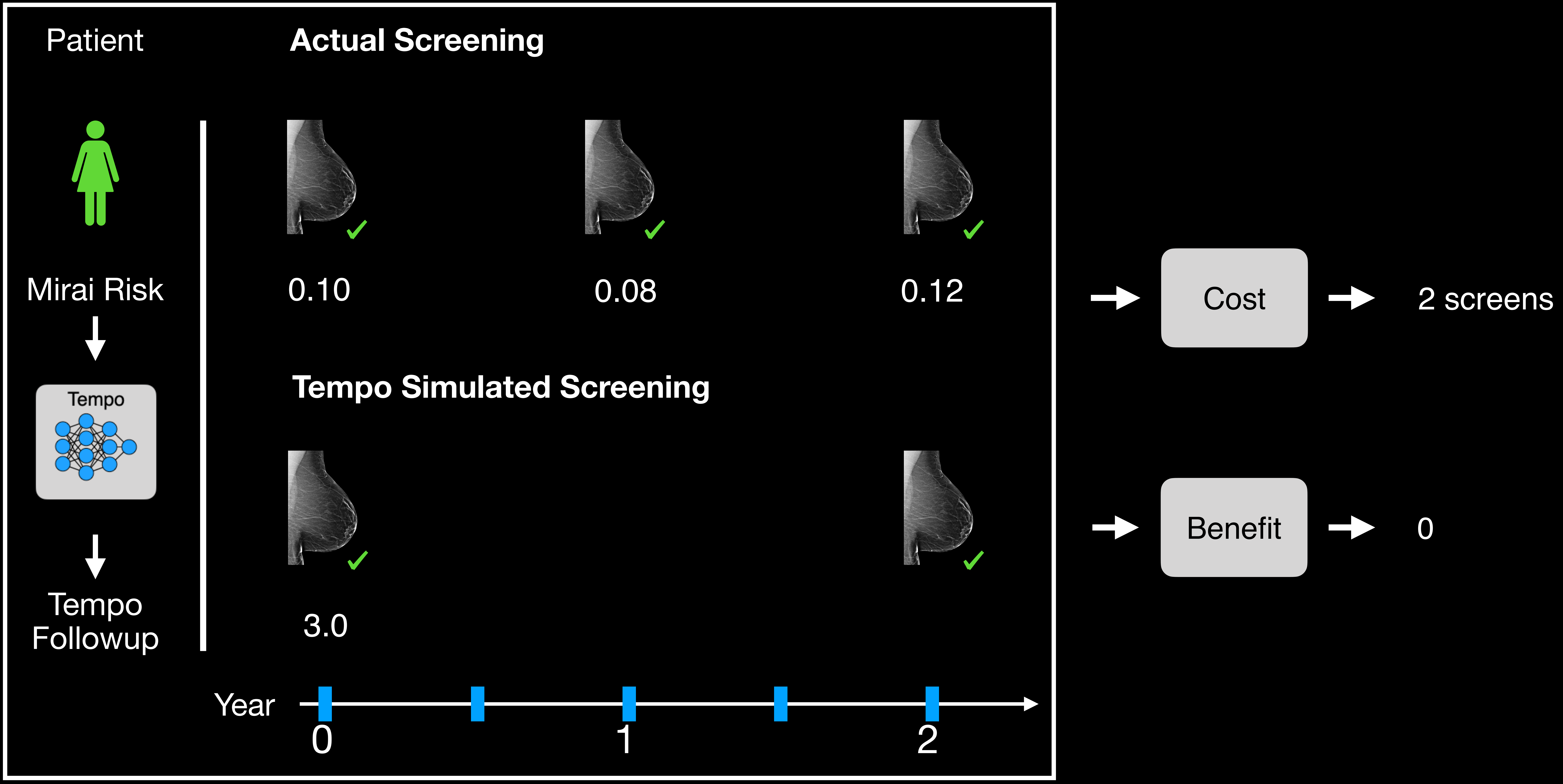
Simulate patient trajectories



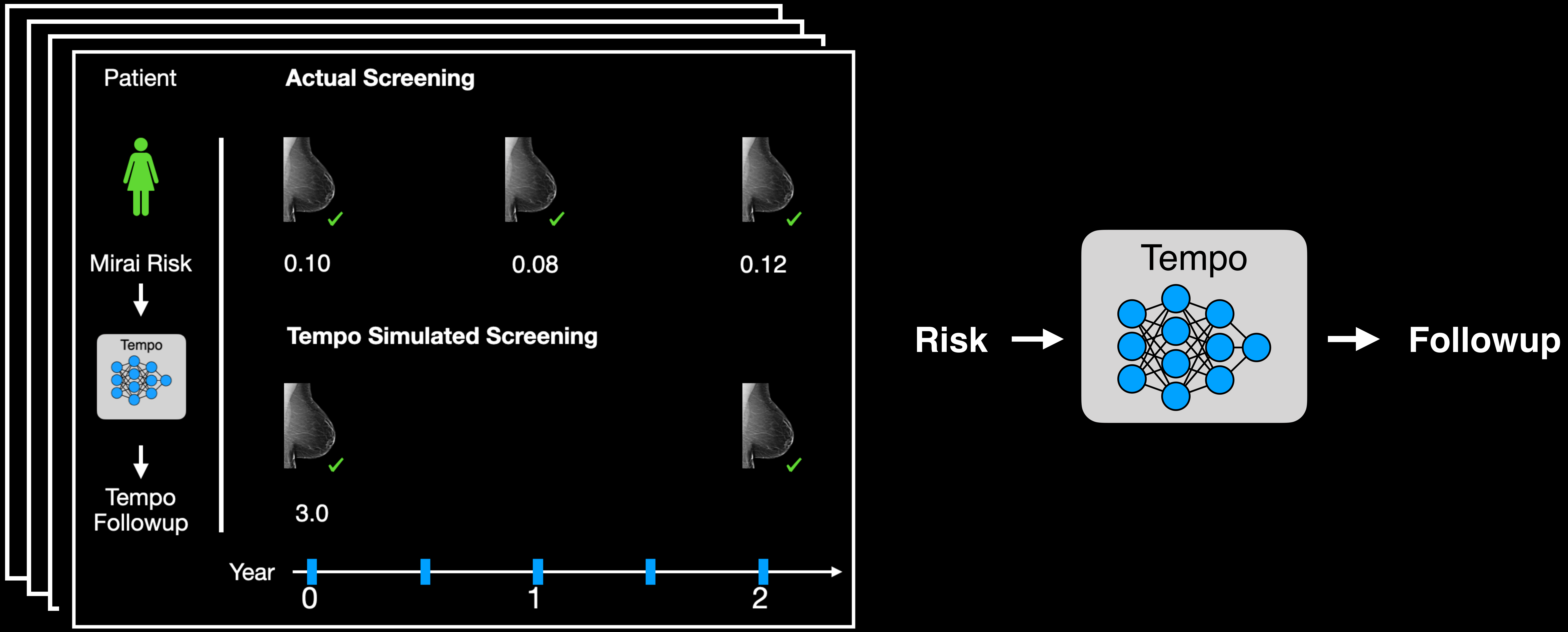
Evaluate individual impact



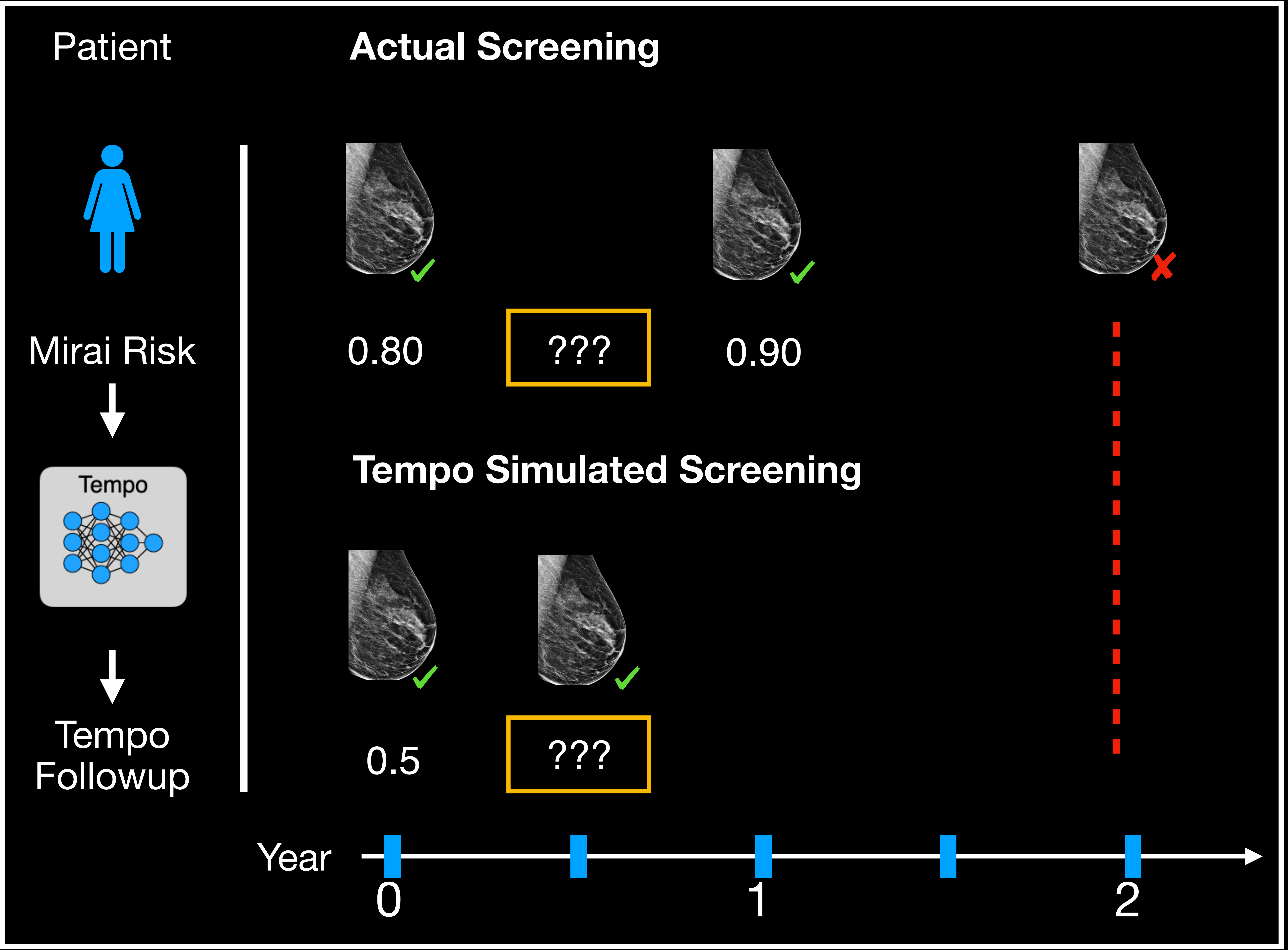
Evaluate individual impact



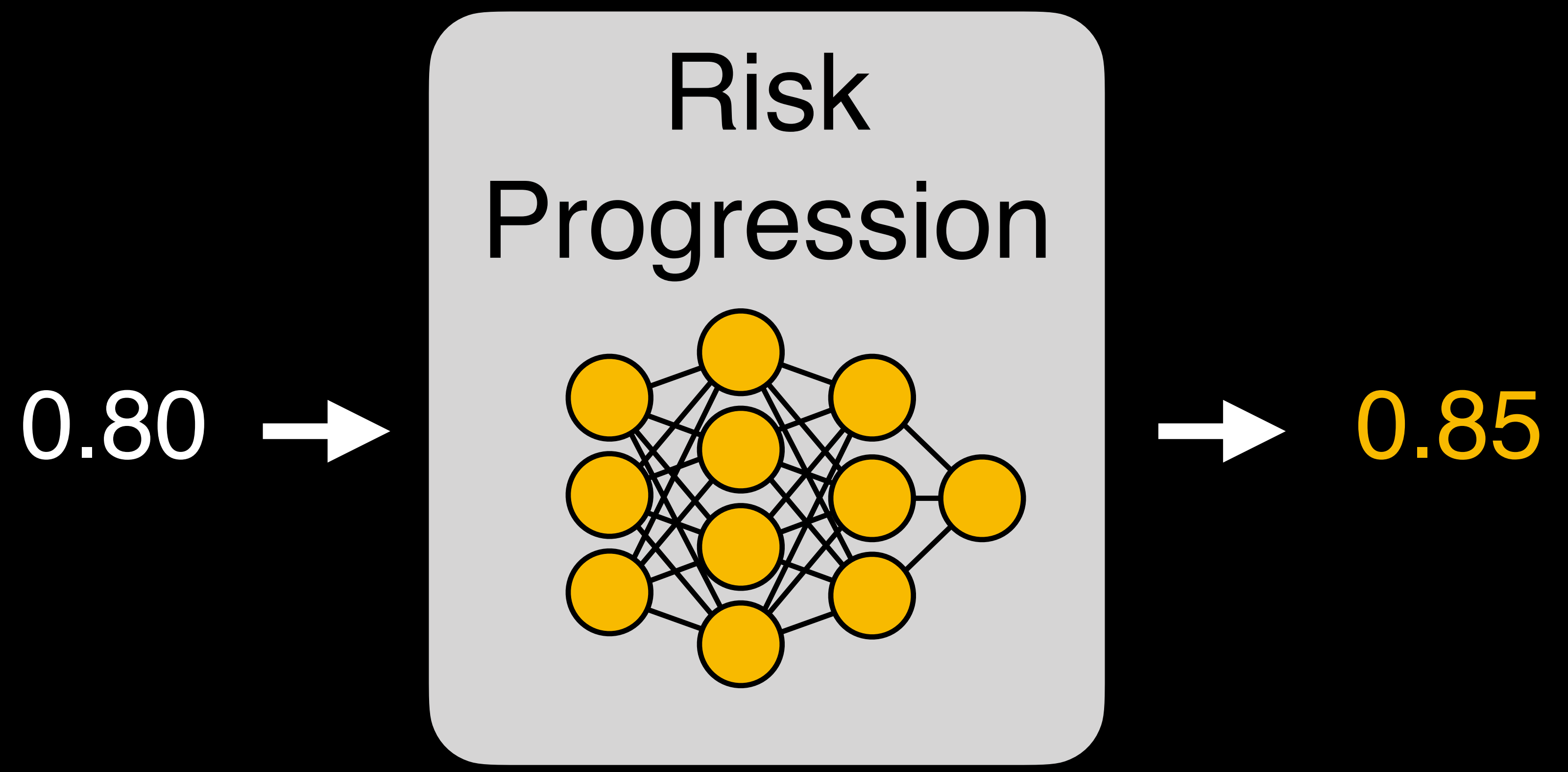
Optimized over population screening



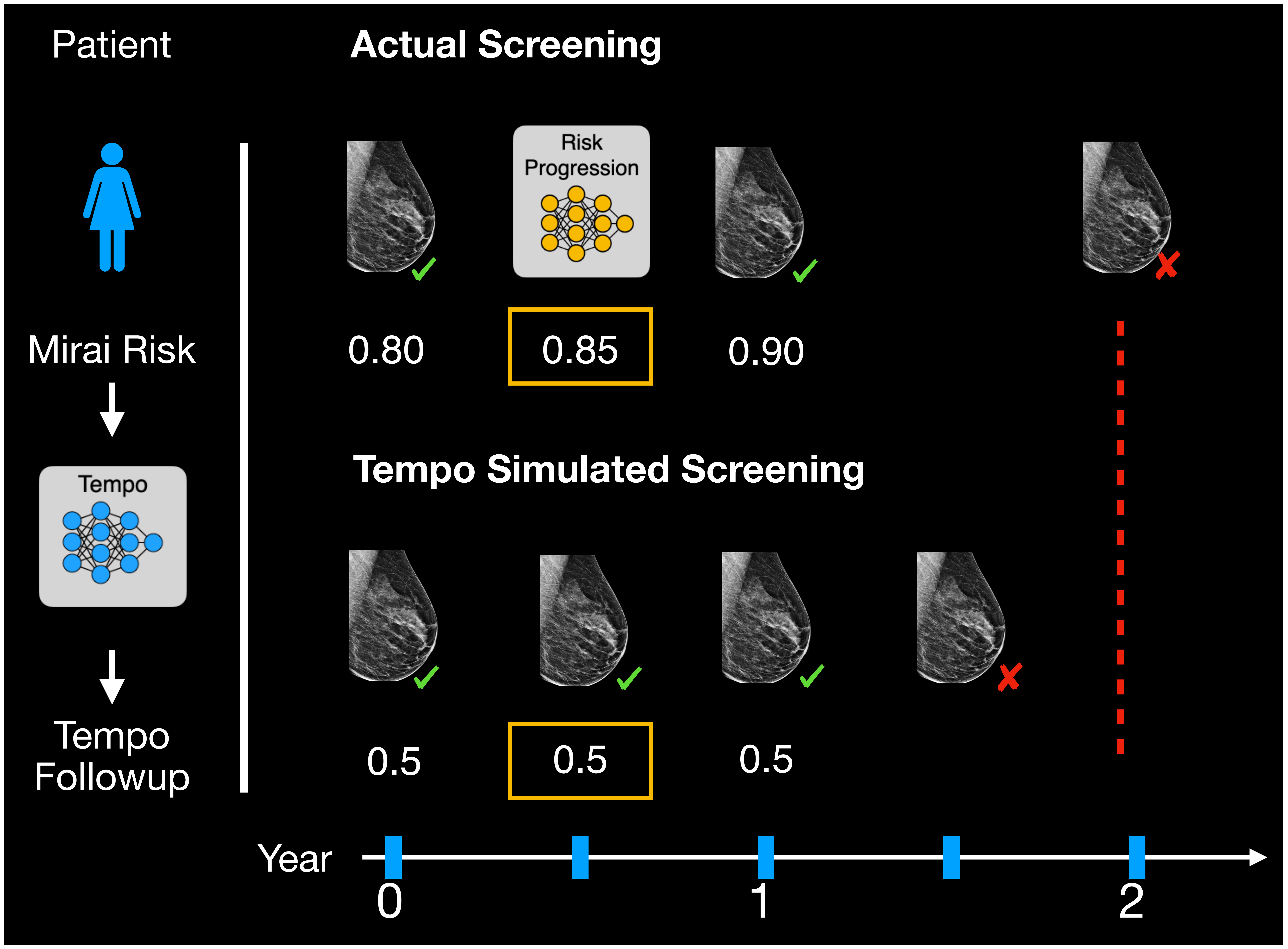
Estimating missing risk assessments



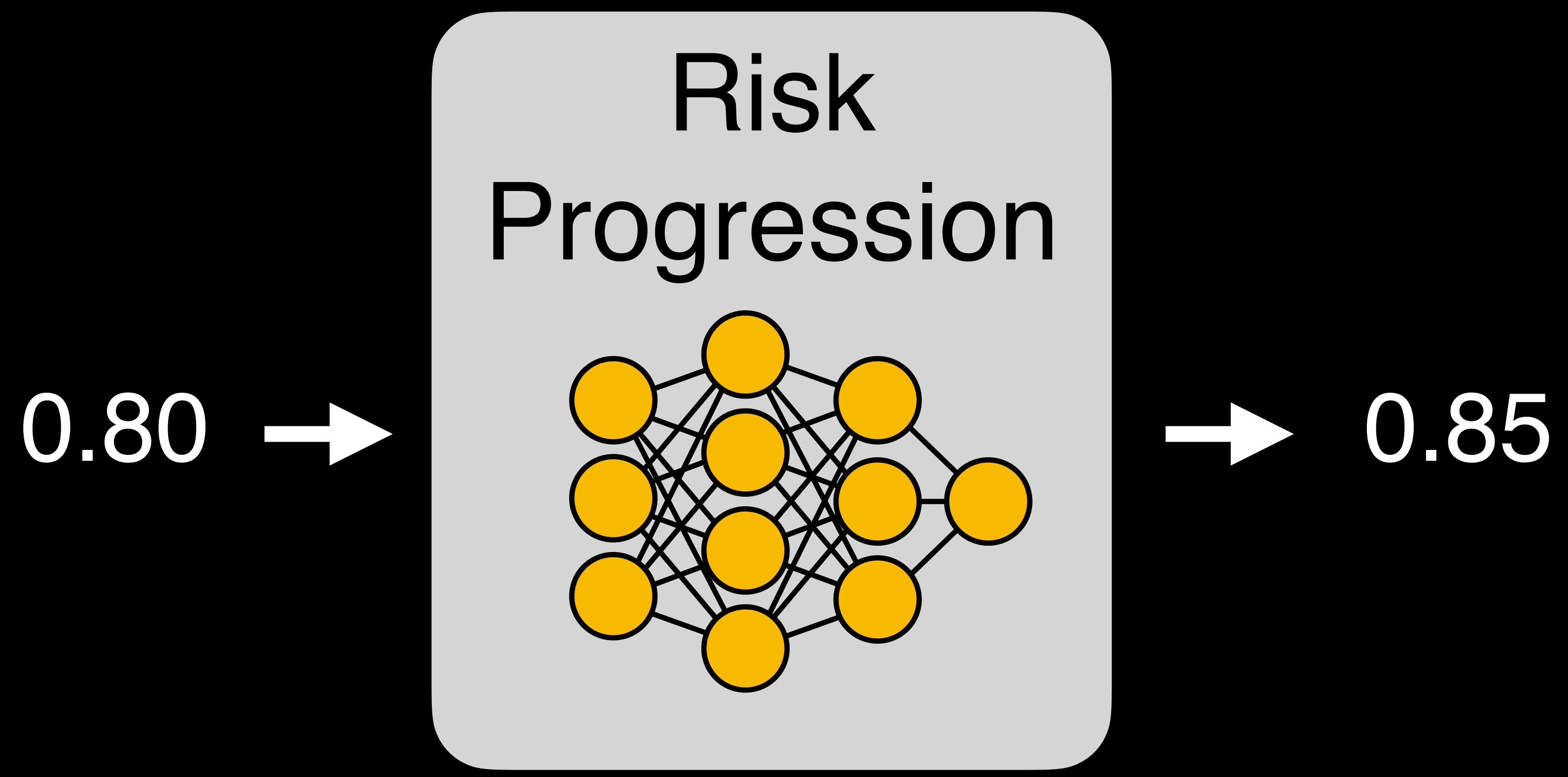
Learn $P(r_t \,|\, r_{t-1}, r_{t-2}, \ldots, r_0)$



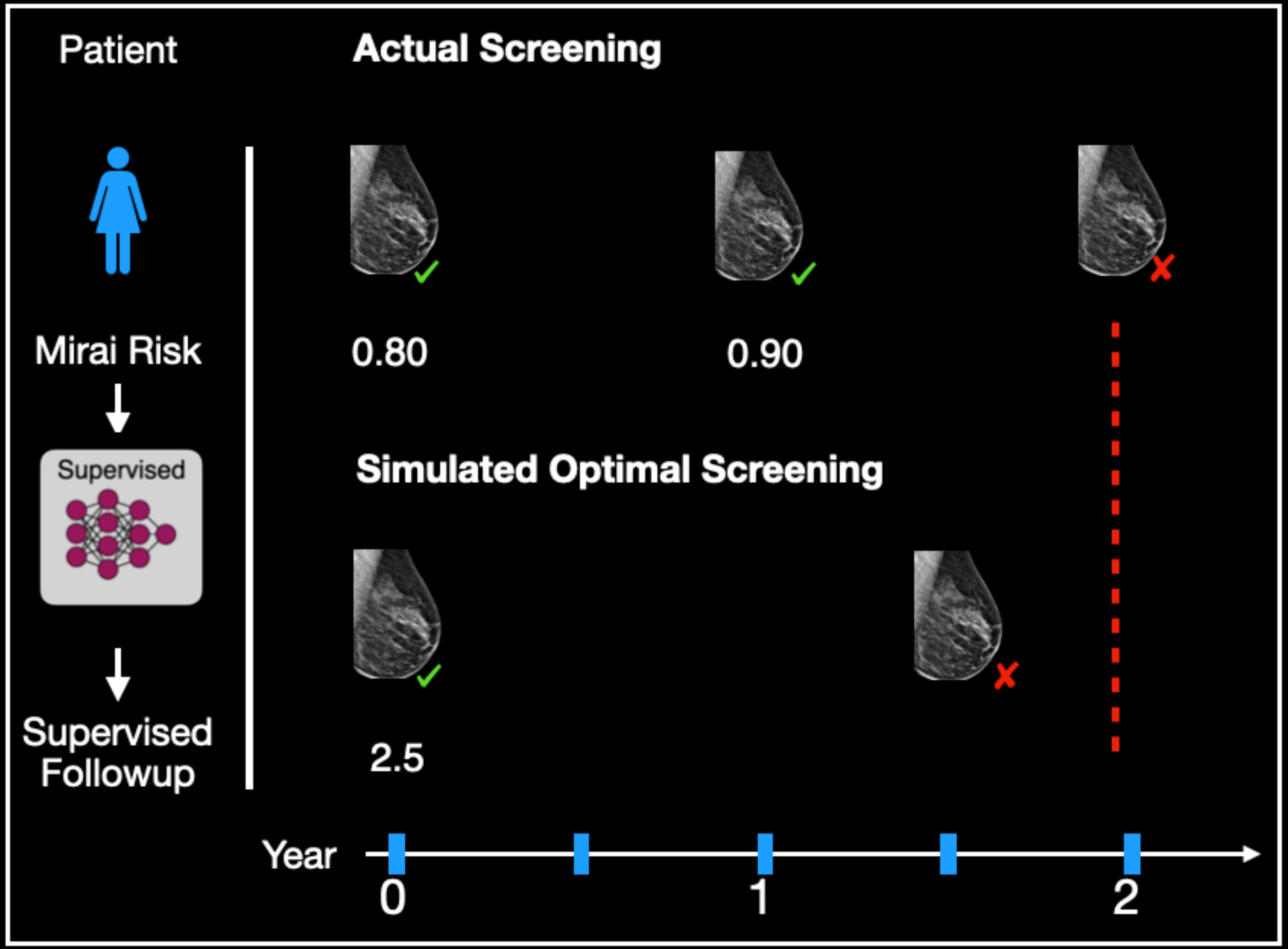
Estimating missing risk assessments

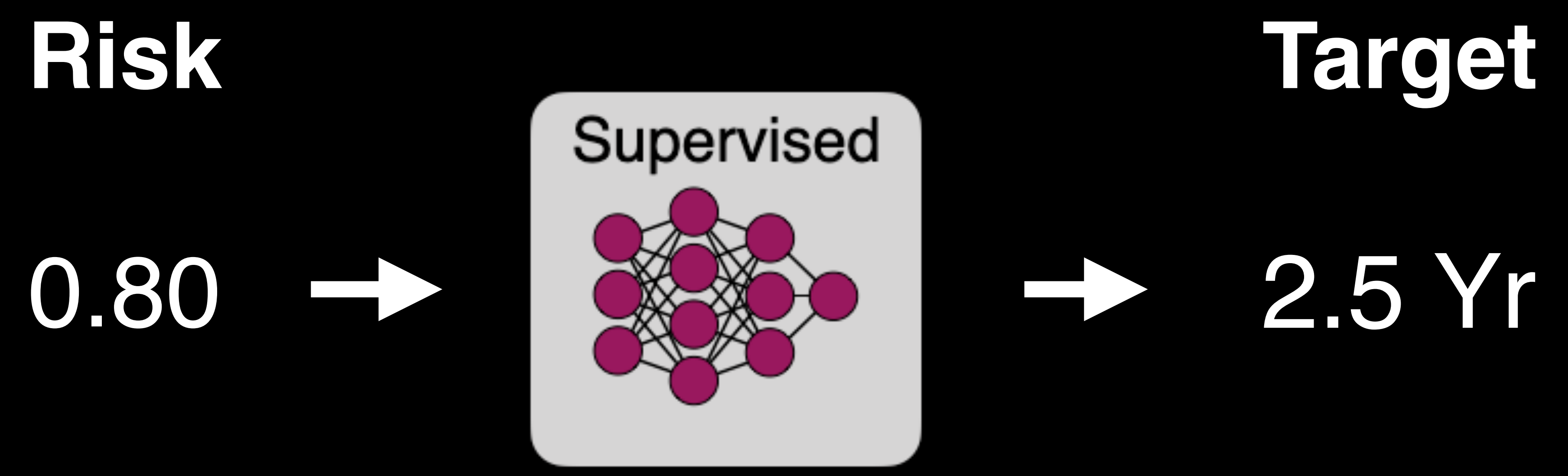


Learn $P(r_t \mid r_{t-1}, r_{t-2}, \ldots, r_0)$



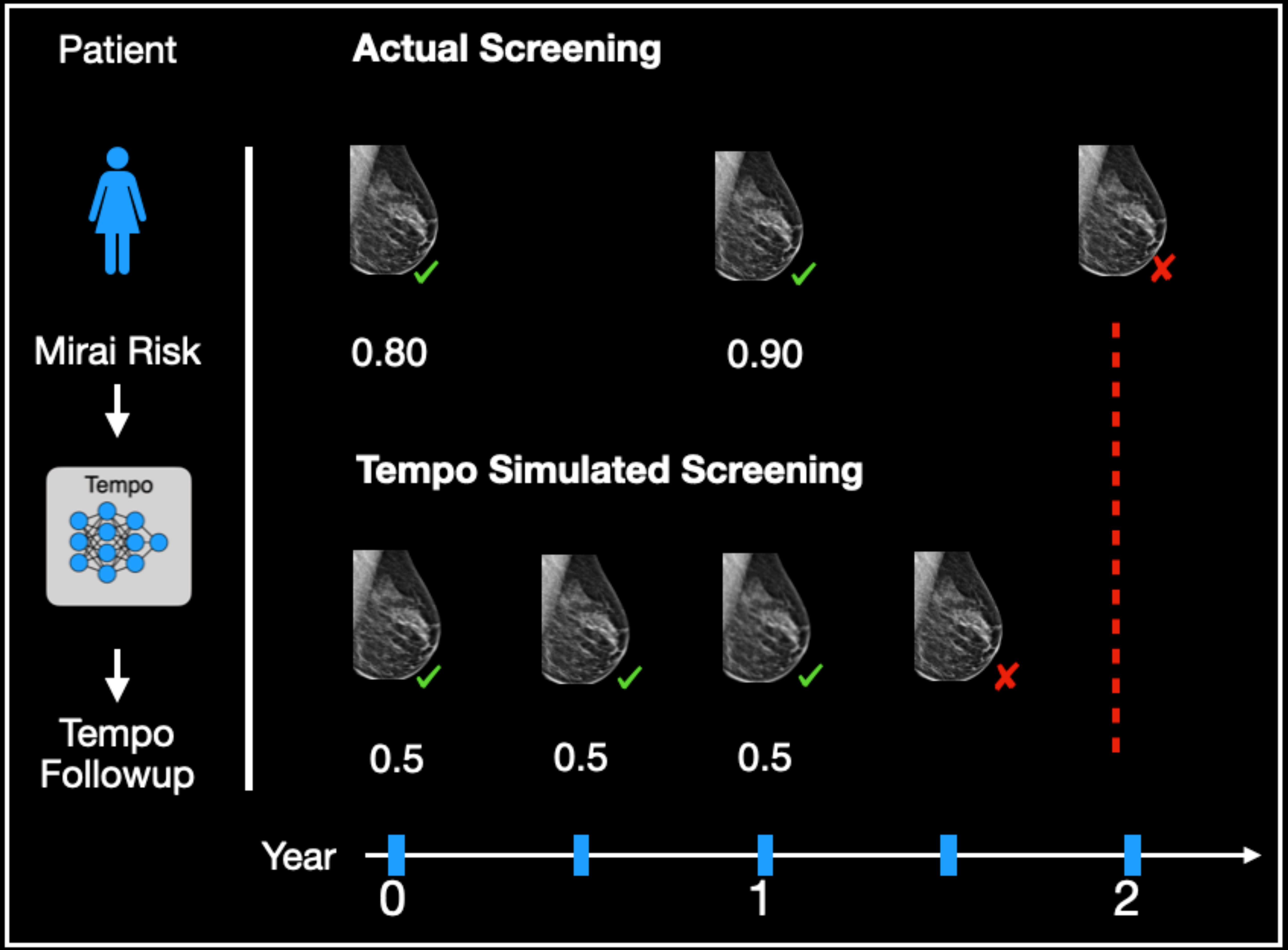
Baseline: Policy Design as Imitation

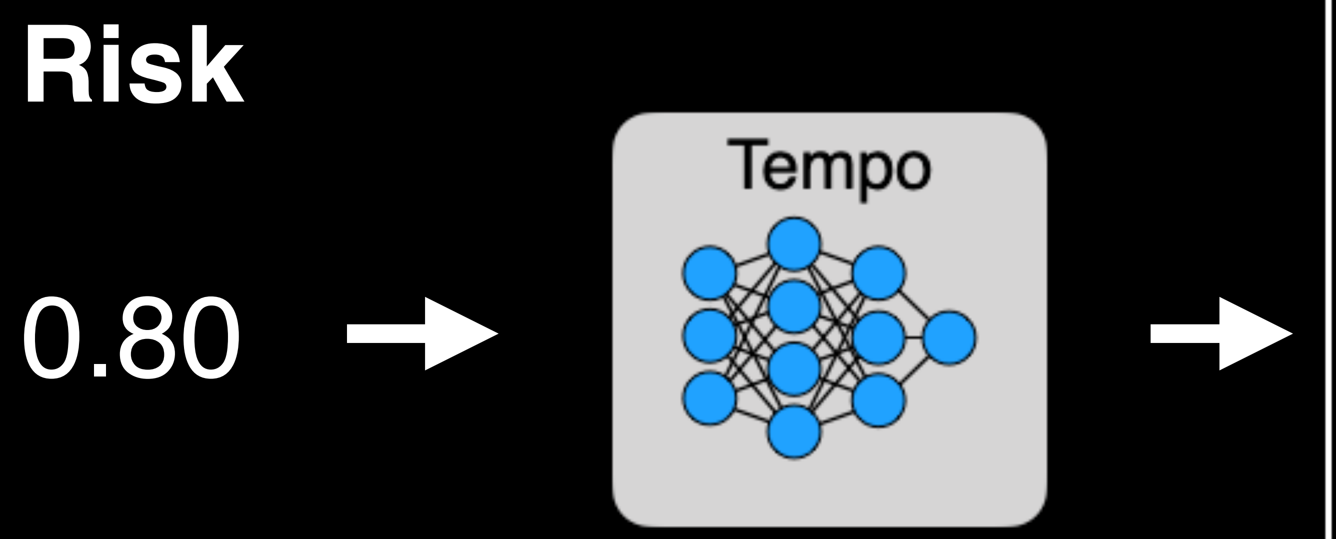




Reward = - || Prediction - Target ||

Tempo: Policy Design as Reinforcement Learning





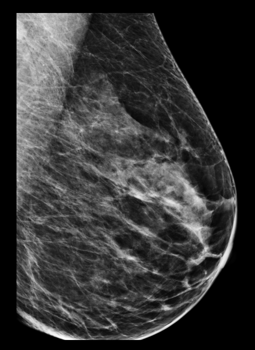
Prediction Target

(action, cost, benefit)		
0.5 Yrs,	4,	6 mo
1.0 Yrs,	2,	0
⋮		
3.0 Yrs,	1,	-12 mo

Reward = λ_1 Benefit - λ_2Cost

Learning for a fixed preference

Patient



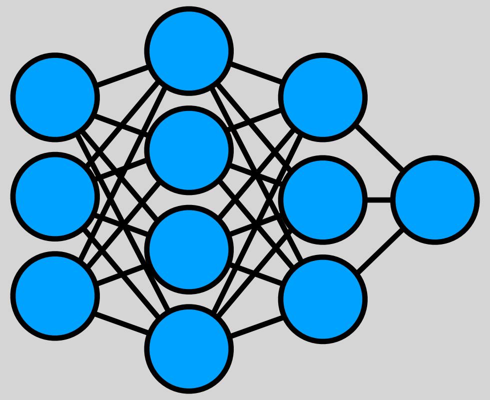
→

Risk

0.80

→

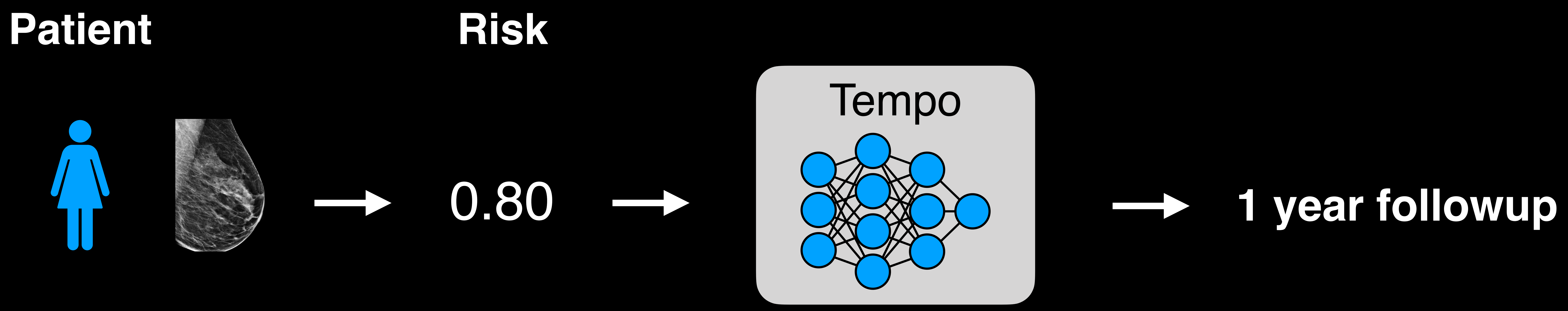
Tempo



→ **1 year followup**

$$\text{Reward} = \lambda_1 \text{ Early Detection Benefit} - \lambda_2 \text{Screening Cost}$$

Learning for a fixed preference: Q Learning

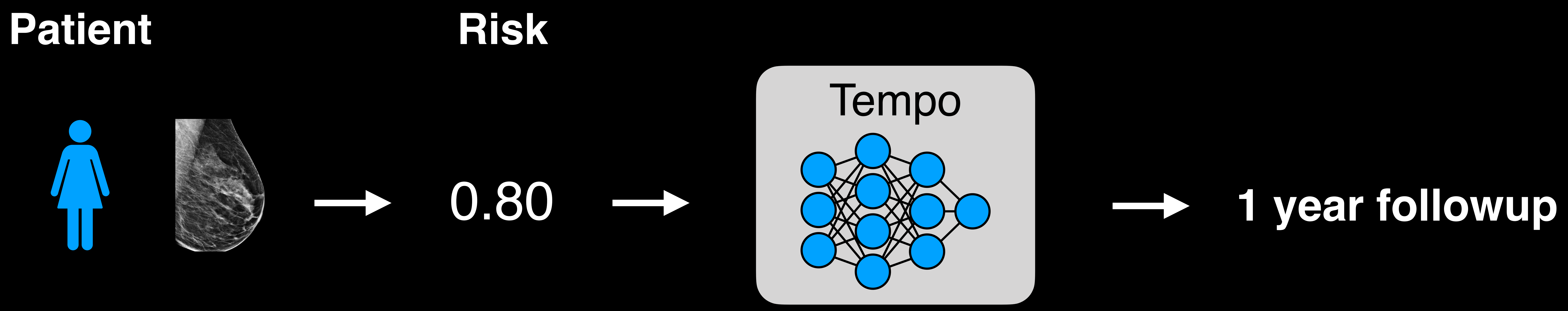


Reward = λ_1 Early Detection Benefit - λ_2Screening Cost

$$Q(s,a) = R(s,a) + \gamma \max_a Q(s',a)$$

$$\mathcal{L}(s,a) = ||R(s,a) + \max_a Q(s',a) - Q(s,a)||^2$$

Tricks for stable training

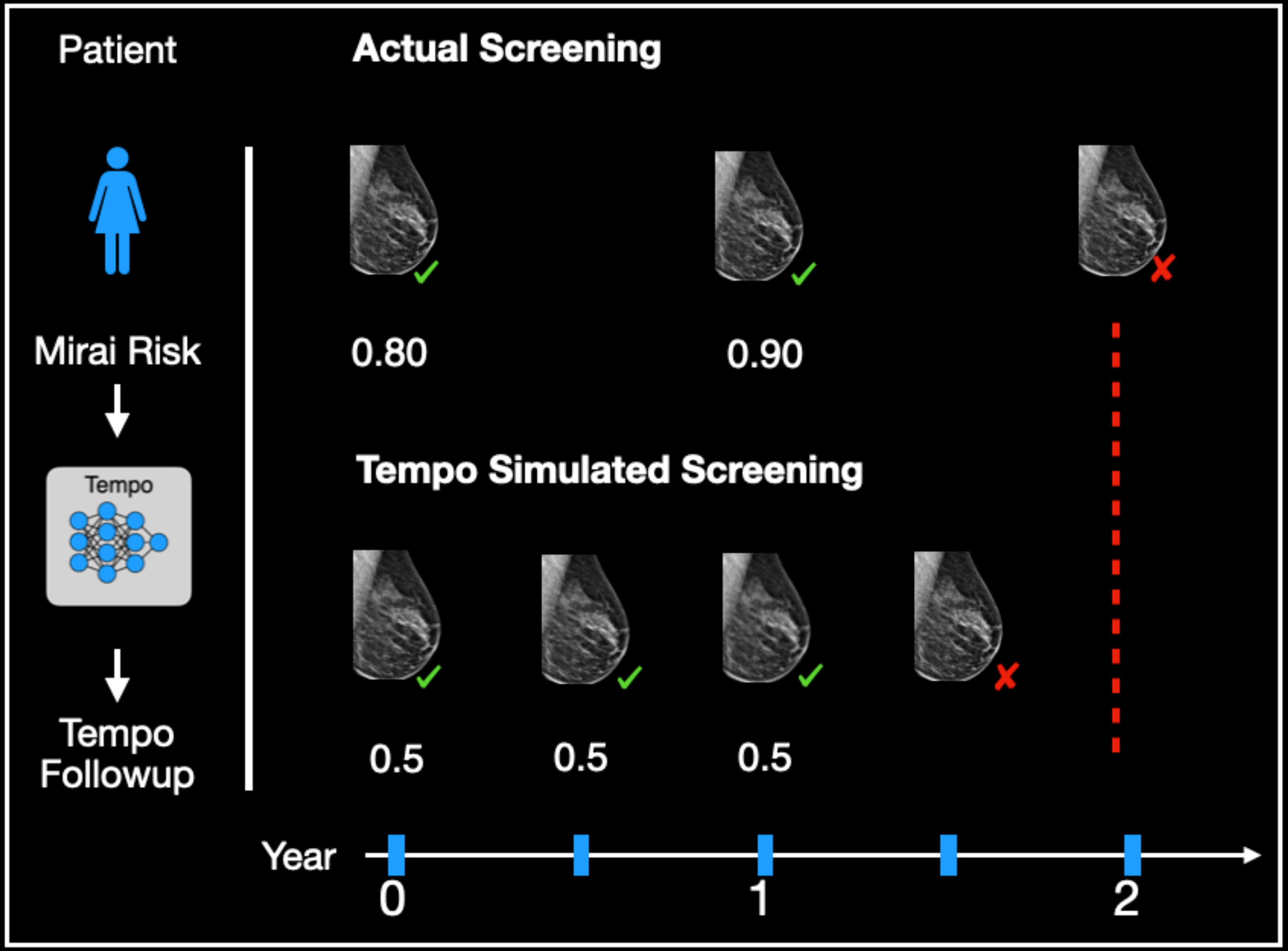


Randomly sample transitions from *Experience Replay Buffers*

Use slowly updated target network (i.e. copy Q every 100~ steps)

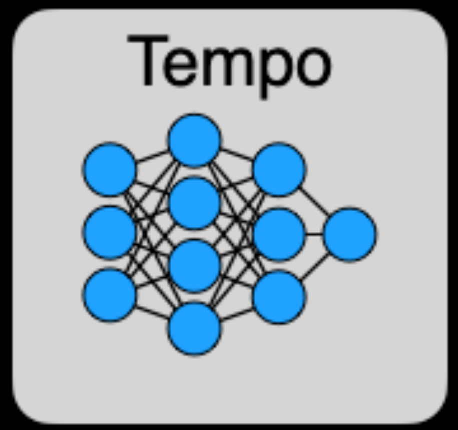
$$\mathscr{L}(s,a) = ||R(s,a) + \text{max}_a Q_{\boxed{target}}(s',a) - Q(s,a)||^2$$

Scaling to unknown preferences



Risk

0.80 →



→

Prediction Target

(action, cost, benefit)		
0.5 Yrs,	4,	6 mo
1.0 Yrs,	2,	0
⋮		
3.0 Yrs,	1,	-12 mo

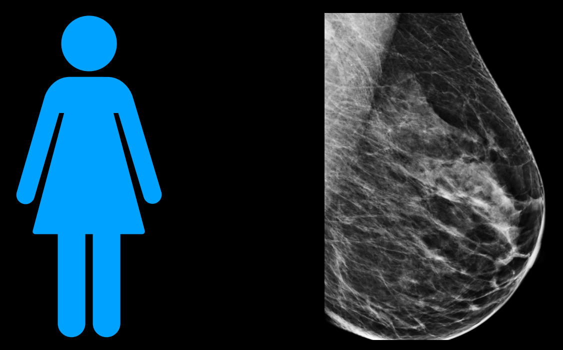
Reward = λ_1 Benefit - λ_2Cost

- λ unknown at training time

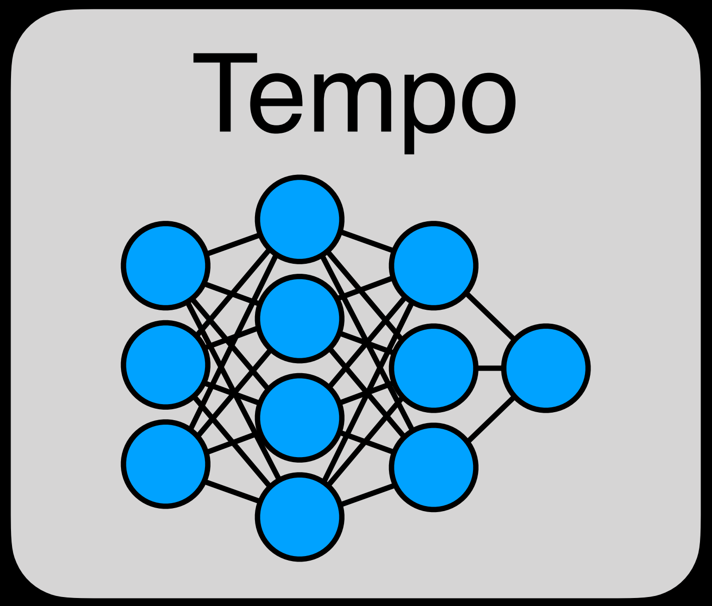
+ We have access to [Benefit, Cost]

Supporting diverse clinical requirements

Patient **Risk**



→ 0.80 →



→ **1 year followup**

↑

preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

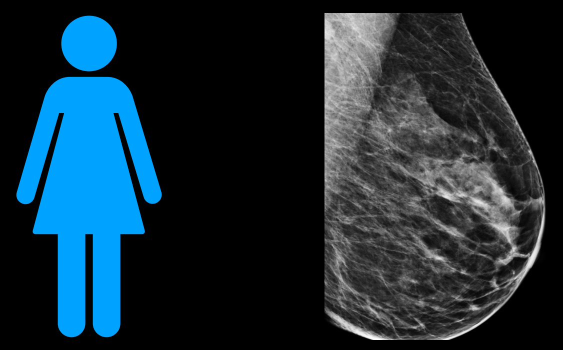
$\vec{r}(s, a) =$ [Early Detection, Screening Cost]

Trained across possible (λ_1, λ_2) to maximize:

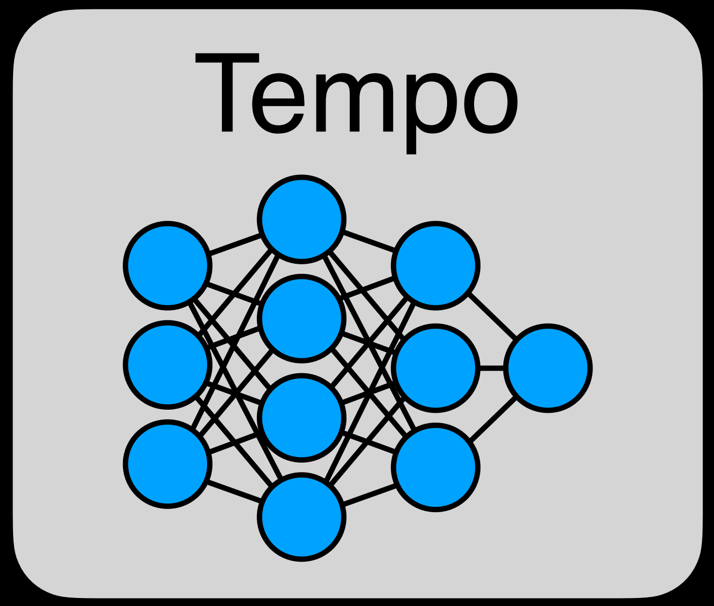
$\lambda \cdot \vec{r} = \lambda_1$ Early Detection Benefit - λ_2Screening Cost

Towards multi-objective RL: Scalarized updates

Patient → **Risk**



→ 0.80 →



→ **1 year followup**

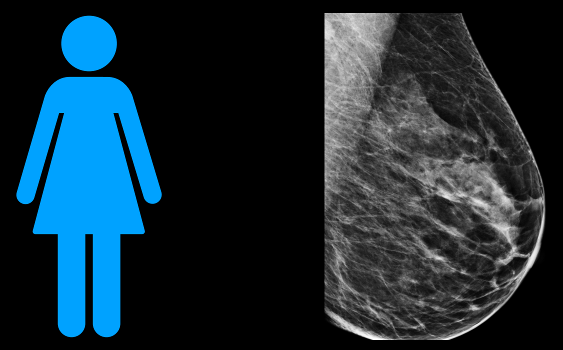
preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

$$\vec{r}(s,a) = [\text{Early Detection, Screening Cost}]$$

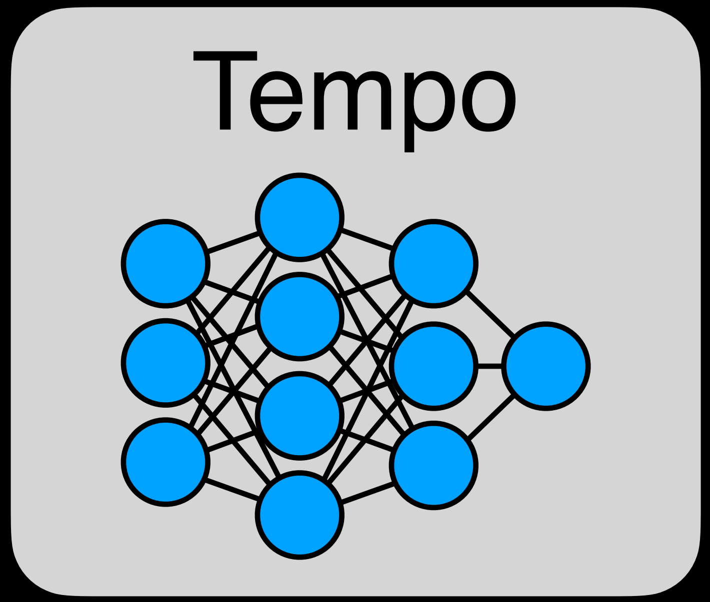
$$Q(s,a,\lambda) = \lambda\vec{r}(s,a) + \gamma\lambda\,\text{max}_a Q(s',a,\lambda)$$

Towards multi-objective RL: Scalarized updates

Patient **Risk**



→ 0.80 →



→ **1 year followup**

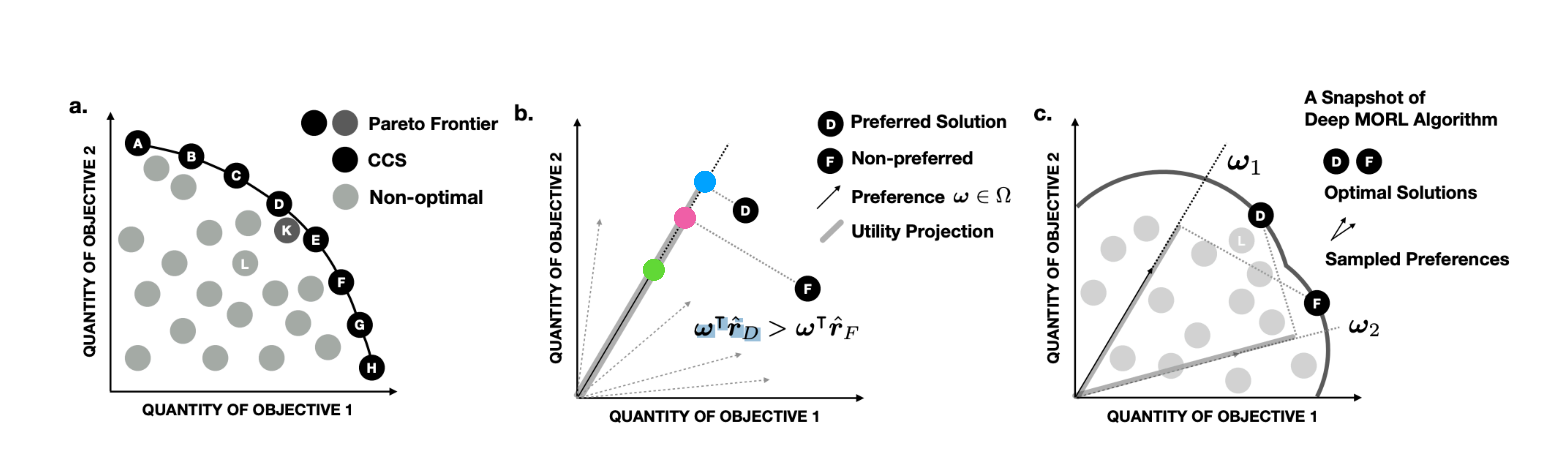
↑

preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

$\vec{r}(s,a) = [\text{Early Detection, Screening Cost}]$

$Q(s,a,\lambda) = \lambda\vec{r}(s,a) + \gamma\lambda\text{max}_a Q(s',a,\lambda)$ ← **Doesn't use relationship between λ**

Envelope Q-Learning



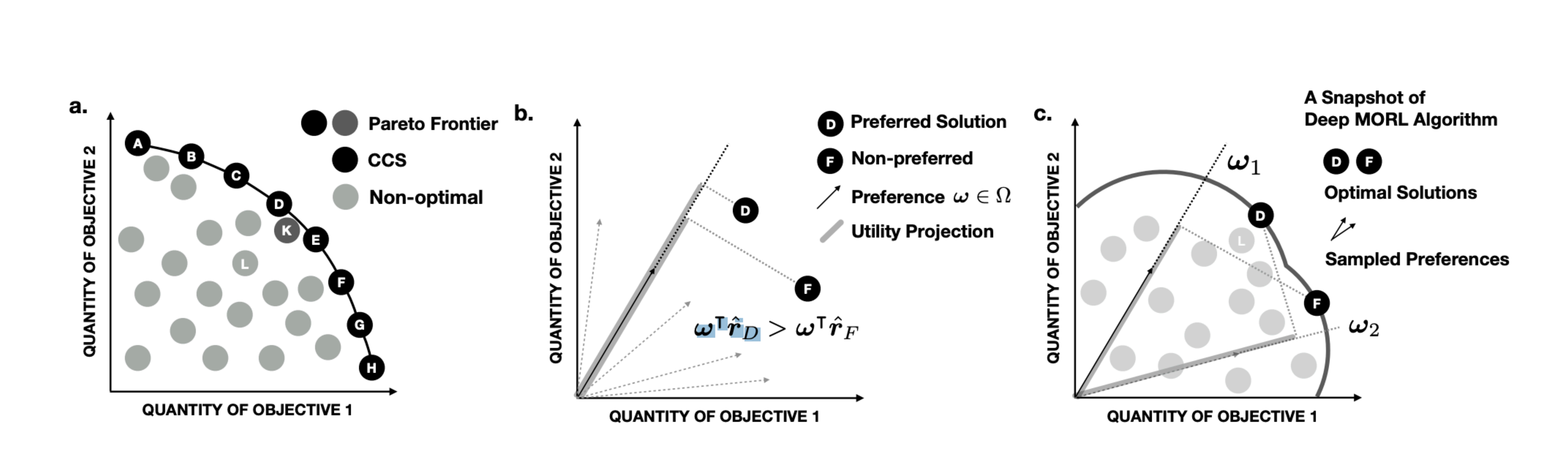
Search *envelope* of policy:

Identify λ' more effective for true target λ

A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation

Runzhe Yang, Xingyuan Sun, Karthik Narasimhan

Envelope Q-Learning



$$y = \vec{r}(s, a) + \gamma \arg_Q \max_{a,\lambda'} \lambda^t \ Q(s', a, \lambda')$$

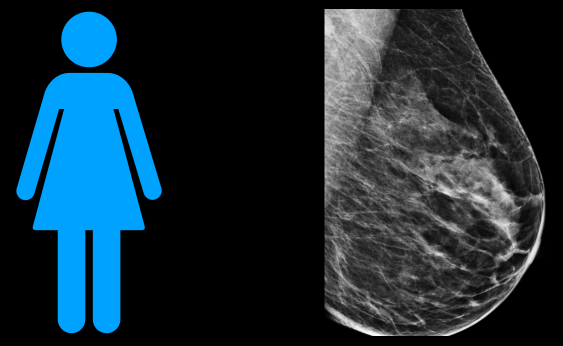
$$\mathscr{L}(s, a, \lambda) = ||y - Q(s, a, \lambda)||^2$$

A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation

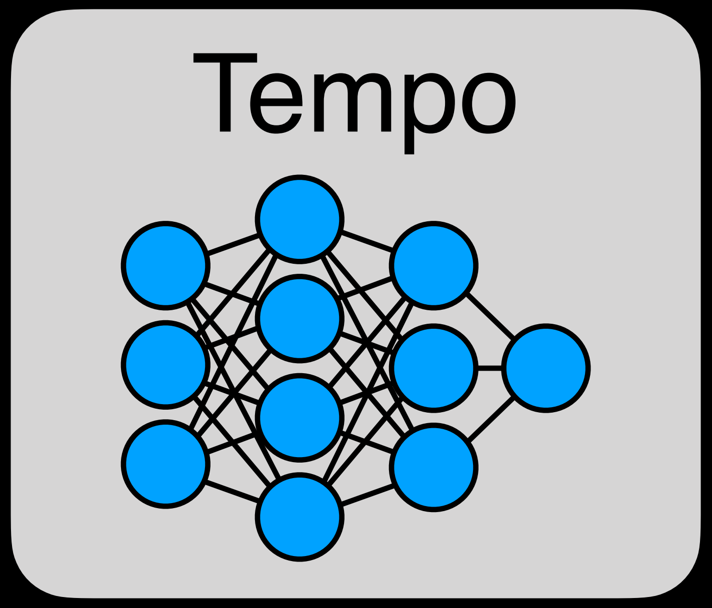
Runzhe Yang, Xingyuan Sun, Karthik Narasimhan

Supporting diverse clinical requirements

Patient **Risk**



→ 0.80 →



→ **1 year followup**

↑

preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

Trained across possible (λ_1, λ_2) to maximize:

λ_1 Early Detection Benefit - λ_2Screening Cost

Experimental Setup

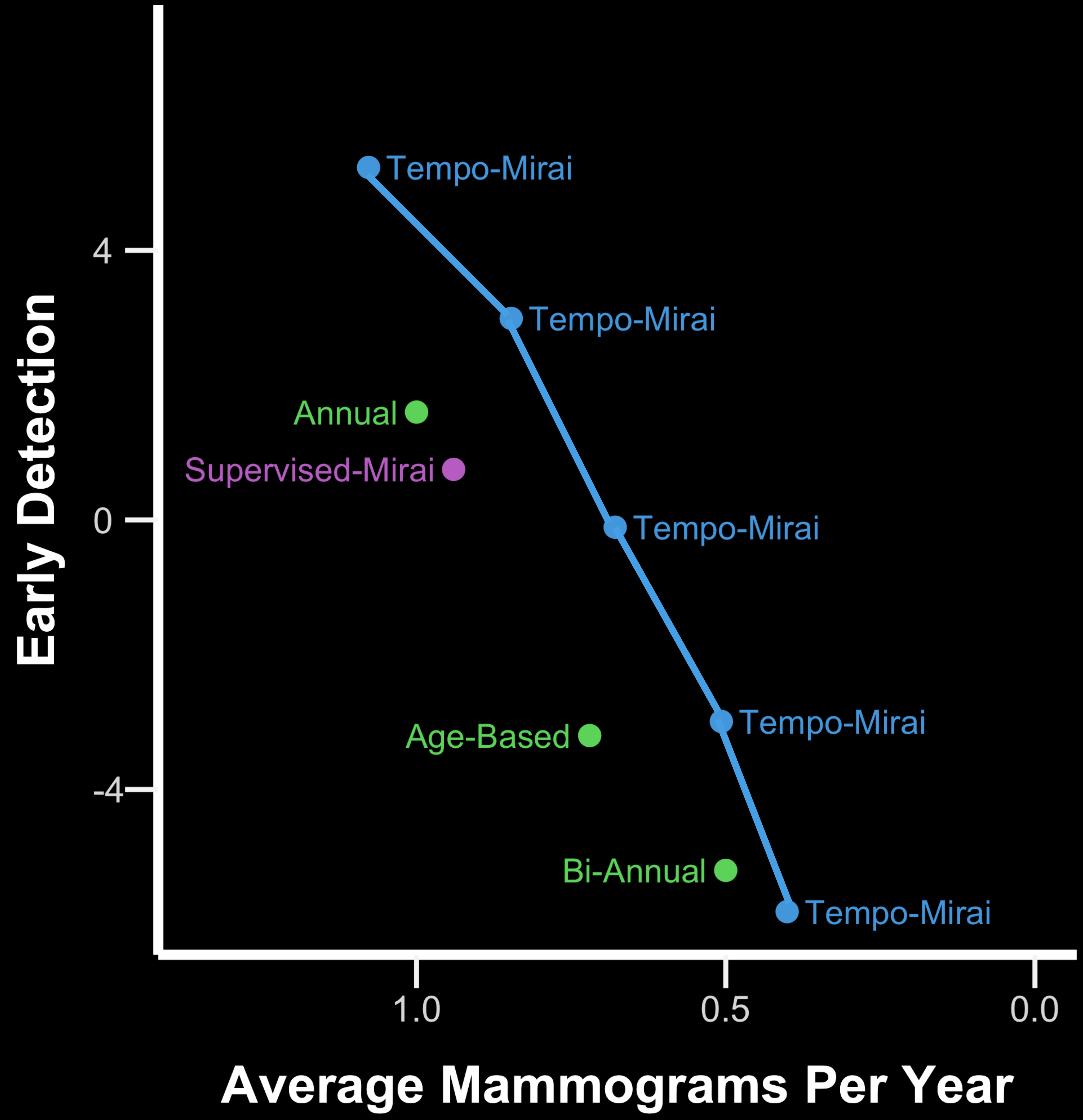
Train all models on MGH training set

Test on MGH, Emory, Karolinska, and CGMH

Evaluate Screening Efficiency: $\dfrac{\textbf{Early Detection}}{\textbf{Avg Mammo per Year}}$

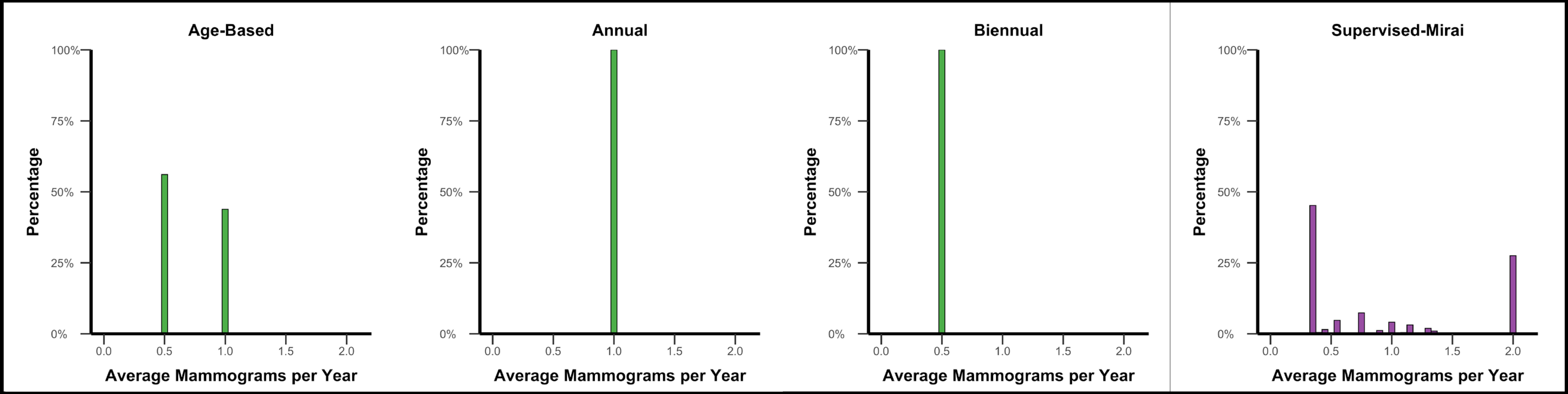
Results: MGH Screening Efficiency

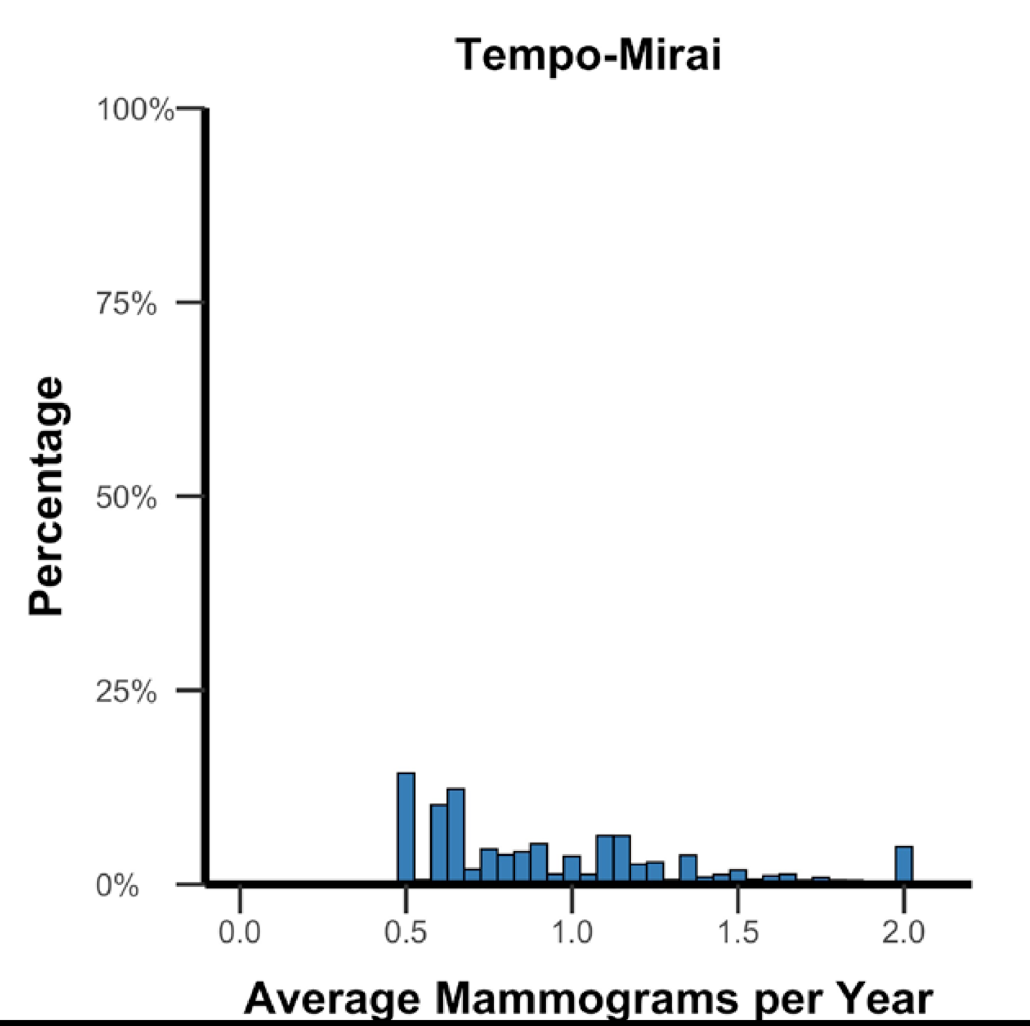
Supporting diverse clinical needs



MGH Test Set

How do the policy behave?



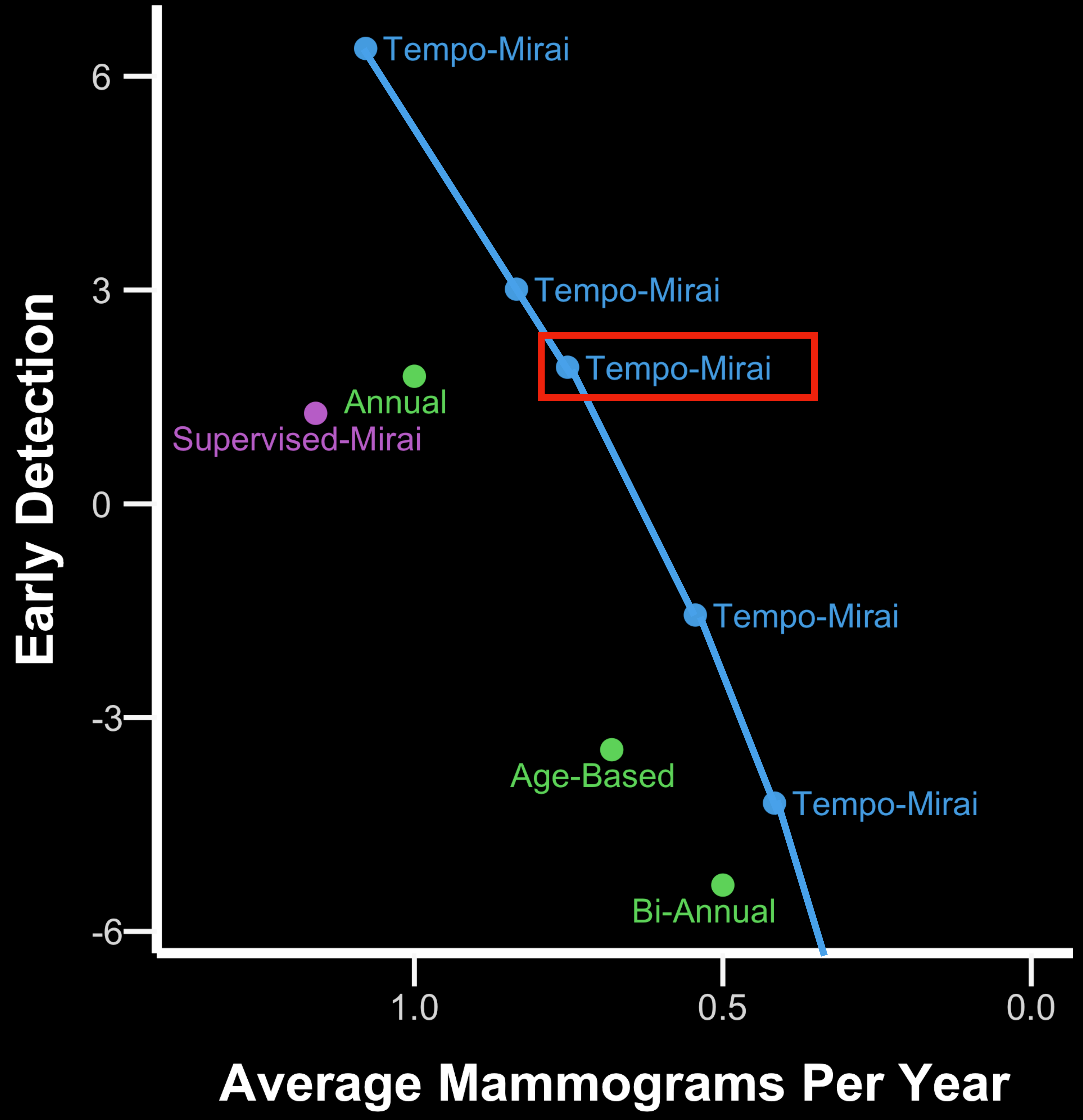


How to Deploy?

Validate on **retrospective data**

Choose desired operating point

Run prospective trials



Emory Test Set

Summary

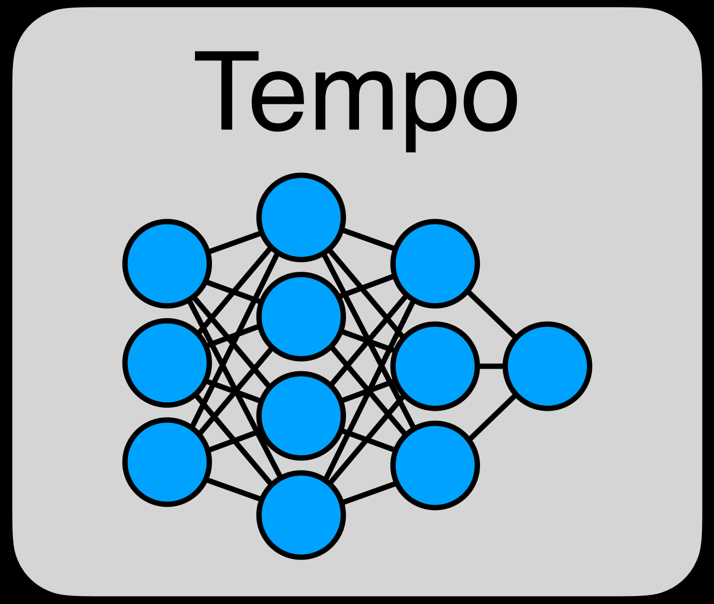
Patient

→

Risk

0.80

→



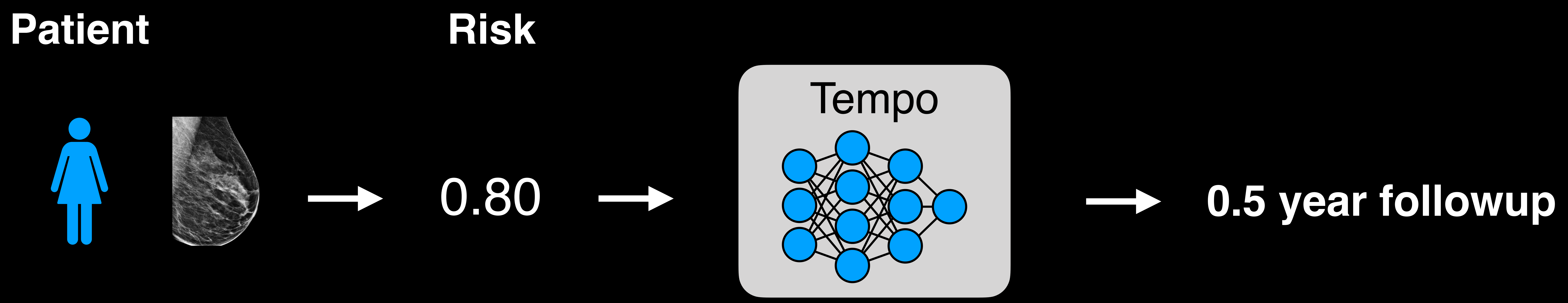
→ **0.5 year followup**

Learn personalized screening policies by modeling individuals

Applicable with arbitrary reward design / choice of risk model

Better **early detection and less overscreening**

What’s missing?



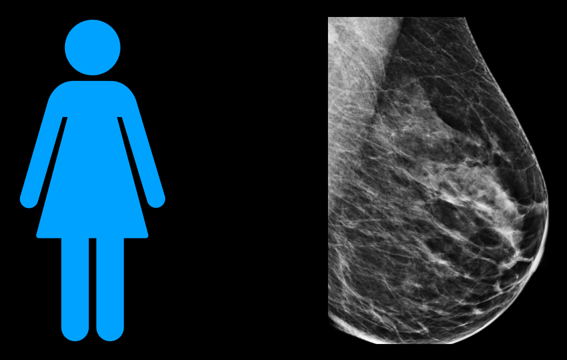
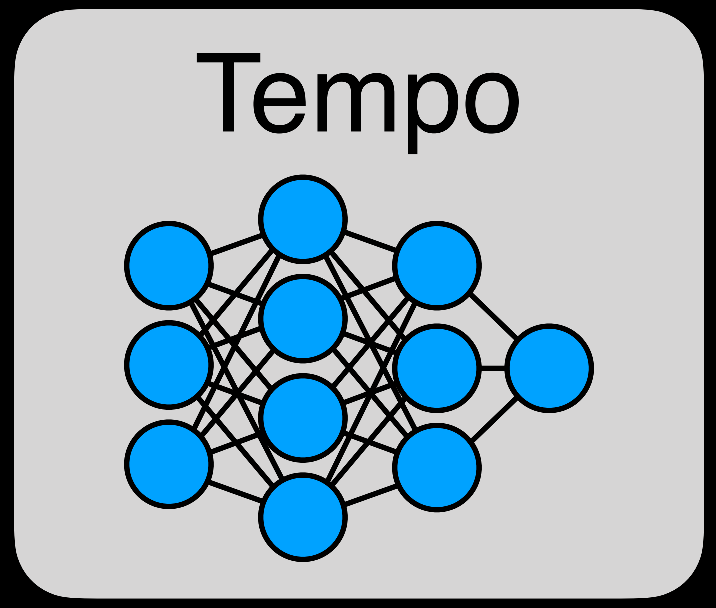
Robustness to risk progression

Preference elicitation

Understanding causal effects from real trajectories

What's missing?

Patient **Risk**

 → 0.80 → 

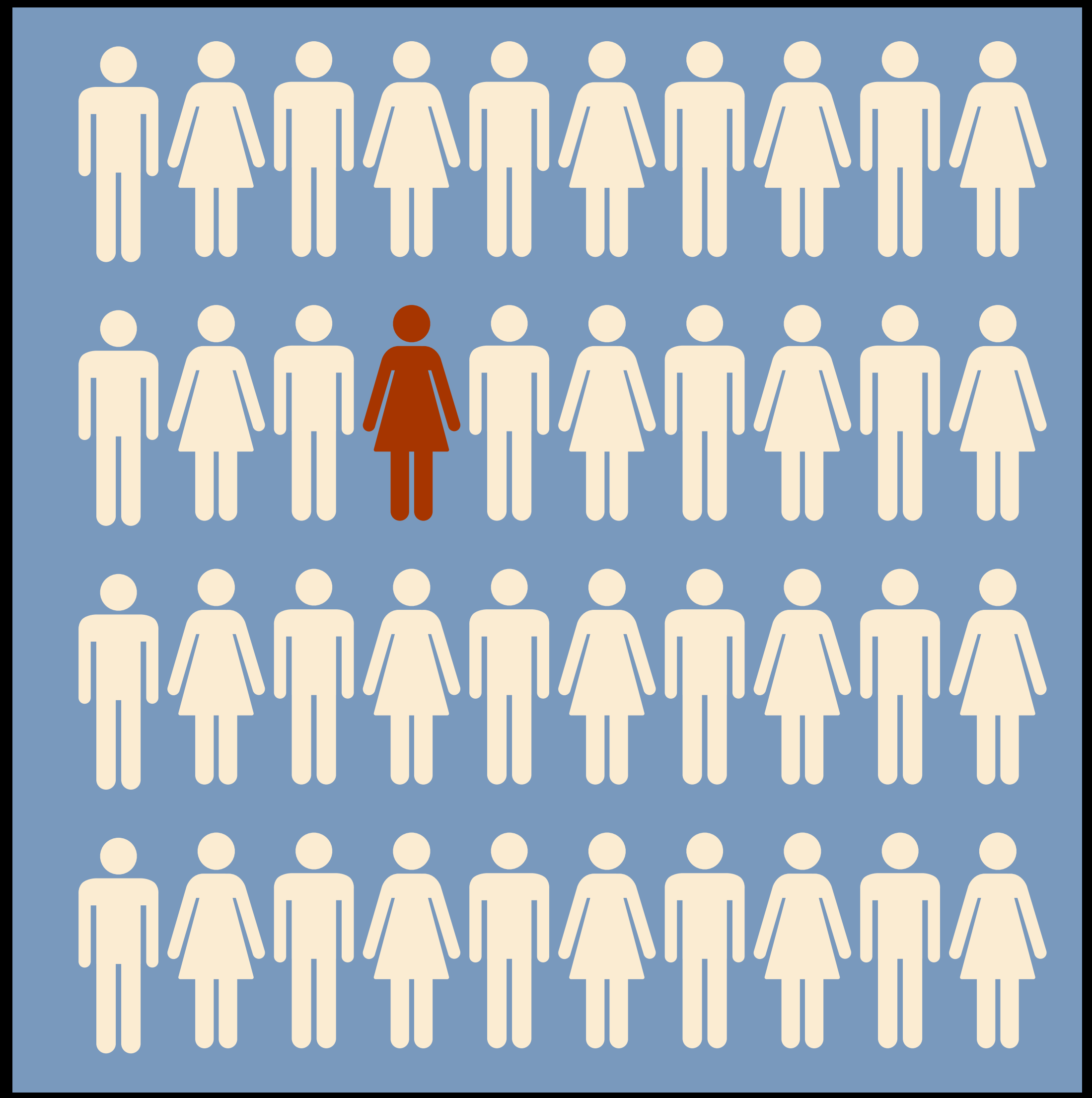
→ **0.5 year followup**

Modeling patient data directly for policy design

Hierarchical Action Spaces

Deployment Safety

What if we succeed?



Rethink screening criteria / guidelines across diseases

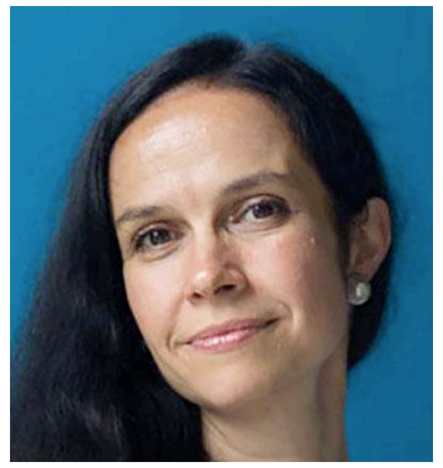
Early detection for “unscreeneable" diseases

New doors for prevention and therapeutic development

Thank you to my collaborators

Regina Barzilay
MIT

Tommi Jaakkola
MIT



Muriel Medard
MIT

Tal Schuster
Google

Peter Mikhael
MIT

Victor Quach
MIT

Kevin Hughes
MGH

Fredrik Strand
Karolinska

Gigin Lin
CGMH

Judy Gichoya
Emory

Hari Traveri
Emory

Michal Guindy
Maccabi

Join us!

computationalhealth.berkeley.edu

Prospective PhD students: Join my lab!

Prospective Faculty: We are now recruiting!

Computational
PRECISION HEALTH

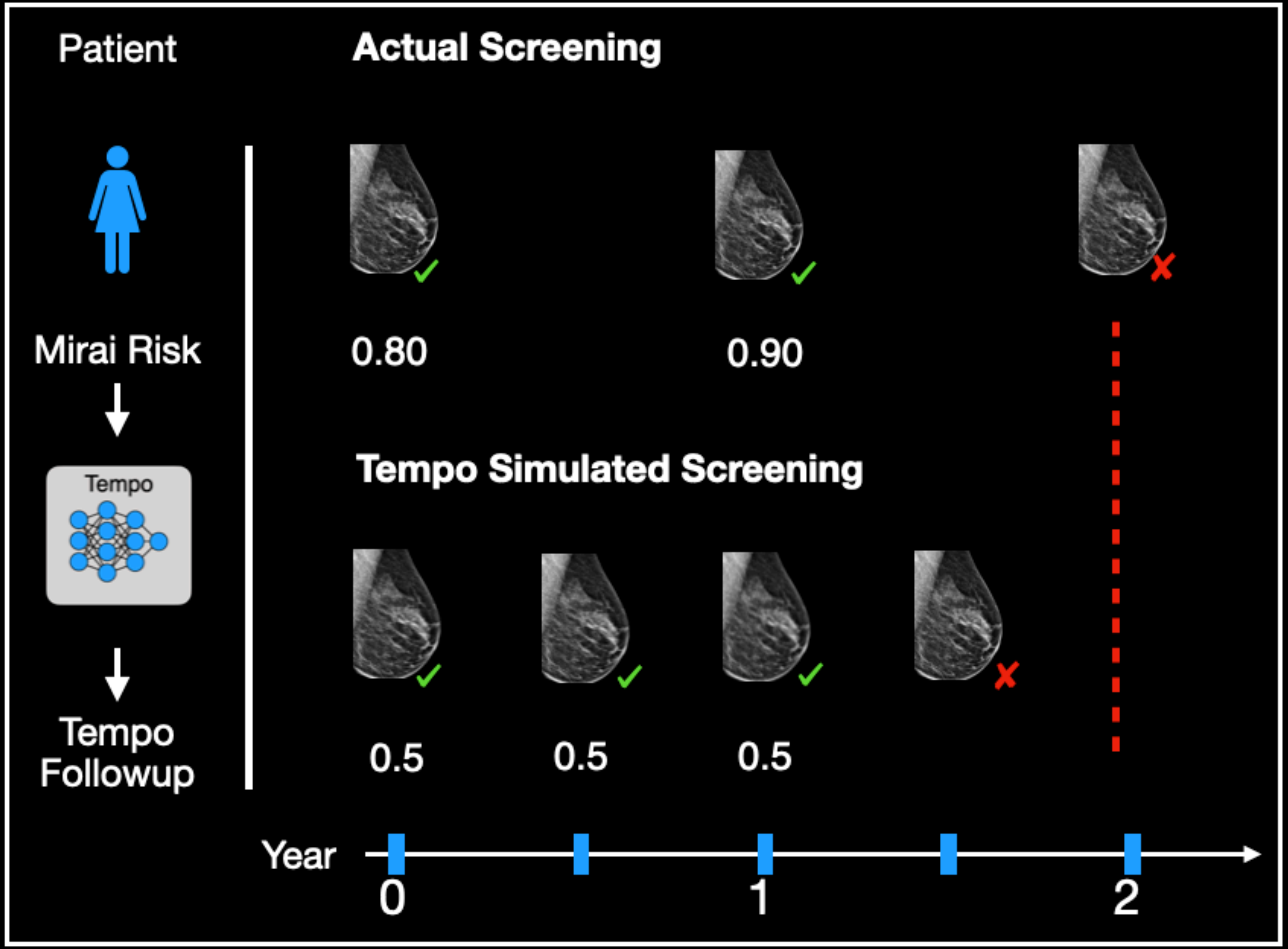
Berkeley | UCSF

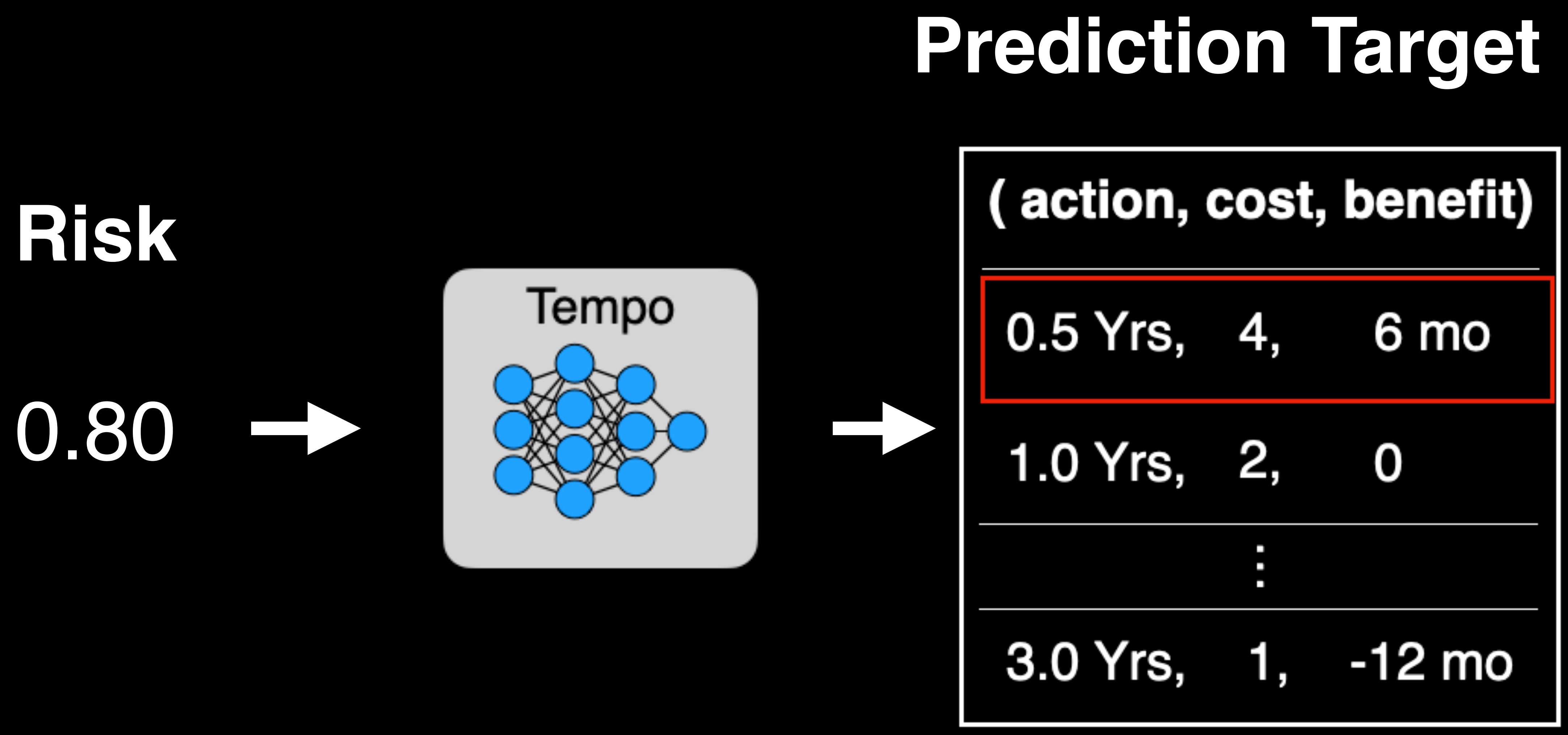
Questions?

Appendix Slides

Early Detection Appendix Slides

Tempo: Policy Design as Reinforcement Learning

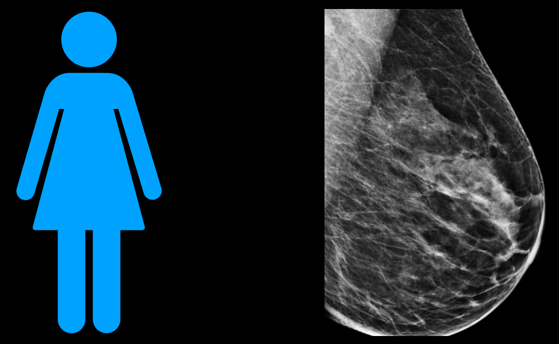




Reward = λ_1 Benefit - λ_2Cost

Learning for a fixed preference

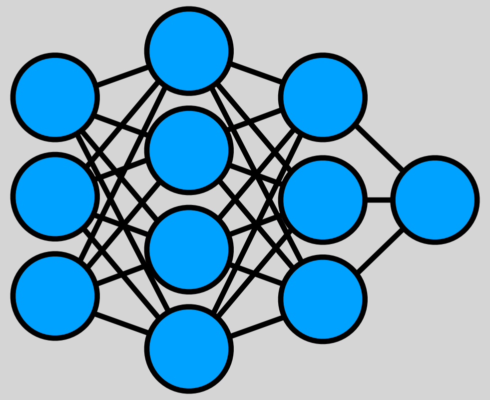
Patient



Risk

0.80 →

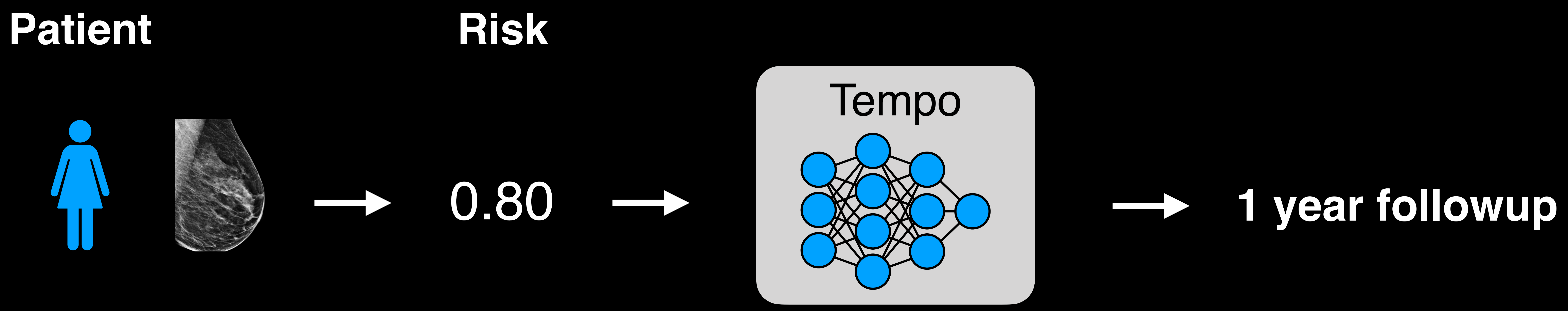
Tempo



→ **1 year followup**

Reward = λ_1 Early Detection Benefit - λ_2Screening Cost

Learning for a fixed preference: Q Learning

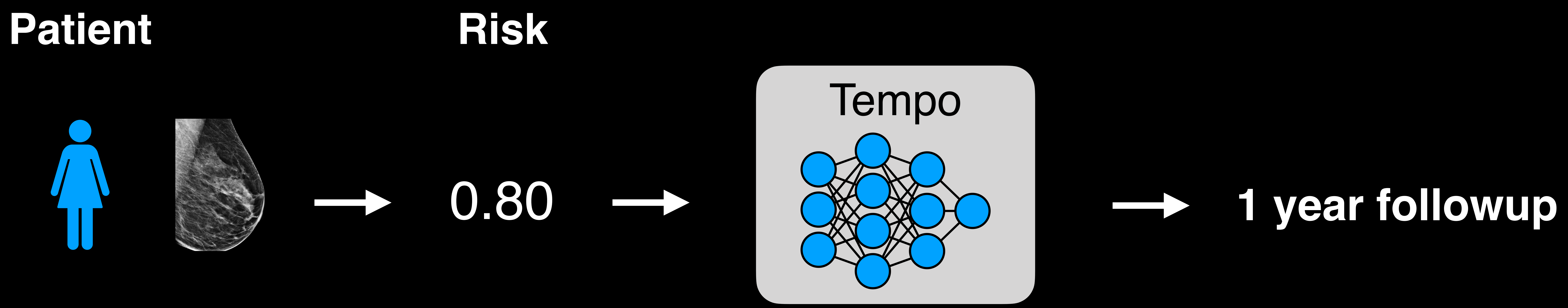


Reward = λ_1 Early Detection Benefit - λ_2Screening Cost

$$Q(s,a) = R(s,a) + \gamma \text{max}_a Q(s',a)$$

$$\mathscr{L}(s,a) = ||R(s,a) + \text{max}_a Q(s',a) - Q(s,a)||^2$$

Tricks for stable training

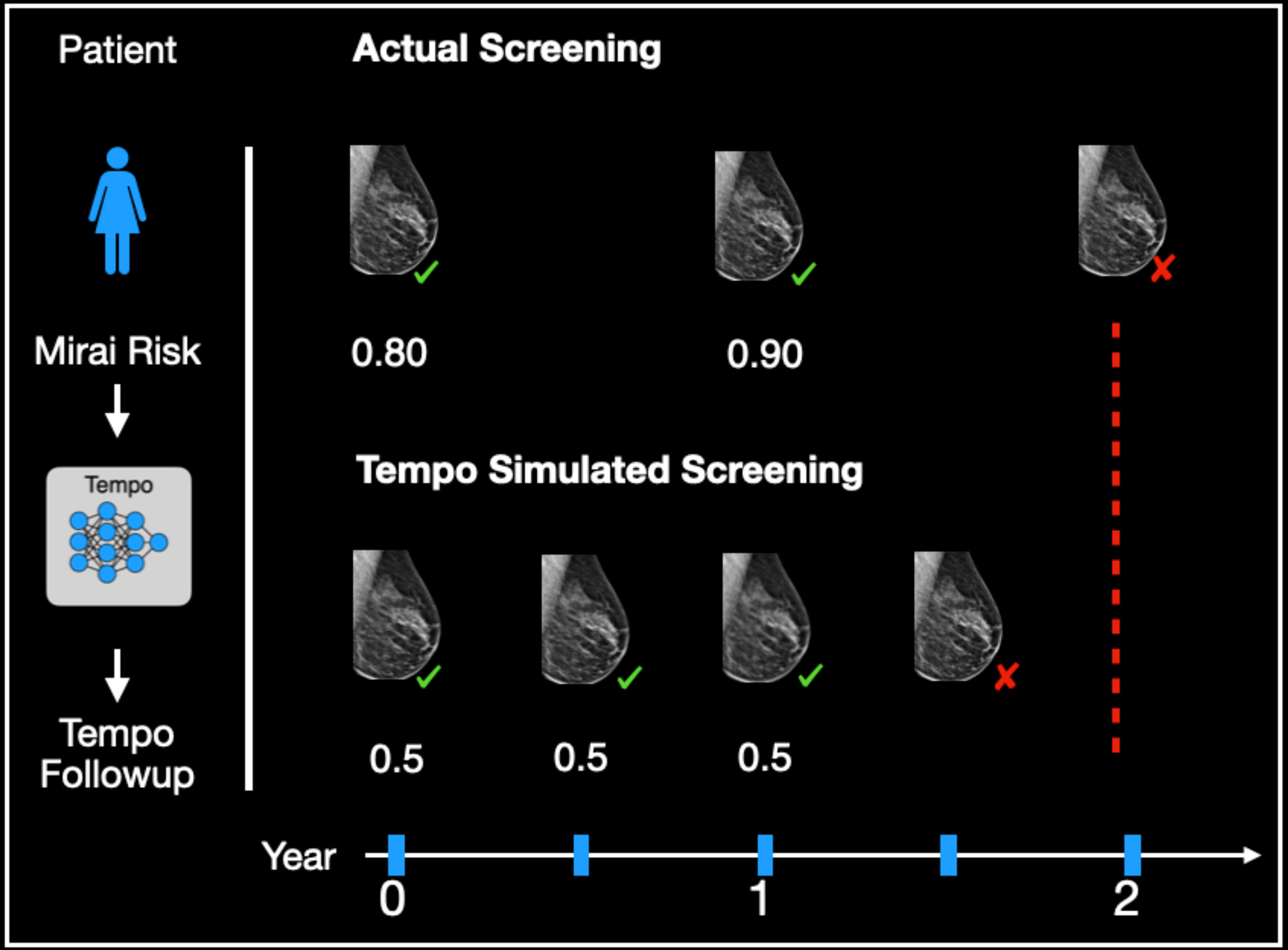


Randomly sample transitions from *Experience Replay Buffers*

Use slowly updated target network (i.e. copy Q every 100~ steps)

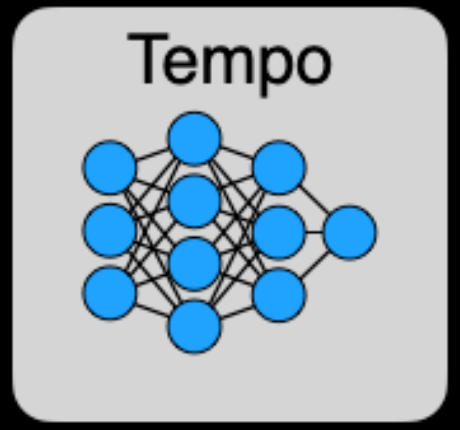
$$\mathcal{L}(s,a) = ||R(s,a) + \text{max}_a Q_{\boxed{target}}(s',a) - Q(s,a)||^2$$

Scaling to unknown preferences



Risk

0.80 →



→

Prediction Target

(action, cost, benefit)		
0.5 Yrs,	4,	6 mo
1.0 Yrs,	2,	0
⋮		
3.0 Yrs,	1,	-12 mo

Reward = λ_1 Benefit - λ_2Cost

- λ unknown at training time
+ We have access to [Benefit, Cost]

Supporting diverse clinical requirements

Patient

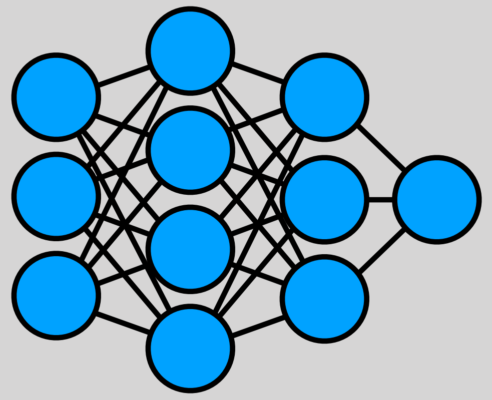
→

Risk

0.80

→

Tempo



→ **1 year followup**

preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

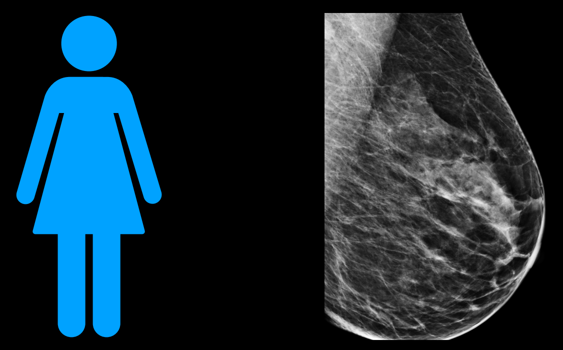
$\vec{r}(s, a) =$ [Early Detection, Screening Cost]

Trained across possible (λ_1, λ_2) to maximize:

$\lambda \cdot \vec{r} = \lambda_1$ Early Detection Benefit - λ_2Screening Cost

Towards multi-objective RL: Scalarized updates

Patient



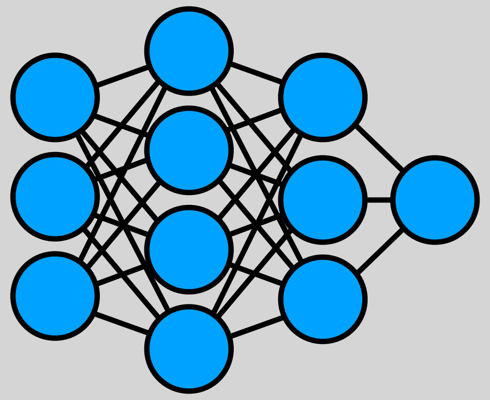
→

Risk

0.80

→

Tempo



→ **1 year followup**

↑

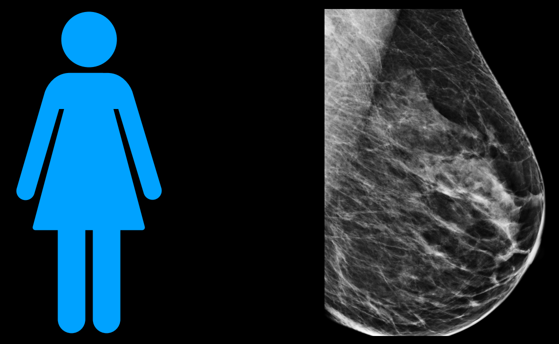
preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

$\vec{r}(s, a) = [\text{Early Detection, Screening Cost}]$

$Q(s, a, \lambda) = \lambda \vec{r}(s, a) + \gamma \lambda \text{max}_a Q(s', a, \lambda)$

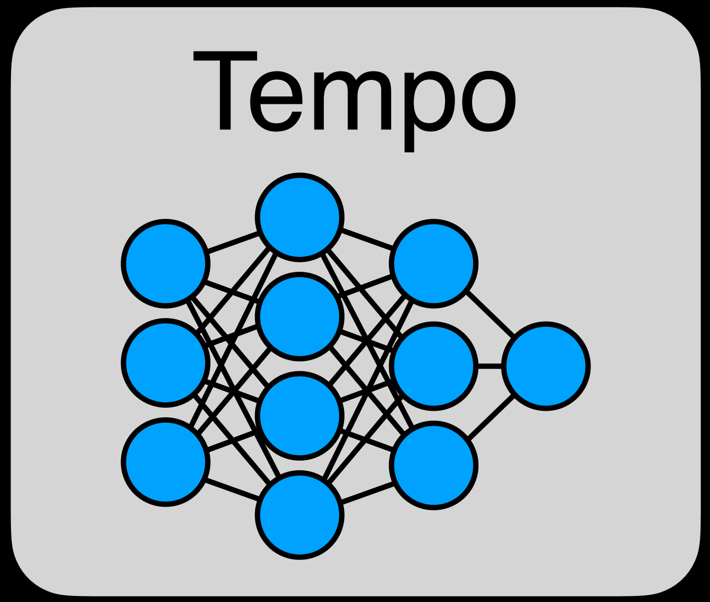
Towards multi-objective RL: Scalarized updates

Patient



Risk

0.80



1 year followup

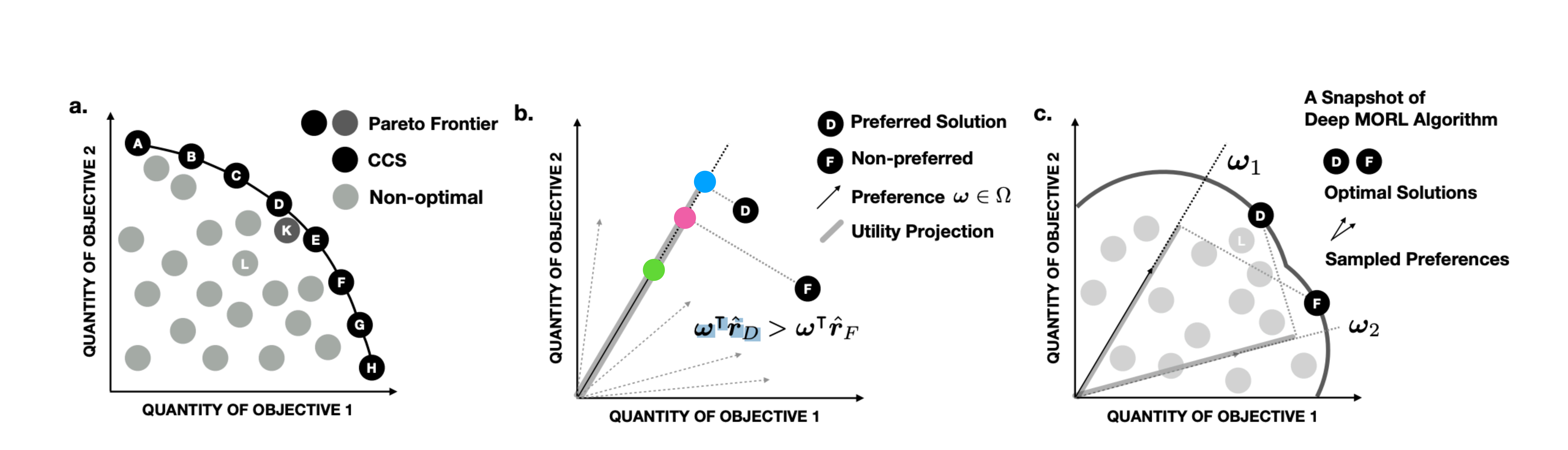
preference: $\lambda = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}$

$\vec{r}(s, a) = [\text{Early Detection, Screening Cost}]$

$Q(s, a, \lambda) = \lambda \vec{r}(s, a) + \gamma \lambda \text{max}_a Q(s', a, \lambda)$

Doesn't use relationship between λ

Envelope Q-Learning



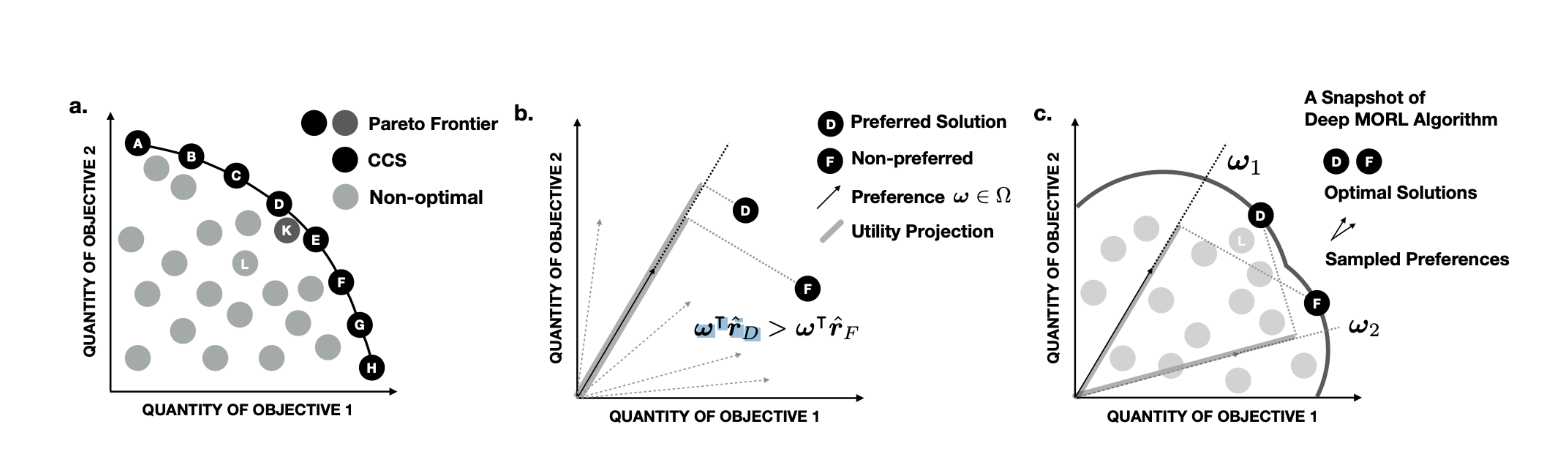
Search *envelope* of policy:

Identify λ' more effective for true target λ

A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation

Runzhe Yang, Xingyuan Sun, Karthik Narasimhan

Envelope Q-Learning



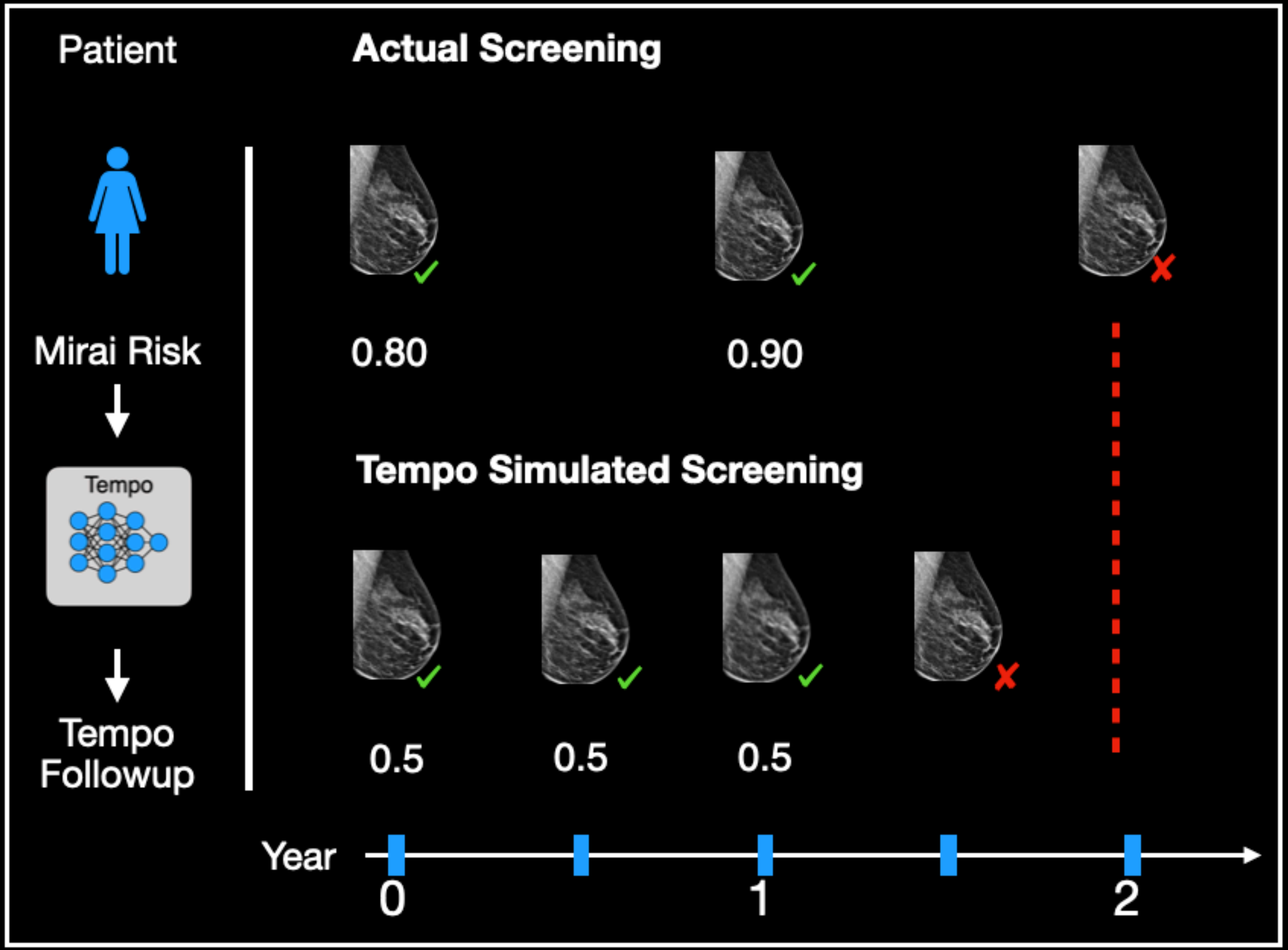
$$y = \vec{r}(s,a) + \gamma \mathrm{arg}_Q \mathrm{max}_{a,\lambda'} \lambda^t \; Q(s', a, \lambda')$$

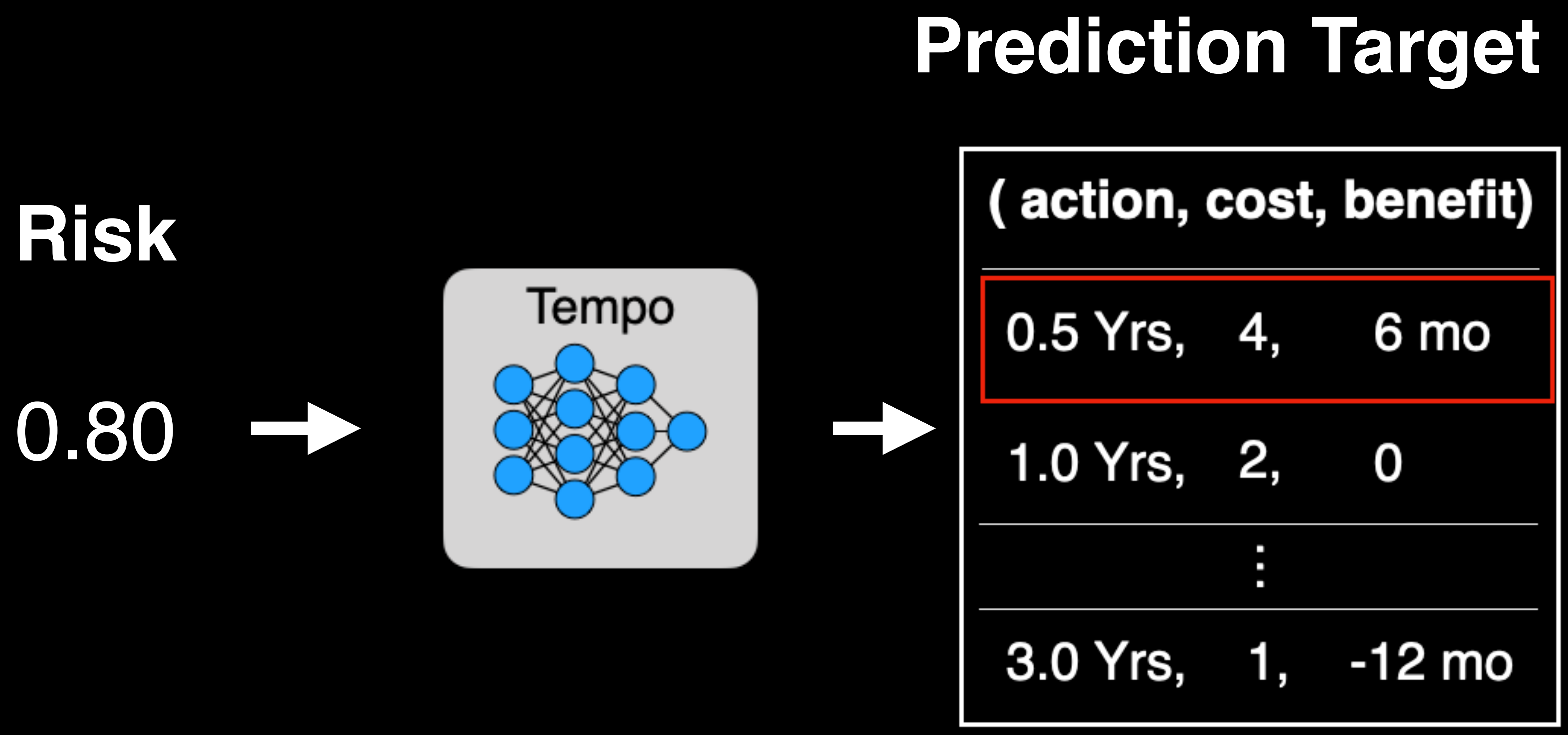
$$\mathscr{L}(s,a,\lambda) = ||y - Q(s,a,\lambda)||^2$$

A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation

Runzhe Yang, Xingyuan Sun, Karthik Narasimhan

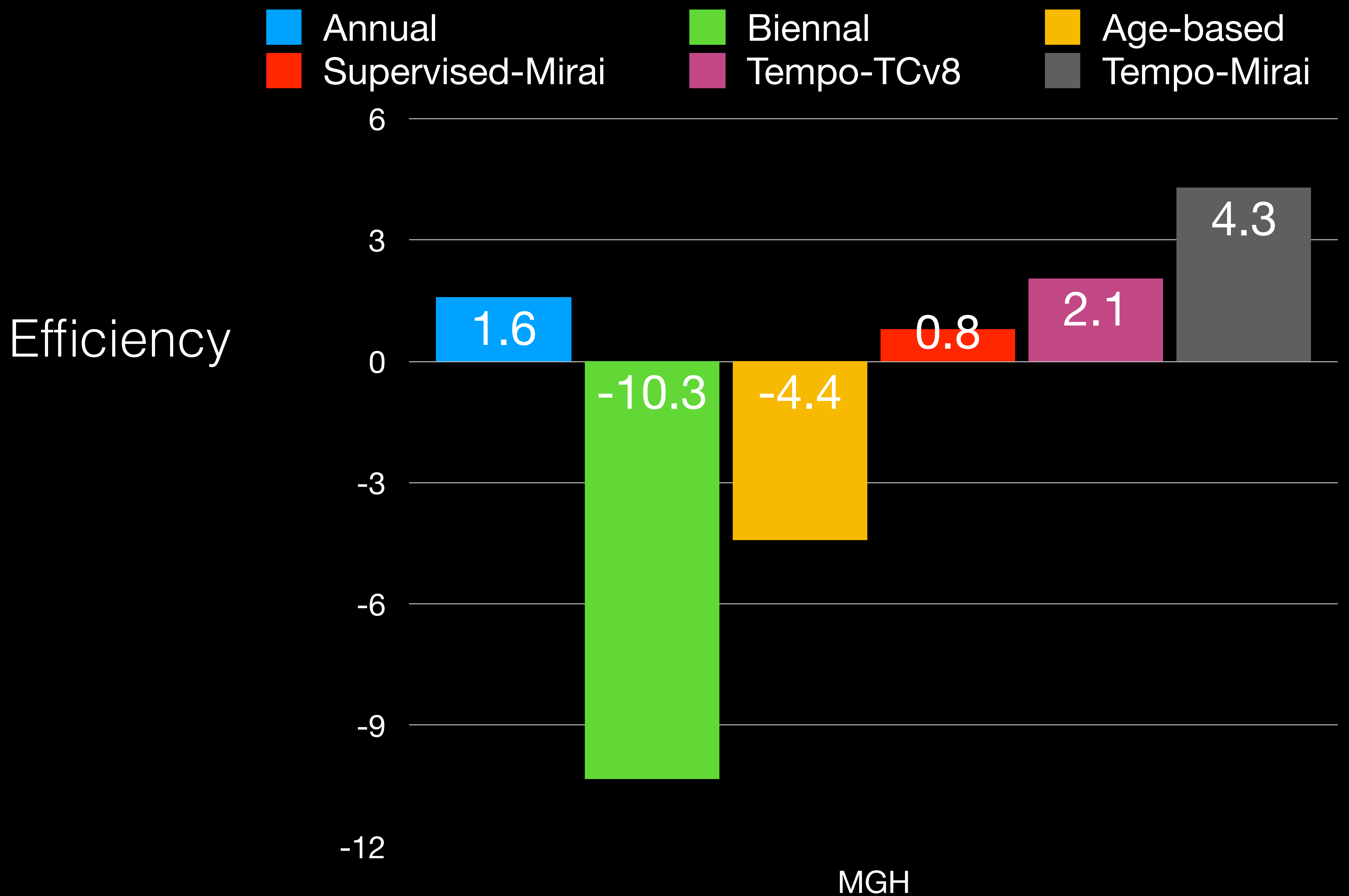
Tempo: Policy Design as Reinforcement Learning



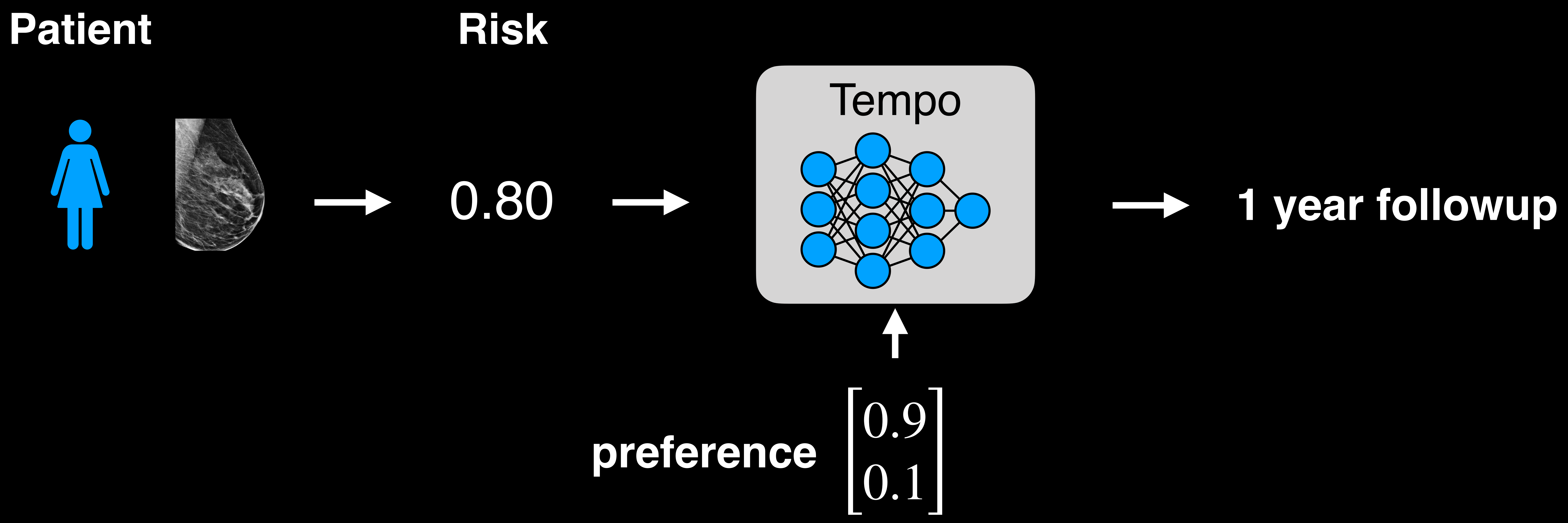


Reward = λ_1 Benefit - λ_2Cost

Results: MGH Screening Efficiency



Supporting diverse clinical requirements



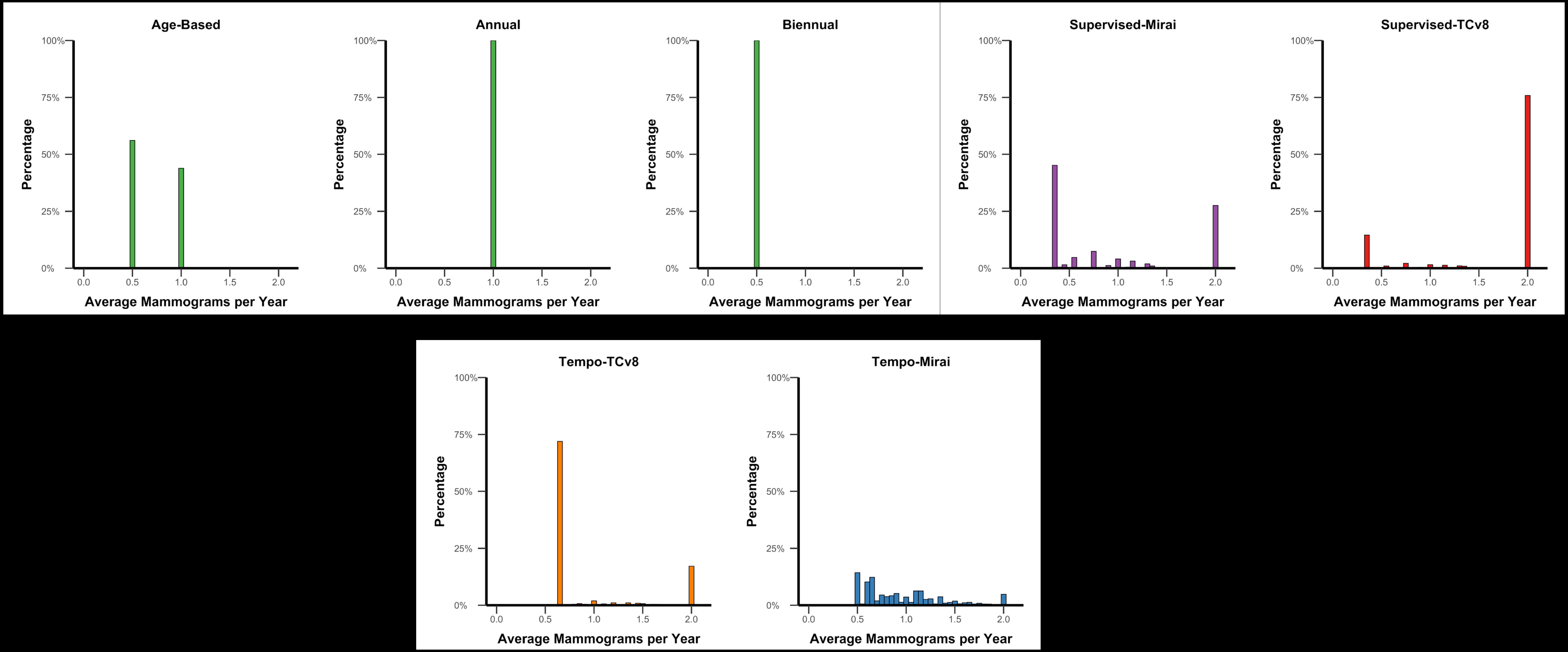
Trained across possible (λ_1, λ_2) to maximize:

Reward = λ_1 Early Detection Benefit - λ_2Screening Cost

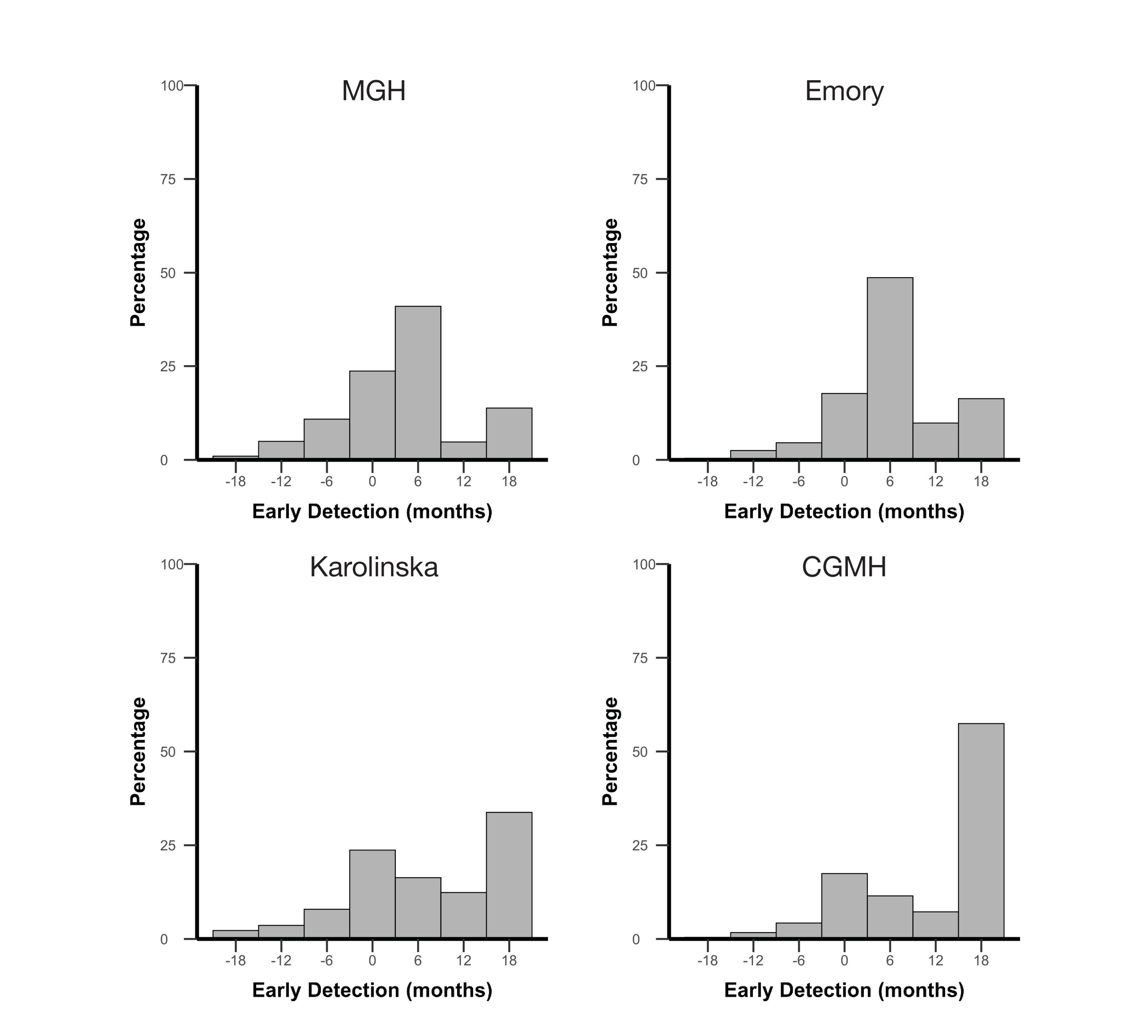
A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation

Runzhe Yang, Xingyuan Sun, Karthik Narasimhan

Results: Screening Frequency



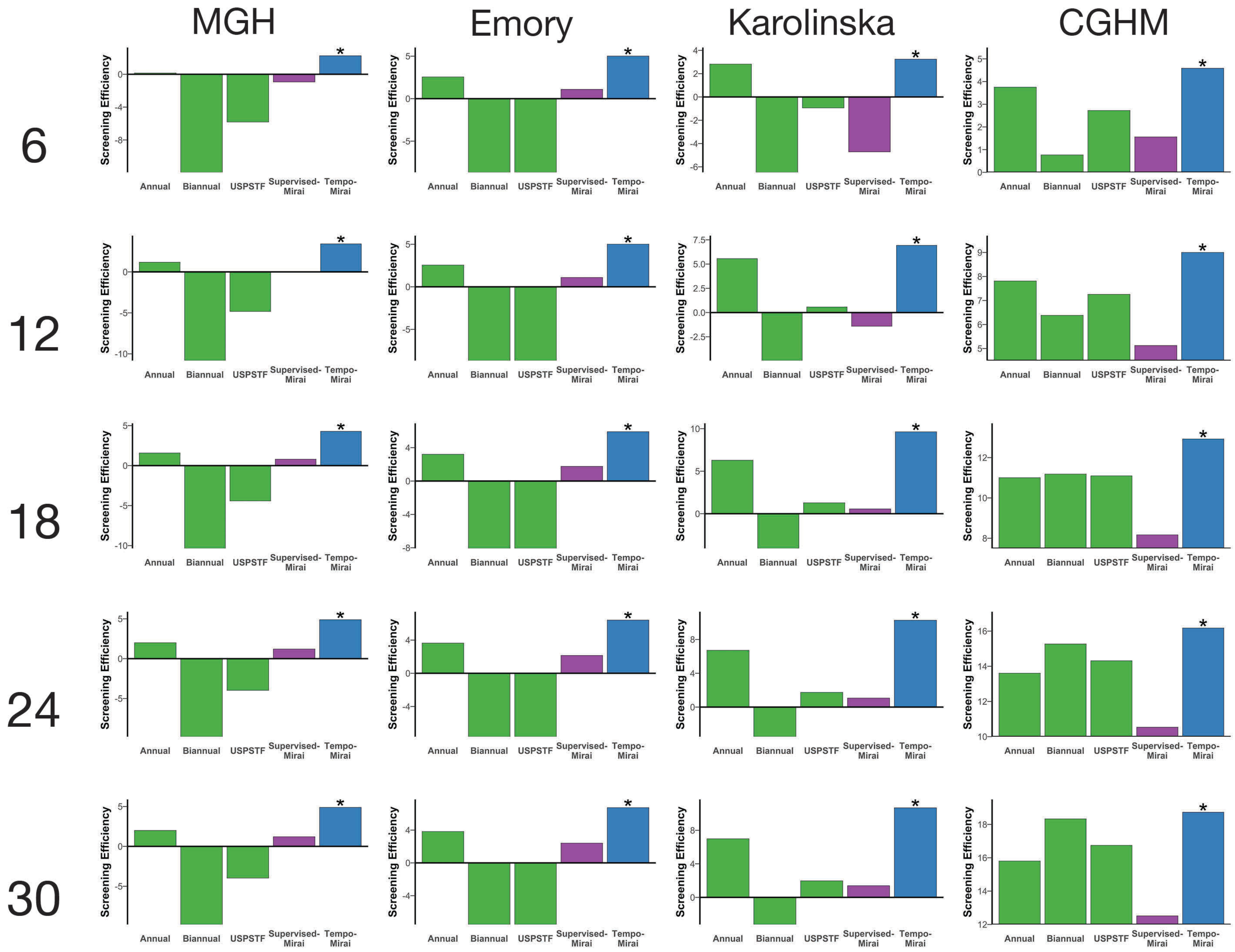
Results: Early Detection



Results: Robustness

- Policies trained with 18 month assumption

- Tempo-Mirai more efficient across assumptions



Results: Risk Progression