

Optimizing risk-based breast cancer screening policies with reinforcement learning

Adam Yala (✉ adamyala@mit.edu)

Massachusetts Institute of Technology

Peter Mikhael

Massachusetts Institute of Technology <https://orcid.org/0000-0002-6030-1636>

Constance Lehman

Massachusetts General Hospital

Gigin Lin

Chang Gung Memorial Hospital <https://orcid.org/0000-0001-7246-1058>

Fredrik Strand

Karolinska University Hospital

Yung-Liang Wang

Chang Gung Memorial Hospital

Kevin Hughes

Massachusetts General Hospital

Siddharth Satuluru

UCLA

Thomas Kim

Georgia Institute of Technology

Imon Banerjee

Emory University <https://orcid.org/0000-0002-3327-8004>

Judy Gichoya

Emory University

Hari Trivedi

Emory University

Regina Barzilay

MIT

Article

Keywords: health screening, breast cancer, risk-based breast cancer screening policies, reinforcement learning

Posted Date: June 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-104592/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Medicine on January 13th, 2022.

See the published version at <https://doi.org/10.1038/s41591-021-01599-w>.

Optimizing risk-based breast cancer screening policies with reinforcement learning

Authors: Adam Yala*¹, Peter G Mikhael¹, Constance Lehman², Gigin Lin³, Fredrik Strand^{4,5}, Yung-Liang Wan³, Kevin Hughes⁶, Siddharth Satuluru⁷, Thomas Kim⁸, Imon Banerjee⁹, Judy Gichoya¹⁰, Hari Trivedi¹⁰, Regina Barzilay¹

Affiliations:

¹ Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

² Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA.

³ Department of Medical Imaging and Intervention, Chang Gung Memorial Hospital at Linkou, Taoyuan 333, Taiwan

⁴ Breast Radiology Unit, Department of Imaging and Physiology, Karolinska University Hospital, Stockholm, Sweden.

⁵ Department of Oncology-Pathology, Karolinska Institute, Stockholm, Sweden.

⁶ Division of Surgical Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA.

⁷ Department of Computer Science, University of California Los Angeles, Los Angeles, CA 90095, USA.

⁸ Department of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332, USA.

⁹ Department of Biomedical Informatics, Emory University, Atlanta, GA 30332, USA.

¹⁰ Department of Radiology, Emory University, Atlanta, GA 30332, USA.

*To whom correspondence should be addressed: adamyala@csail.mit.edu

Abstract

Screening programs must balance the benefits of early detection against the costs of over screening. Achieving this goal relies on two complementary technologies: (1) the ability to assess patient risk, (2) the ability to develop personalized screening programs given that risk. While methodologies for assessing patient risk have significantly improved with new advances in deep learning applied to imaging and genetics, our ability to personalize screening policies still lags behind. Here, we introduce a novel reinforcement learning-based framework for personalized screening, Tempo, and demonstrate its efficacy in the context of breast cancer. We trained our risk-based screening policies on a large screening mammography dataset from Massachusetts General Hospital (MGH) USA and validated them on held-out patients from MGH, and on external datasets from Emory USA, Karolinska Sweden and Chang Gung Memorial Hospital (CGMH) Taiwan. Across all test sets, we found that a Tempo policy combined with an image-based AI risk model was significantly more efficient than current regimes used in clinical practice in terms of simulated early detection per screen frequency. Moreover, we showed that the same Tempo policy can be easily adapted to a wide range of possible screening preferences, allowing clinicians to select their desired early detection to screening cost trade-off without training a new policy. Finally, we demonstrated Tempo policies based on AI-based risk models out performed Tempo policies based on less accurate clinical risk models. Altogether, our results show that pairing AI-based risk models with agile AI-designed screening policies has the potential to improve screening programs, advancing early detection while reducing over-screening.

Introduction

For multiple diseases, early detection significantly improves patient outcomes [1, 2]. This motivates considerable investments in population-wide screening programs [3, 4] such as mammography for breast cancer. To be effective and economically viable, these programs must find the right balance between early detection and overscreening. This capacity builds on two complementary technologies: (1) ability to accurately assess patient risk at a given time point; (2) ability to design screening regimes based on this risk. With recent advances in deep learning, imaging and genetics, risk assessment technologies are rapidly improving [5, 6]. However, our ability to utilize these predictions to personalize screening regimes lags behind. This deficiency is particularly apparent when the screening system has limited throughput.

In this paper, we focus on the design of screening regimes attuned to the increased capacity of the modern risk assessment models. The need for new methods to personalize screening is motivated by a substantial change in risk assessment algorithms. Traditional risk assessment models rely on a number of categorical variables encoding patient demographics and clinical history combined with traditional statistical models to predict risk [7, 8]. These scores are relatively static throughout a patient's lifetime, with changes typically driven by the patient's age. Moreover, the limited predictive capacity of these risk models restricts the scope of recommendations they support and, consequently, their impact on the screening regime. Current guidelines divide the population into a few large groups, most often discriminating predicted high-risk patients from the rest, and recommend the same screening frequency to all the members of that cohort [9-11]. As a result, there remains large opportunities to further personalize care.

The power of novel, AI-driven risk models [6, 12-14] has given us an opportunity to fundamentally transform population screening. Deep learning algorithms enable these risk models to operate over raw patient data such as imaging in addition to traditional expert-specified categorical variables. Moreover, these models can detect highly complex dependencies which further strengthens their predictive capacity relative to traditional methods. One distinctive feature of these risk models is that their predictions may fluctuate over time as the patient's raw data evolves. This suggests that screening regimes need to be flexibly adjusted with changes in risk and be optimized over a patient's lifetime. We hypothesize that by pairing AI-based risk models and agile AI-based screening regimes, we can improve early detection, while lowering the overall cost of screening. This paper presents empirical findings that support this hypothesis in the area of breast cancer screening. The core methodology is applicable to other disease areas and other types of risk models beyond imaging.

Results

Overview of Algorithm

In computational terms, we can view breast cancer screening as a sequential decision task, where we wish to develop a policy, i.e., a screening guideline, that predicts a followup recommendation for each patient to maximize their chances at early detection while minimizing

their screening costs. Intuitively, such a policy should recommend infrequent screenings for low risk patients, while prescribing a higher frequency for patients at increased risk. The question is how to personalize screening intervals based on a patient's risk profile. More formally, we can cast the screening problem as a Markov Decision Process (MDP), where a patient's state is their risk assessment, the possible actions are different follow-up recommendations (e.g., six months, two years, etc.), and rewards are a combination of expected early detection benefits minus screening costs. This formulation enables us to find the best possible policy for this MDP with Reinforcement Learning (RL) algorithms [15, 16]. RL algorithms are commonly used in game playing [17, 18] and robotics [19-23]. The closest to our work is research on predictive maintenance [24-26] where the goal is to establish an optimal regime for prophylactic maintenance of equipment while controlling for costs and the likelihood of failure. We show how to leverage these methods for determining effective cancer screening policies.

Applying RL in this context poses a unique challenge, namely the estimation of patient trajectories from retrospective data. The training data pertaining to individual patients only contains information about their risk at the time points when their mammogram was taken. However, to determine whether the algorithm makes the correct recommendation, we need to know the risk assessment at intermediate points. Therefore, we design an algorithm that learns to extrapolate a patient's risk at unobserved time points from their observed screenings. This estimation evolves as new mammograms of the patient become available. With the access to these predictions, we can guide our reinforcement learner to adjust its actions according to the estimated risk. Using the retrospective trajectories as our simulation environment, we train screening policies to maximize the future reward given the patient's evolving risk assessments, as illustrated in Figure 1. In doing so, our trained screening policies are specialized to the dynamics and subtleties of the underlying risk model.

Our full framework, named Tempo, is depicted in Figure 2. As described above, we first train a risk progression neural network to predict future risk assessments given previous assessments. This model is then used to estimate patient risk at unobserved timepoints. Next, we train a screening policy neural network to maximize the expected future reward, i.e combination of early detection and screening cost, using envelope-Q learning [27], a multiple objective reinforcement learning algorithm. The input of our screening policies is the patient's risk assessment, and desired weighting between rewards (i.e screening preference). The output of the policy is a recommendation for when to return for the next screen, ranging from six months to three years in the future, in multiples of six months. Our reward balances two contrasting aspects, one reflecting the imaging cost, i.e., the average mammograms a year recommended by the policy, and one modeling early detection benefit relative to the retrospective screening trajectory. Our early detection reward measures the time difference in months between each patient's recommended screening date, if it was after their last negative mammogram, and their actual diagnosis date. We evaluate screening policies by simulating their recommendations for heldout patients. The exact reward details as well as the neural network architectures used are elaborated in the *Materials and Methods*.

Experimental Setup

We developed Tempo using the Massachusetts General Hospital (MGH) dataset, which consists of 137,682, 16,634 and 17,119 exams from 43,749, 5,399 and 5,525 patients, for the training, validation, and testing sets respectively. For each exam, we had access to Mirai [12] and TCv8 [8] risk assessments. Mirai [12] is a recently proposed mammography-based risk model which predicts risk at multiple time points, and TCv8 [8] is a traditional risk model that combines a variety of risk factors including age, family history, and hormonal factors. For Tempo to be broadly applicable, its screening policies must be validated in new clinical environments and patient populations. To this end, we also validated Tempo on representative datasets from Emory in the US consisting of 22,094 exams from 10,369 patients, Karolinska in Sweden consisting of 14,356 exams from 7,191 patients and Chang Gung Memorial Hospital (CGMH) in Taiwan consisting of 12,280 exams from 12,280 patients. For each exam in the Emory, Karolinska and CGMH datasets, we obtained Mirai risk assessments. We note that the Emory, Karolinska and CGMH datasets were only used for held-out testing. The demographics for all test sets are reported in Table 1 and more detailed demographics for each dataset are shown in Tables S1, S2, S3 and S4. All datasets are described in detail in the *Materials and Methods*.

Our primary objective was to develop personalized screening policies that would outperform current guidelines, improving early detection while reducing screening costs. To this end, we compared Tempo, developed with both TCv8 and Mirai risk assessments, to existing screening guidelines, including annual screening, biennial screening, and a hybrid screening strategy recommended by the USPSTF [11] which switches from annual screening to biennial screening at age 55. We refer to Tempo with Mirai and TCv8 risk assessments as Tempo-Mirai and Tempo-TCv8, respectively. We utilized a GRU to estimate risk progression for Tempo Mirai and a deterministic model, Static risk, to estimate risk progression for Tempo-TCv8. Our risk progression methods are described in detail in the *Materials and Methods* and are validated in Table S5. To quantify the benefit of using our reinforcement learning approach to develop screening policies, we also evaluated policies trained with supervised learning, named Supervised-Mirai and Supervised-TCv8 respectively. We note that given full screening trajectories, we can easily compute the optimal screening recommendations in hindsight. Instead of maximizing the overall reward with reinforcement learning, Supervised-Mirai and Supervised-TCv8 are trained to predict these optimal followup recommendations at each time step. These baselines are detailed in the *Materials and Methods*. For each policy, we measure its screening cost in terms of the average number of mammograms it recommends per year and its early detection benefit in months relative to historical screening. Our early detection metric assumed that early screening, following a patient's last negative mammogram, could offer a maximum early detection benefit of 18 months. We note that our early detection benefit metric is local and institution specific, as different institutions have different screening patterns. To directly compare policies that recommend differing numbers of mammograms, we also evaluated the efficiency of each policy, as measured by the early detection benefit in months divided by the number of mammograms per year recommended.

Evaluating personalized screening policies

The results of all screening policies across the MGH, Emory, Karolinska and CGMH test sets are illustrated in Table 2. We utilized the same Tempo-Mirai operating point across all test sets. We illustrate the performance of Tempo across different operating points, i.e screening preferences, in all test sets in Figure 3.

On the MGH test set, the annual and USPSTF guidelines obtained screening efficiencies, i.e. early detection benefit per screening cost, of 1.58 (95% CI 0.54, 2.58) and -4.42 (95% CI -5.83, -3.12). In contrast, Tempo-Mirai, Tempo-TCv8, and Supervised-Mirai obtained screening efficiencies of 4.29 (95% CI 3.17, 5.25), 2.16 (95% CI 1.18, 3.40), and 0.80 (95% CI -0.58, 2.12) respectively. We found that Tempo-Mirai was significantly more efficient than Tempo-TCv8, Supervised-Mirai and annual screening ($p < 0.001$, $p < 0.001$, $p < 0.001$), obtaining higher early detection per screening cost. Specifically, Tempo-Mirai obtained an early detection benefit of 4.10 (95% CI 3.06, 4.96) months while recommending 0.96 (95% CI 0.94, 0.96) mammograms per year, while the annual guideline obtained an early detection benefit of 1.58 (95% CI 0.54, 2.58) months while recommending 1.0 mammograms per year.

In addition to overall performance on the test sets, we also studied the histogram of early detection benefits in Figure S1, and the histogram of recommended screening frequencies in Figure 4 and Figure S2. We note that all trained policies (e.g. Tempo-Mirai, Supervised-Mirai) have the same set of possible recommendations ranging from a six month to three year screening followup, but we found that Supervised-Mirai only selected two options, recommending either 6 months or three years of followup. In contrast, Tempo-Mirai at our chosen operating point leveraged follow up recommendations of 6 months, one year and two years. As shown on Figure 4, we found that Tempo-Mirai offers a wider range of recommended frequencies than other methods, reflecting a larger degree for personalization. This reflects the optimization differences between the two policies. Tempo-Mirai is optimized to maximize overall reward across patient trajectories, as measured by early detection and screening cost, and does not receive any explicit guidance on the correct recommendation given a specific risk assessment. As a result, Tempo-Mirai has the flexibility to explore a wide range of possible recommendations during training to identify high performing policies. In contrast, Supervised-Mirai has a rigid modeling objective; it is instead trained to predict the optimal (i.e. correct in hindsight) screening recommendation from each risk assessment which is difficult given the uncertainty of real world risk models.

To understand the flexibility of Tempo-based policies, we plotted the performance of each policy in Figure 3 while varying the screening preference (i.e. operating point), which specifies the desired balance between early detection and screening cost. Across a wide range of possible operating points, Tempo-Mirai outperformed other policies in increasing early detection and reducing screening costs, demonstrating that the policy can be easily adapted to suit clinical requirements without retraining.

Next, we analyzed Tempo-Mirai's ability to generalize to new populations. To this end, we tested Tempo-Mirai, which was trained on MGH data, on test sets from Emory, Karolinska and CGMH. In the Emory test set, Tempo-Mirai, Supervised-Mirai and annual screening obtained efficiencies of 5.92 (95% CI 5.06, 6.54), 1.76 (95% CI 0.41, 2.86) and 3.21 (95% CI 2.37, 3.88) respectively. In the Karolinska test set, Tempo-Mirai, Supervised-Mirai and annual screening obtained efficiencies of 9.63 (95% CI 8.53, 10.72), 0.56 (95% CI -0.98, 1.55) and 6.29 (95% CI 5.76, 6.85) respectively. In the CGMH test set, Tempo-Mirai, Supervised-Mirai and annual screening obtained efficiencies of 12.92 (95% CI 11.54, 14.41), 8.17 (95% CI 6.30, 10.53) and 11.00 (95% CI 9.86, 12.28) respectively. Tempo-Mirai was significantly more efficient than Supervised-Mirai and annual screening in all test sets with $p < 0.001$, $p < 0.001$ at Emory, $p < 0.001$, $p < 0.001$ at Karolinska and $p < 0.001$, $p = 0.02$ at CGMH.

While the above results show that Tempo-Mirai consistently improved over alternate policies in screening efficiency, we also observed that the absolute magnitude of early detection varied substantially across different datasets. For instance, annual screening obtaining early detection benefits of 1.58 (95% CI 0.54, 2.58), 3.21 (95% CI 2.37, 3.88), 6.29% (95% CI 5.76, 6.85) and 11.0 (95% CI 9.86, 12.28) months in the MGH, Emory, Karolinska and CGMH test sets respectively. This difference can be attributed to the different rates of screening across the datasets; patients with future cancer at MGH, Emory, Karolinska and CGMH obtained an average of 0.93, 0.94, 0.80, and 0.66 mammograms per year. These differences are further detailed in Tables S1, S2, S3 and S4, which report detailed demographics of each dataset.

We also noted that Tempo-Mirai recommended different amounts of screening across the MGH, Emory, Karolinska and CGMH test sets, recommending an average of 0.96 (95% CI 0.94, 0.97), 1.08 (95% CI 1.07, 1.08), 0.75 (95% CI 0.74, 0.76), 0.88 (95% CI 0.87, 0.89) mammograms per year respectively. This difference can be attributed to differences in cancer incidence between the different centers. The five-year cancer incidence at MGH, Emory and CGMH are 2.2%, 3.0% and 1.8% respectively, and we expect Tempo to recommend higher rates of screening for higher risk populations. However, the model can offer a diverse set of possible operating points across all test sets, as illustrated in Figure 3, our results indicate that different hospitals may need to input different operating points to obtain the same average screening volume.

Subgroup Analysis

We also investigated how our policies performed for different patient subgroups by age and breast density on the MGH test set in Table S6. We highlight the results of Tempo-Mirai, which obtained an efficiency of 4.29 (95% CI 3.17, 5.25) on the entire MGH test set. When grouping patients by age, Tempo-Mirai obtained efficiencies of 3.41 (95% CI 1.44, 5.53) and 4.45 (95% CI 3.49, 6.03) for patients aged below and above 55 respectively. When grouping patients by breast density category, Tempo-Mirai obtained efficiencies of 4.10 (95% CI 2.85, 5.48) and 4.49 months (95% CI 2.86, 6.35) for patients with non-dense and dense breasts respectively, where non-dense refers to the BI-RADS categories of almost entirely fatty or scattered areas of fibroglandular tissue and dense refers to the BI-RADS categories of heterogeneously dense or extremely dense.

Robustness to Assumptions

Our empirical results across the different test sets depend on the exact choice of assumptions of our early detection metric. As illustrated in Figure 1, our early detection metric, measured the time difference in months between each patient's recommended screening date, and their diagnosis date. Our metric assumed that the maximum early detection benefit obtained through earlier screening was 18 months. To test our models robustness to this assumption, we also evaluated Tempo-Mirai, Supervised-Mirai and annual screening across all test sets when setting our maximum early detection benefit assumption to 6 months, 12 months, 18 months and 24 months. We note that we did not re-train Tempo-Mirai for this analysis, and that Tempo-Mirai was originally trained using the 18 month assumption. For each policy, we measured its screening efficiency, i.e the early detection benefit divided by the number of mammograms per year recommended, to enable head-to-head comparison between policies that recommend different screening volumes. As shown in Figure S3, Tempo Mirai is more efficient than annual screening across all datasets across all assumptions. This result is further supported by the histogram of early detection benefits shown in Figure S1.

Discussion

We developed a reinforcement learning framework for personalized screening, Tempo, to predict followup recommendations from patient risk assessments. We demonstrated that a Tempo policy based on Mirai risk assessments was significantly more efficient than annual screening, achieving earlier detection per screening cost. Moreover, we showed that the same Tempo policy can be adapted to a wide range of possible screening preferences, and that policies that leverage more accurate risk models, i.e Mirai, outperform those based on less accurate risk models, i.e Tyrer-Cuzick. We found that policies developed on data from MGH USA generalized to held-out test sets in Emory USA, Karolinska Sweden and Chang Gung Memorial Hospital Taiwan, and significantly outperformed annual screening and our supervised learning baseline there as well. Finally, we demonstrated our results were robust across a range of possible assumptions for our early detection metric.

Our screening policies can be easily implemented in any screening clinic where Mirai risk assessments are collected. Clinicians can retrospectively validate our trained screening policies across a wide range of possible operating points on their own screening population and choose an operating point to achieve the desired balance between screening volume and early detection benefit. The installed policy can then offer clinicians suggested risk-based followup intervals immediately following a patient's risk assessment. Depending on clinical requirements, Tempo can be utilized to significantly reduce the volume of screening for a fixed early detection target or to significantly improve early detection for a fixed screening budget. For instance, we showed that Tempo-Mirai could obtain better early detection than annual screening at Karolinska and CGMH while reducing screening by 25% and 12% respectively. Due to the scale and cost of breast cancer screening, even modest improvements in screening guidelines have the potential to benefit a wide patient population.

Our study is complementary to a rich body of work surrounding risk-based screening [7, 28-30]. Several guidelines already recommend supplemental imaging or chemo-prevention based on risk assessments [29, 31, 32] and recent results from the DENSE trial [33] have shown that a breast density based screening strategy could significantly reduce interval cancers compared to current screening. Our work is most closely related to the MyPeBS trial [34], which prospectively compares a personalized screening followup strategy based either Tyrer-Cuzick [8] or MammoRisk [35] risk assessments to current national recommendations. These studies point to substantial clinical interest in risk-based screening, however, current methods for devising screening policies rely on categorizing patients into a few coarse categories (e.g low and high risk), limiting personalization. Our study provides a data-driven alternative for clinical decision making and can be easily integrated into a screening trial and routine patient care.

Our work is also related to a large volume of modeling studies focused on breast cancer [36-42]. Typically, these approaches operate over a model of disease progression that characterizes how patients transition between healthy and disease states. The transitions are informed by patient features and impact the likelihood of different observations, such as a palpable lump.

Their probabilities can be estimated from retrospective data or retrieved from the literature. The approaches then work to identify the optimal screening policy under the specified disease progression model. While these approaches were first to demonstrate the feasibility of developing personalized screening policies, they have several limitations that restrict their practical use in clinical settings. First, the postulated disease progression model does not capture the full complexity and uncertainty in our understanding of cancer. Second, the methods generally assume that a patient's features are fixed and do not evolve over their lifetime [42]. This assumption does not hold in general and is not applicable to modern AI-based risk models that are sensitive to changes in patient health. In contrast, our framework does not assume a complete disease progression model; instead, it assumes access to a risk model (rather than discrete set of states) and to a reward function that measures the performance of a screening trajectory given observational data. This relaxed assumption allows us to optimize screening policies directly on observed patient trajectories, which contain the full diversity of cancer diagnoses, and to validate our policies on held-out patient populations, which may differ in their cancer characteristics, such as Emory, Karolinska and CGMH. Moreover, because our policies are optimized across a large set of longitudinal patient trajectories, they are adapted to dynamics of the underlying risk models.

This study focuses on breast cancer screening using image-based risk models. However, our framework is flexible and can be readily utilized for other diseases, other forms of risk models, and other definitions of early detection benefit. For instance, it can easily incorporate richer representations of the cancer outcomes. Recent work has highlighted concerns about the potential over-treatment of ductal carcinoma in-situ [43]. Tempo policies can take these differences into account by leveraging separate reward metrics for the early detection of invasive and in-situ cancers. In this scenario, Tempo policies would be trained using three reward metrics, i.e early detection of invasive cancers, early detection of in-situ cancers and screening cost, and clinicians would select a Tempo operating point (i.e screening preference) that achieves the desired balance between the three metrics. In a similar fashion, our framework can be used to optimize more refined definitions of early detection benefit that account for properties of the cancer (e.g tumor size, grade) at the time of diagnosis. For instance, given access to a patient's tumor properties, a cancer mortality model and a cancer growth model, a sophisticated early detection metric could directly estimate the reduced mortality risk if the patient's cancer had been at an earlier time point. Given a patient's age, this metric could also directly be tied to Quality Adjusted Life Years (QALYs). Similarly, more sophisticated measures of screening cost, that take into account varying false positive risks depending on patient characteristics such as breast density, could be used to further refine screening policies. In this sense, prior work in modeling cancer mortality and screening benefits [36-41], is complementary to our own. We expect that the utility of Tempo, which is agnostic to the underlying choice of screening metrics and risk model, will increase as risk models and outcomes metrics are further refined across more diseases.

There are multiple future directions that can further improve personalized screening algorithms. While our policies were trained to maximize the average early detection across the population,

one could instead optimize the model to also minimize the worst case risk of delaying a diagnosis; in this sense, the diverse work in optimization literature [44, 45] can be complementary as well. There also are opportunities to leverage more rich patient information in developing personalized screening policies. While our method focused on predicting followup recommendations given risk estimates from established risk models, one could instead directly input rich patient information such as a patient's mammograms and family history, into the screening policy. Directly learning to interpret this information for the purpose of personalized screening in an end-to-end fashion may result in more accurate policies. Moreover, the action space of our method could be expanded to also include different types of screening recommendations, such as leveraging MRI or mammograms, and future work could separately model the costs and benefits of each modality. Finally, given improved screening policies, future work could also recalculate the earliest and latest age such that screening is still cost-effective for a patient.

This study has several limitations. Our early detection metric assumed that cancer is detectable up to a fixed time (18 months) before diagnosis. While we found that the trends reported in our study were robust to different values of this assumption, ranging from six months to 24 months, none of these assumptions are individually correct across all cancers as the early detection potential of a tumor depends on that tumor's characteristics at the time of diagnosis. Our simulations also did not account for the sensitivity of screening mammography or the probability of a patient entering the clinic with a palpable lump if their diagnosis is overly delayed. Moreover, our early detection metric is defined relative to historical screening, which is institution specific. As a result, it is not possible to directly compare a screening policy's performance across multiple sites. While our framework is agnostic to the specifics of how the rewards are formulated, further research using more refined early detection metrics, such as QALY's, that explicitly model tumor characteristics at the time of detection and tumor growth is needed. Finally, prospective trials are necessary to assess the efficacy of these models in clinical care before wide-spread adoption.

Materials and Methods

Study Design

The primary objective of this study was to develop a new algorithm to create flexible personalized screening policies. We hypothesized that our Tempo policies could offer improved early detection benefits over annual screening without requiring more screening. Moreover, we hypothesized that these policies would generalize to new institutions.

Dataset Description

To develop Tempo, we collected consecutive screening mammograms and detailed risk information at the time of mammography from 80,134 patients screened between January 1st, 2009 and December 31st, 2016 at Massachusetts General Hospital (MGH) under approval of MGH's Institutional review board with a waiver for written informed consent and in compliance with the Health Portability and Accountability Act. We obtained outcomes through linkage to a

local five-hospital registry in the Massachusetts General Brigham healthcare system, alongside pathology findings from MGH's mammography electronic medical record. We excluded patients who were diagnosed with other cancers (e.g., sarcoma, etc.) in the breast, or did not have all four views (L CC, L MLO, R CC, R MLO). For patients who develop cancer, we excluded exams within six months of diagnosis. For patients who did not develop cancer, we excluded exams within three years of the last followup screen. We note that six months and three years are the minimum and maximum follow up recommendations for Tempo, and so this exclusion enabled us to ensure that simulations always occur within the bounds of observed data. This resulted in 54,673 patients who were randomly split into 43,749 for training, 5,399 for development, and 5,525 for testing. We note that this dataset was also used to develop Mirai [12], and so we used the same training, development and testing splits. Since each patient had multiple exams, this in turn resulted in 137,682, 16,634, and 17,119 exams of training, development, and testing, respectively. All mammograms were acquired on Hologic machines. For each exam, we obtained Mirai [12] risk assessments, as well as Tyrer-Cuzick version 8 (TCv8) risk assessments. Detailed demographics of this dataset are available in Table S1 and the dataset construction procedure is shown in Figure S4.

To evaluate the ability of Tempo policies to generalize to new populations, we collected the Emory, Karolinska and CGMH datasets under approval of the relevant Institutional Review Boards with a waiver for written informed consent. To create the Emory test set, which contains a large representation of african american women, we extracted eight years of mammograms from an institutional database of all comers for screening mammography from 2013-2020 and randomly selected 30% of women, totalling 28,994 patients. All mammograms were acquired on Hologic machines. We collected outcomes from pathology findings from Emory's mammography electronic medical record. We obtained Mirai risk assessments for each exam. As with the MGH dataset, we excluded exams within six months of diagnosis. For patients who did not develop cancer, we excluded exams within three years of the last followup screen. This resulted in a total of 22,030 exams from 10,340 patients. Detailed demographics of this dataset are shown in Table S2 and the dataset construction procedure is shown in Figure S4.

The Karolinska test set was extracted from the Cohort of Screen-Aged Women (CSAW) [46]. All women aged 40-74 within the Karolinska University uptake area who had attended screening and were diagnosed with breast cancer, without implants and without prior breast cancer, from 2008-2016 were included, as well as a random sample of controls with at least two years follow-up, from the same time period. The full Karolinska case-control and validation datasets included 11,301 and 2,580 women, respectively. A random subset of 9,484 patients in total were selected for inclusion in this study. We included all mammograms, acquired on Hologic machines, from 2008-2016 for the included women that contained all four views (L CC, L MLO, R CC, R MLO), resulting in 14,362 exams from 7,193 patients. We excluded exams within six months of a cancer diagnosis. For patients who did not develop breast cancer, we excluded exams within three years of the last screening followup. Because of the case-control dataset design, this dataset has a much higher ratio of patients that develop cancer, relative to the 1.9% incidence reported in the CSAW cohort [45]. To take this into account, we randomly resampled patients who did not develop cancer from our cohort to produce a larger dataset with a 1.9% cancer incidence, resulting in a total of 93,052 exams from 7,193 patients. Detailed

demographics of this dataset are shown in Table S3 and the dataset construction procedure is shown in Figure S4. For each exam, we obtained Mirai [12] risk assessments.

To create the CGMH test set, which consisted of 12,280 exams from 12,280 patients, we selected random women undergoing screening mammography there between 2010-2011 that were aged 45-70. Women aged 40-44 were also included if they had a family history of breast cancer, following local screening guidelines. All mammograms were acquired on Hologic machines. Cancer outcomes were obtained from the national cancer registry. More details for this dataset is available in Table S4 and Figure S4. We excluded patients with unknown age. For each patient that developed cancer, we also manually collected all the dates of their future screenings from 2010-2020 through chart review. This allowed us to estimate early detection benefits relative to historical screening. We did not collect future screening dates for patients who did not develop cancer. For patients who developed cancer, we excluded exams within six months of diagnosis, while for patients who did not develop cancer, we excluded exams within three years of the last followup screen. For each exam, we obtained Mirai [12] risk assessments.

For patients with multiple exams in a dataset, we considered each exam as a possible trajectory starting point and evaluated screening policies across all starting points. We account for patients contributing multiple correlated trajectories in our statistical analysis by using a clustered bootstrap procedure with 5000 samples. For each trajectory, we considered its censor time as either the date of cancer diagnosis via biopsy or the date of last screening follow up. We designed our screening policies to offer a minimum followup recommendation of six months and a maximum followup recommendation of three years. Because our followup intervals were in increments of six months, we discretized time across all trajectories into six month timesteps. To ensure that our simulations always occurred within the time frame of the observed data, we excluded starting points where cancer was diagnosed in less time than the minimum action (six months). For screening trajectories without a cancer diagnosis, we excluded starting points where the time to the last screening followup was less than the maximum action (three years).

Reward Design

We considered two rewards in our simulation environment, one measuring imaging cost and early detection benefit respectively. We modeled our imaging cost reward as the negative amount of mammograms per year recommended by a policy. To model early detection benefits, we measured the time difference in six-month timesteps between each patient's recommended screening date, if it was after their last negative mammogram, and their actual diagnosis date. We then converted this value into months. We defined a patient's diagnosis date as the date of their positive biopsy. Negative values of this reward imply a delayed diagnosis, and positive values imply relative screening benefit over the retrospective trajectory. We capped maximum early detection benefit for any patient to be 18 months and did not cap the possible screening delay. As a result, if a patient's last negative mammogram was three years before their cancer diagnosis, and a screening policy recommended a mammogram two years and one year before a patient's cancer diagnosis, we assigned this trajectory an early detection benefit of 18 months.

We provide additional analysis for different possible assumptions for the maximum screening benefit in Figure S3.

Risk Progression Models

We considered two possible methods to estimate risk progression, namely Static Risk, which always predicted that a patient's risk at the next time step would be the same as at the last time step, and a Recurrent Neural Network (RNN). For our RNN model, we utilized a Gated Recurrent Unit (GRU) [47] with an additive hazard layer [12] to estimate future risk. The GRU was trained to take previous risk observations, observed or predicted, to predict risk at the next time step, and minimize the Kullback-Leibler (KL) divergence between predicted risk assessments and observed ones on the MGH training set. We experimented with different learning rates, hidden sizes, numbers of layers and dropout, and chose the model that obtained the lowest validation KL divergence on the MGH validation set. Our final risk progression GRU had two layers, a hidden dimension size of 100, a dropout of 0.25, and was trained for 30 epochs with a learning rate of $1e-3$ using the Adam optimizer. The outputs of our risk progression model for Tempo-Mirai are visualized in Figure S5. Given a trained risk progression model, we can now estimate unobserved risk assessments auto-regressively. At each time step, the model takes as input the previous risk assessment, the prior hidden state, using the previous predicted assessment if the real one is not available, and predicts the risk assessment at the next time step. We validated our risk progression network on the MGH, Emory and Karolinska test sets in Table S5 and note that our GRU outperformed the static risk baseline in all datasets. Since we only collected only one exam for each patient in the CGMH test set, we could not validate the risk progression network on that test set. The implementation for each risk progression as well as our hyper-parameter search is available in our code release.

Personalized Screening Models

We implemented our personalized screening policy as multiple layer perceptron, which took as input a risk assessment and weighting between rewards and predicted the Q-value for each action, i.e follow up recommendation, across the rewards. This network was trained using Envelope Q-Learning [27]. Following recent work in deep reinforcement learning [17, 48], we used an experience replay buffer to reduce correlation between our training batches, and we utilized a target Q-network [17] to stabilize training updates. We experimented with different numbers of layers, hidden dimension sizes, learning rates, dropouts, exploration epsilons, target network reset rates and weight decay rates. We note that we conducted the same grid searches for Tempo-Mirai and Tempo-TCv8, and chose each model to maximize the average reward on the MGH validation set. Our final Tempo-Mirai model had 6 layers, each with 256 hidden units, followed by ReLU nonlinearities. It was trained for 30 epochs using a learning rate of $1e-3$, a dropout of 0.25 and a weight decay of 0.01 using the Adam optimizer, and the target network was reset every 1000 batches. Our final Tempo-TCv8 model had 4 layers, each with 256 hidden units, followed by ReLU nonlinearities. It was trained for 30 epochs using a learning rate of $1e-3$, a dropout of 0.25 and a weight decay of 0 using the Adam optimizer, and the target network was reset every 1000 batches. The implementation of each risk policy, the training code as well as our hyper-parameter searches are available in our code release. For both

Tempo-Mirai and Tempo-TCv8, we chose a reward weighting to approximately match the screening cost of annual screening on the MGH development set. We used reward weights for screening cost and early detection benefit of (0.50, 3.0) and (0.77, 3.0) for Tempo-Mirai and Tempo-TCv8 respectively. For Tempo-TCv8, this resulted in a reward weight of 0.77 and 3.0. for Tempo-TCv8. We utilized the same reward weighting across all test sets.

Supervised Learning Baseline

We implemented our supervised learning baselines, Supervised-Mirai and Supervised-TCv8, as a multiple layer perceptron, which took as input a risk assessment and predicted a probability distribution across followup recommendations. This network was trained to minimize the cross entropy loss between its actions, and the optimal sequence of actions. We computed optimal actions for each patient to maximize our rewards metrics. For patients who did not develop cancer within the time period of the maximum follow up recommendation, the optimal action was the maximum followup recommendation of three years. For patients who developed cancer, the optimal action was to recommend a screening followup in the time step following the last negative mammogram. Unlike Tempo-Mirai which is trained to maximize trajectory level rewards using reinforcement learning, Supervised-Mirai is trained to maximize the likelihood of the optimal sequence of actions. As a result, Supervised-Mirai does not benefit from observing how its own errors compound across the trajectory at training time. For each supervised model, we experimented with different numbers of layers, hidden dimension sizes, learning rates, dropouts, and weight decays. To enable fair comparison against Tempo models, we searched the same space of hyper-parameters and selected the hyper-parameters that achieved the best average reward on the MGH validation set. Our final Supervised-Mirai model had 8 layers, each with 512 hidden units, followed by ReLU nonlinearities. It was trained for 30 epochs using a learning rate of $1e-3$, a dropout of 0.25 and a weight decay of 0.1 using the Adam optimizer. Our final Supervised-TCv8 model also had 8 layers, each with 512 hidden units, followed by ReLU nonlinearities. It was trained for 30 epochs using a learning rate of $1e-4$, a dropout of 0.25 and a weight decay of 0.1 using the Adam optimizer. The implementation of each risk policy, the training code as well as our hyper-parameter searches are available in our code release.

Statistical Analysis

To calculate confidence intervals while accounting for patients having multiple trajectories, we used a clustered bootstrap approach with 5000 samples. To assess significance in the difference between two metrics, we used a two-tailed t-test with a predefined p-value of 0.05 for significance.

List of Supplementary Materials

Table S1: Detailed demographics of MGH dataset.

Table S2: Detailed demographics of Emory test set.

Table S3: Detailed demographics of Karolinska test set. Because the Karolinska dataset was collected in a case-control design, it has a much higher cancer incidence than reported in the CSAW cohort [45]. To take this into account, we randomly resampled this dataset to produce a larger dataset with 1.9% cancer incidence.

Table S4: Detailed demographics of CGMH test set.

Table S5: Testing risk progression models on the MGH, Emory and Karolinska test sets.

Table S6: Results for all screening policies on subgroups of the MGH test set. For each policy, we report the average number of mammograms per year, the early detection benefit in months relative to historical screening (higher positive number means earlier), and the screening efficiency (higher positive number is better). We defined screening efficiency as the early detection benefit divided by the average number of mammograms per year. All metrics are followed by their 95% confidence interval.

Figure S1: Histogram of early detection benefit in months relative to historical screening for patients who developed cancer in the MGH (top left), Emory (top right), Karolinska (bottom left), and CGMH (bottom right) test sets

Figure S2: Histogram of screening recommendations for each screening policy. MGH (top left), Emory (top right), Karolinska (bottom left), CGMH (bottom right).

Figure S3. Our early detection metric assumed that a cancer could be caught up to 18 months before diagnosis. To test the robustness of our results to this assumption, we also evaluated our screening policies when changing this assumption to six months, 12 months and 24 months. For each policy, we report its screening efficiency, which is defined as its early detection benefit in months divided by the amount of mammograms it recommends per year. We use a * to denote the policy with the highest screening efficiency.

Figure S4: Dataset construction flow chart for the MGH dataset (top left), Emory (top right), Karolinska test set (bottom left), and CGMH test set (bottom right).

Figure S5: Estimated (circle) and observed (square) Mirai five-year risk for four random patients in the MGH test set. We estimated unobserved risk observations using a recurrent neural network, which was optimized to predict future risk assessments from past risk assessments on the MGH training set.

Data availability:

All datasets were used under license to the respective hospital system for the current study and are not publicly available.

Code availability:

All models and code used for training, evaluating and developing Tempo is publicly available at learningtocure.csail.mit.edu.

Author contributions: A.Y., and R.B. designed the research goals and aims. A.Y and R.B designed the model. A.Y., and R.B designed the evaluation methodology. A.Y. wrote the software. C.L, G.L, F.S, Y.W, S.S, T.K, I.B, J.G, and H.T curated the datasets. A.Y. and P.G.M performed the analysis. P.G.M. created the visualizations. All authors contributed to manuscript writing. R.B. supervised the project.

Funding:

The authors are grateful to be supported by grants from Susan G Komen, Breast Cancer Research Foundation, Quanta Computing, an Anonymous Foundation and the MIT J-Clinic. This work was also supported by Chang Gung Medical Foundation Grant SMRPG3K0051 and by Stockholm Läns Landsting HMT Grant 201708002.

Competing interests:

All authors have no competing interests relevant to this article to declare.

Tables

	MGH	Emory	Karolinska	CGMH
All exams	17119 (608)	22030 (723)	14362 (1768)	12280 (235)
Age				
<40	120 (2)	237 (7)	0 (0)	0 (0)
40-50	4710 (91)	4523 (114)	5921 (558)	3656 (74)
50-60	5271 (187)	6210 (162)	4200 (499)	5816 (109)
60-70	4728 (198)	7018 (231)	3903 (652)	2801 (52)
70-80	1997 (96)	3532 (195)	338 (59)	7 (0)
80<	313 (34)	510 (14)	0 (0)	0 (0)

Table 1: Demographics of Massachusetts General Hospital (MGH), Emory, Karolinska, and Chang Gung Memorial Hospital (CGMH) test sets. Each number is followed by the number of exams eventually followed by a cancer diagnosis.

Screening Policy	Risk model	Average number of Mammograms per Year	Earlier Detection in Months	Efficiency
MGH Test Set: 17,119 exams from 5,525 patients. 210 patients develop cancer.				
Annual	Age	1.0 (1.00, 1.00)	1.58 (0.54, 2.58)	1.58 (0.54, 2.58)
Biennial	Age	0.5 (0.50, 0.50)	-5.17 (-6.22, -4.13)	-10.34 (-12.44, -8.26)
USPSTF	Age	0.72 (0.71, 0.73)	-3.18 (-4.23, -2.22)	-4.42 (-5.83, -3.12)
Supervised	TCv8	1.66 (1.65, 1.69)	4.55 (3.51, 6.08)	2.74 (2.08, 3.70)
	Mirai	0.94 (0.92, 0.96)	0.75 (-0.55, 1.94)	0.80 (-0.58, 2.12)
Tempo	TCv8	0.96 (0.94, 0.97)	2.06 (1.14, 3.20)	2.16 (1.18, 3.40)
	Mirai	0.96 (0.94, 0.97)	4.10 (3.06, 4.96)	4.29 (3.17, 5.25)
Emory Test Set: 22,030 exams from 10,340 patients. 333 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	3.21 (2.37, 3.88)	3.21 (2.37, 3.88)
Biennial	Age	0.5 (0.5, 0.5)	-4.03 (-5.07, -3.22)	-8.07 (-10.14, -6.5)
USPSTF	Age	0.68 (0.67, 0.69)	-2.11 (-2.97, -1.40)	-3.12 (-4.36, -2.08)
Supervised	Mirai	1.16 (1.15, 1.18)	2.05 (0.48, 3.29)	1.76 (0.41, 2.86)
Tempo	Mirai	1.08 (1.07, 1.08)	6.39 (5.49, 6.99)	5.92 (5.06, 6.54)
Karolinska Test Set: 14,353 exams from 7,191 patients. 919 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	6.29 (5.76, 6.85)	6.29 (5.76, 6.85)
Biennial	Age	0.5 (0.5, 0.5)	-2.04 (-2.66, -1.41)	-4.07 (-5.32, -2.82)
USPSTF	Age	0.79 (0.79, 0.80)	1.02 (0.37, 1.63)	1.28 (0.46, 2.08)
Supervised	Mirai	0.60 (0.59, 0.61)	0.34 (-0.60, 1.24)	0.56 (-0.98, 2.11)
Tempo	Mirai	0.75 (0.74, 0.76)	7.23 (6.46, 7.97)	9.63 (8.53, 10.72)
CGMH Test Set: 12280 exams from 12280 patients. 235 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	11.00 (9.86, 12.28)	11.00 (9.86, 12.28)

Biennial	Age	0.5 (0.5, 0.5)	5.59 (4.11, 7.05)	11.18 (8.22, 14.09)
USPSTF	Age	0.78 (0.77, 0.78)	8.63 (7.15, 10.06)	11.10 (9.14, 13.01)
Supervised	Mirai	0.98 (0.97, 0.99)	8.02 (6.23, 10.21)	8.17 (6.30, 10.53)
Tempo	Mirai	0.88 (0.87, 0.89)	11.36 (10.21, 12.59)	12.92 (11.54, 14.41)

Table 2: Results for all screening policies on the MGH, Emory, Karolinska and CGMH test sets. For each policy, we report the average number of mammograms per year, the early detection benefit in months relative to historical screening (higher positive number means earlier), and the screening efficiency (higher positive number is better). We defined screening efficiency as the early detection benefit divided by the average number of mammograms per year. All metrics are followed by their 95% confidence interval.

Figures:

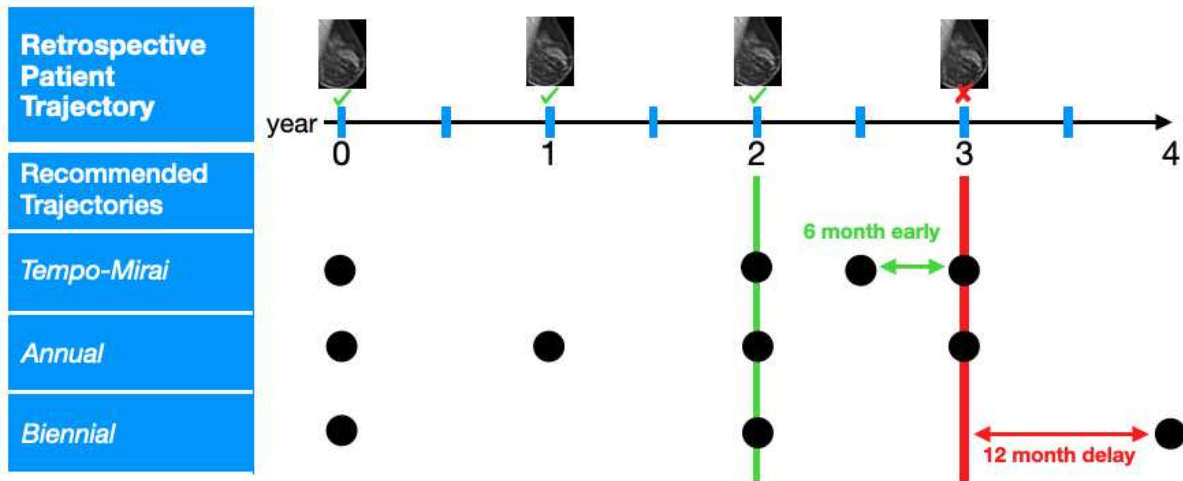


Figure 1: Retrospective patient trajectory from MGH test set compared to recommended trajectories by different guidelines. This patient was screened every year, from years zero to year three, and was diagnosed with breast cancer in year three. The red “x” and red line indicate the known time of cancer diagnosis. The green check marks indicate screening negative mammograms, and the green line indicate the last known negative time-point, i.e year two. For each recommended trajectory, we can compute the screening cost and early detection benefit relative to the historical screening. We measure the early detection benefit of a policy, by comparing its recommended screening dates to the last known negative date and the known cancer date. In our simulation, Tempo-Mirai, annual screening and biennial screening obtained an early detection benefit of 6.0 months, 0 months and -12.0 months respectively while recommending an average of 1.0, 1.0 and 0.5 mammograms per year for this patient.

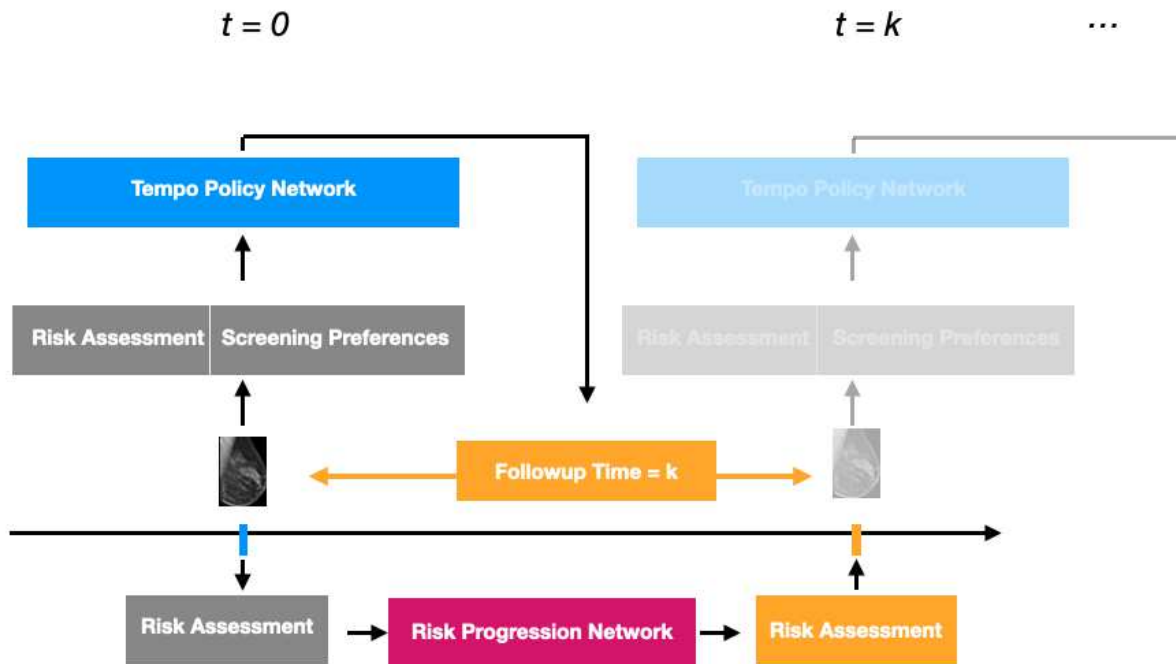


Figure 2: Overview of Tempo. Each screening policy, e.g Tempo-Mirai, takes as input a risk assessment and a weighting between rewards, and outputs a recommended followup time. We refer to the weighting between rewards as the screening preference. If a risk assessment is not available at the next recommended time-step, we estimate the risk assessment using our risk progression network.

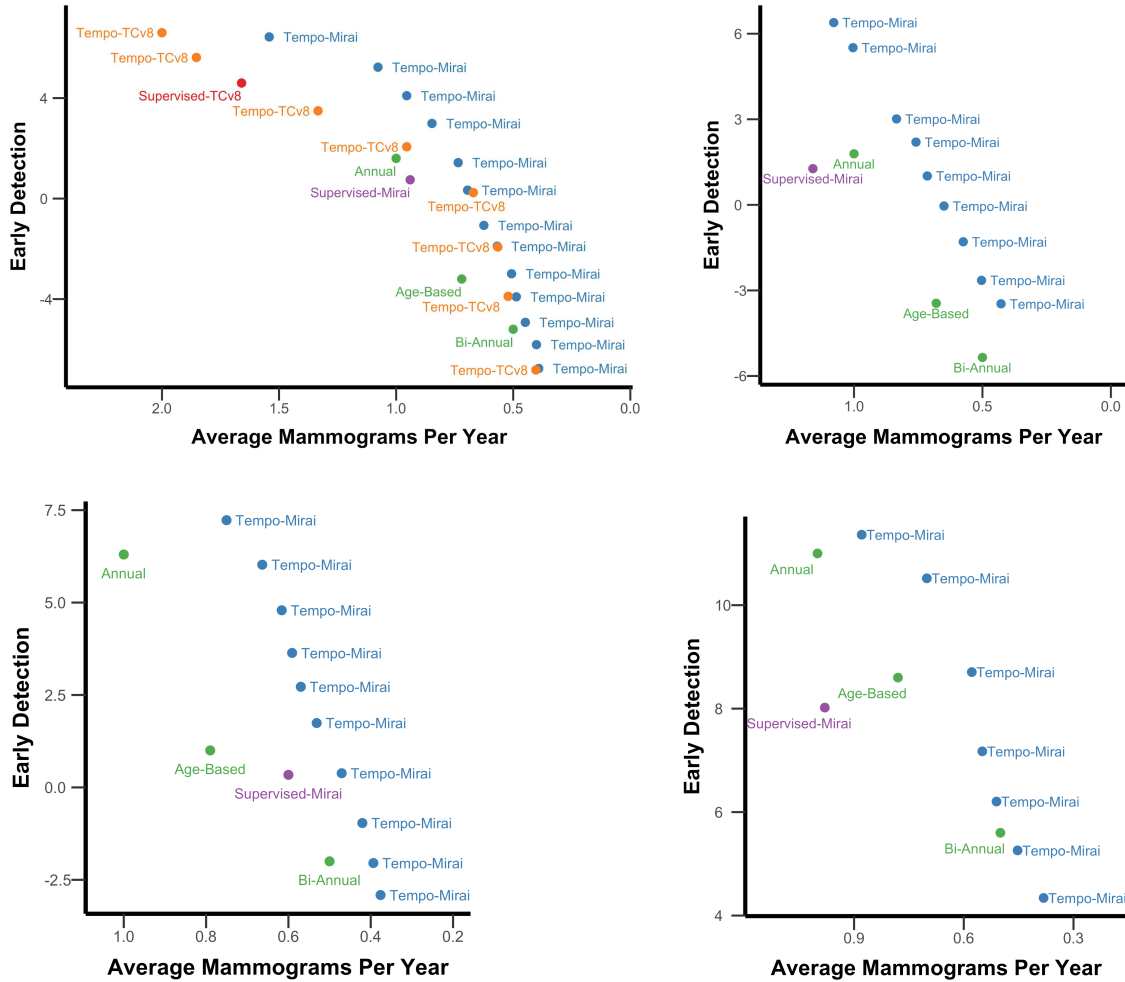


Figure 3: Early detection vs the number of mammograms per year at MGH (top left), Emory (top right), Karolinska (bottom left), CGMH (bottom right). Each point for a Tempo model (e.g Tempo-Mirai) corresponds to an alternative preference in the trade-off between early detection and screening frequency. Tempo policies (i.e Tempo-Mirai, Tempo-TCv8) are all trained using our reinforcement learning framework and Supervised policies (i.e Supervised-Mirai, Supervised-TCv8) are trained using a supervised learning baseline. Mirai and TCv8 policies refer to policies that leverage Mirai[12], and Tyrer-Cuzick version 8 [8] risk assessments respectively.

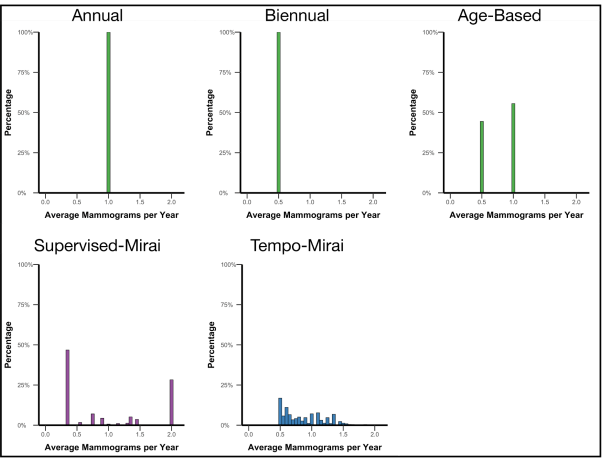
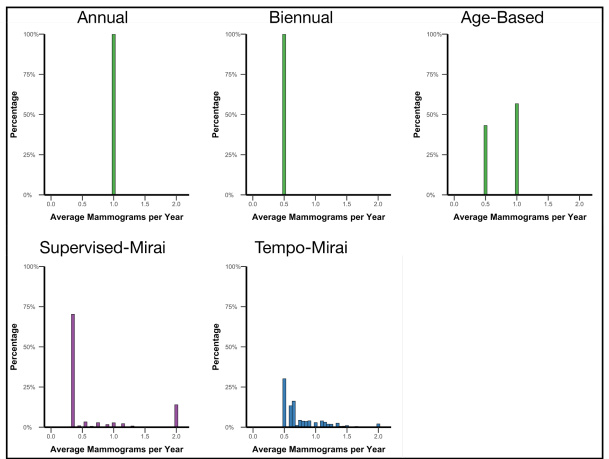
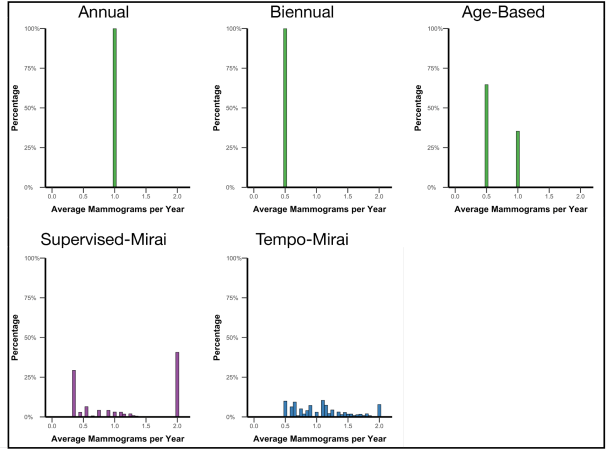
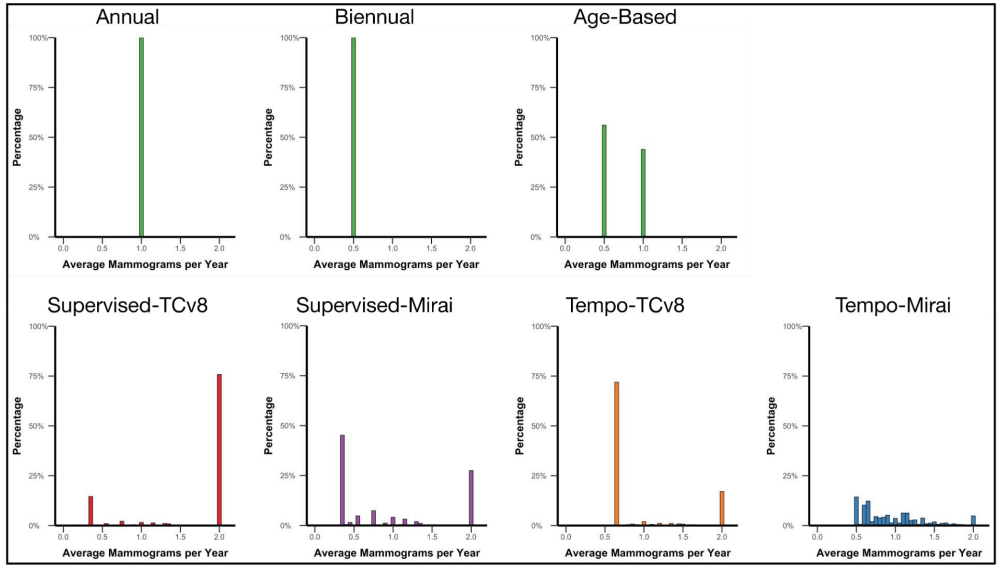


Figure 4: Histograms of screening frequency, i.e average number mammograms per year, as recommended by each screening policy for patients across the MGH (top), Emory (middle), Karolinska (bottom left) and CGMH (bottom right) test sets.

Supplementary Tables

	MGH Train Set		MGH Validation Set		MGH Test Set	
Characteristics	All	Cancer	All	Cancer	All	Cancer
All exams	137682 (100.0)	5202 (3.8)	16634 (100.0)	613 (3.7)	17119 (100.0)	608 (3.6)
Age						
<40	948 (0.7)	28 (3.0)	114 (0.7)	2 (1.8)	120 (0.7)	2 (1.7)
40-50	36971 (26.9)	1051 (2.8)	4483 (27.0)	151 (3.4)	4710 (27.5)	91 (1.9)
50-60	42425 (30.8)	1331 (3.1)	5153 (31.0)	155 (3.0)	5271 (30.8)	187 (3.5)
60-70	37715 (27.4)	1763 (4.7)	4585 (27.6)	181 (3.9)	4728 (27.6)	198 (4.2)
70-80	16663 (12.1)	798 (4.8)	1958 (11.8)	107 (5.5)	1977 (11.5)	96 (4.9)
80<	2960 (2.1)	231 (7.8)	341 (2.1)	17 (5.0)	313 (1.8)	34 (10.9)
Density						
Almost entirely fatty	12639 (9.2)	294 (2.3)	1499 (9.0)	47 (3.1)	1569 (9.2)	30 (1.9)
Scattered areas of fibroglandular tissue	65353 (47.5)	2496 (3.8)	8007 (48.1)	250 (3.1)	8112 (47.4)	293 (3.6)
Heterogeneously dense	52991 (38.5)	2171 (4.1)	6255 (37.6)	276 (4.4)	6633 (38.7)	265 (4.0)
Extremely dense	6623 (4.8)	239 (3.6)	867 (5.2)	39 (4.5)	797 (4.7)	20 (2.5)
Race						
White	112055 (81.4)	4495 (4.0)	13432 (80.8)	518 (3.9)	13932 (81.4)	534 (3.8)
African American	6585 (4.8)	204 (3.1)	792 (4.8)	27 (3.4)	807 (4.7)	26 (3.2)
Asian or Pacific Islander	6055 (4.4)	136 (2.2)	779 (4.7)	16 (2.1)	817 (4.8)	21 (2.6)

Hispanic	1542 (1.1)	52 (3.4)	181 (1.1)	5 (2.8)	145 (0.8)	4 (2.8)
Other Race	11445 (8.3)	315 (2.8)	1450 (8.7)	47 (3.2)	1418 (8.3)	23 (1.6)
Time to Next Exam						
< 1 year	1313 (1.0)	295 (22.5)	172 (1.0)	50 (29.1)	159 (0.9)	32 (20.1)
1-2 years	114503 (83.2)	4264 (3.7)	13791 (82.9)	486 (3.5)	14192 (82.9)	514 (3.6)
2-3 years	12377 (9.0)	429 (3.5)	1489 (9.0)	58 (3.9)	1536 (9.0)	41 (2.7)
>= 3 years	9489 (6.9)	214 (2.3)	1182 (7.1)	19 (1.6)	1232 (7.2)	21 (1.7)
Time to Cancer						
0-1 year	61 (0.0)	61 (100.0)	8 (0.0)	8 (100.0)	13 (0.1)	13 (100.0)
1-2 years	298 (0.2)	298 (100.0)	40 (0.2)	40 (100.0)	34 (0.2)	34 (100.0)
2-3 years	508 (0.4)	508 (100.0)	50 (0.3)	50 (100.0)	60 (0.4)	60 (100.0)
3-4 years	632 (0.5)	632 (100.0)	80 (0.5)	80 (100.0)	97 (0.6)	97 (100.0)
4-5 years	724 (0.5)	724 (100.0)	92 (0.6)	92 (100.0)	94 (0.5)	94 (100.0)
5-10 years	2979 (2.2)	2979 (100.0)	343 (2.1)	343 (100.0)	310 (1.8)	310 (100.0)

Table S1: Detailed demographics of MGH dataset.

	Emory Dataset	
Characteristics	All	Cancer
All exams	22,030 (100.0)	723 (3.3)
Age		
<40	237 (1.1)	7 (3.0)
40-50	4,523 (20.5)	114 (2.5)
50-60	6,210 (28.2)	162 (2.6)
60-70	7,018 (31.9)	231 (3.3)
70-80	3,532 (16.0)	195 (5.5)
80<	510 (2.3)	14 (2.7)
Race		
White	9,780 (44.4)	348 (3.6)
African American	10,436 (47.4)	343 (3.3)
Asian	994 (4.5)	15 (1.5)
Native Hawaiian or Other Pacific Islander	122 (0.6)	9 (7.4)
American Indian or Alaskan Native	21 (0.1)	NA
Multiple	47 (0.2)	NA
Time to Next Exam		
< 1 year	529 (2.4)	48 (9.1)
1-2 years	16,557 (75.2)	546 (3.3)
2-3 years	2,628 (11.9)	82 (3.1)
>= 3 years	2,316 (10.5)	47 (2.0)
Time to Cancer		
0-1 year	16 (0.1)	16 (100.0)
1-2 years	96 (0.4)	96 (100.0)

2-3 years	124 (0.6)	124 (100.0)
3-4 years	110 (0.5)	110 (100.0)
4-5 years	132 (0.6)	132 (100.0)
5-10 years	245 (1.1)	245 (100.0)

Table S2. Detailed demographics of Emory test set

Characteristics	Karolinska Dataset before resampling		Karolinska Dataset after resampling	
	All	Cancer	All	Cancer
All exams	14362 (100.0)	1768 (12.3)	93052 (100.0)	1768 (1.9)
Age				
40-50	5921 (41.2)	558 (9.4)	39433 (42.4)	558 (1.4)
50-60	4200 (29.2)	499 (11.9)	27514 (29.6)	499 (1.8)
60-70	3903 (27.2)	652 (16.7)	24010 (25.8)	652 (2.7)
70-80	338 (2.4)	59 (17.5)	2095 (2.3)	59 (2.8)
Time to Next Exam				
< 1 year	90 (0.6)	84 (93.3)	134 (0.1)	84 (62.7)
1-2 years	5421 (37.7)	618 (11.4)	35380 (38.0)	618 (1.7)
2-3 years	7087 (49.3)	912 (12.9)	45844 (49.3)	912 (2.0)
>= 3 years	1764 (12.3)	154 (8.7)	11694 (12.6)	154 (1.3)
Time to Cancer				
0-1 year	25 (0.2)	25 (100.0)	25 (0.0)	25 (100.0)
1-2 years	94 (0.7)	94 (100.0)	94 (0.1)	94 (100.0)
2-3 years	257 (1.8)	257 (100.0)	257 (0.3)	257 (100.0)
3-4 years	204 (1.4)	204 (100.0)	204 (0.2)	204 (100.0)
4-5 years	352 (2.5)	352 (100.0)	352 (0.4)	352 (100.0)
5-10 years	836 (5.8)	836 (100.0)	836 (0.9)	836 (100.0)

Table S3: Detailed demographics of Karolinska test set. Because the Karolinska dataset was collected in a case-control design, it has a much higher cancer incidence than reported in the CSAW cohort [46]. To take this into account, we randomly resampled this dataset to produce a larger dataset with 1.9% cancer incidence.

	CGMH Dataset	
Characteristics	All	Cancer
All exams	12280 (100.0)	235 (1.9)
Age		
40-50	3656 (29.8)	74 (2.0)
50-60	5816 (47.4)	109 (1.9)
60-70	2801 (22.8)	52 (1.9)
70-80	7 (0.1)	NA
Time to Next Exam		
< 1 year	NA	13 (100.0)
1-2 years	NA	31 (100.0)
2-3 years	NA	50 (100.0)
>= 3 years	NA	141 (1.2)
Time to Cancer		
0-1 year	11 (0.1)	11 (100.0)
1-2 years	24 (0.2)	24 (100.0)
2-3 years	42 (0.3)	42 (100.0)
3-4 years	26 (0.2)	26 (100.0)
4-5 years	36 (0.3)	36 (100.0)
5-6 years	96 (0.8)	96 (100.0)

Table S4: Detailed demographics of CGMH test set.

Risk Model	Progression Model	KL Divergence on Test Set (95% Confidence interval)
MGH Test Set: 17,119 exams from 5,525 patients. 210 patients develop cancer.		
Mirai	Static Risk	0.038 (0.036, 0.040)
	RNN	0.028 (0.026, 0.029)
Emory Test Set: 22,094 exams from 10,369 patients. 333 patients develop cancer.		
Mirai	Static Risk	0.035 (0.034, 0.036)
	RNN	0.029 (0.028, 0.030)
Karolinska Test Set: 14,353 exams from 7,191 patients. 919 patients develop cancer.		
Mirai	Static Risk	0.029 (0.027, 0.031)
	RNN	0.026 (0.025, 0.027)

Table S5: Testing risk progression models on the MGH, Emory and Karolinska test sets. Static Risk assumes that patient risk does not change, i.e risk assessments at future time steps will equal the last observed risk assessment. RNN is an auto-regressive recurrent neural network that was trained to predict future risk assessments from prior assessments on the MGH training set. For each model, we report the Kullback-Leibler (KL) divergence (lower is better), between the risk progressions model predicted risk and the observed risk. All metrics are followed by their 95% confidence interval.

Screening Policy	Risk model	Average Mammograms per Year	Early Detection in Months	Efficiency
Age: <=55. 8,038 exams from 3,016 patients. 84 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	1.16 (-0.34, 2.62)	1.16 (-0.34, 2.62)
Biennial	Age	0.5 (0.5, 0.5)	-5.50 (-7.11, -3.90)	-11.00 (-14.22, -7.80)
USPSTF	Age	0.97 (0.96, 0.97)	0.848 (-0.76, 2.29)	0.88 (-0.78, 2.37)
Supervised	TCv8	1.45 (1.42, 1.48)	2.17 (0.39, 4.30)	1.50 (0.26, 3.03)
	Mirai	0.769 (0.75, 0.80)	-2.17 (-4.83, 1.01)	-2.82 (-6.08, 1.35)
Tempo	TCv8	0.96 (0.94, 0.99)	1.76 (0.21, 3.77)	1.83 (0.21, 4.01)
	Mirai	0.86 (0.84, 0.87)	2.92 (1.25, 4.63)	3.41 (1.44, 5.53)
Age: > 55. 9081 exams from 2959 patients. 148 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	1.77 (0.68, 2.83)	1.77 (0.68, 2.83)
Biennial	Age	0.5 (0.5, 0.5)	-5.02 (-6.23, -3.89)	-10.04 (-12.47, -7.78)
USPSTF	Age	0.5 (0.5, 0.5)	-5.02 (-6.23, -3.88)	-10.04 (-12.47, -7.78)
Supervised	TCv8	1.85 (1.83, 1.87)	5.64 (4.28, 7.44)	3.06 (2.30, 4.07)
	Mirai	1.08 (1.06, 1.12)	2.09 (0.38, 3.93)	1.92 (0.34, 3.72)
Tempo	TCv8	1.72 (1.70, 1.74)	4.84 (3.20, 6.80)	2.82 (1.84, 4.02)
	Mirai	1.04 (1.03, 1.06)	4.63 (3.49, 6.03)	4.45 (3.29, 5.88)
Density: Non-dense. 9681 exams from 3370 patients. 120 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	1.54 (0.21, 2.65)	1.54 (0.21, 2.65)
Biennial	Age	0.5 (0.5, 0.5)	-5.26 (-6.64, -3.97)	-10.51 (-13.28, -7.95)
USPSTF	Age	0.67 (0.66, 0.67)	-4.05 (-5.43, -2.80)	-6.08 (-8.05, -4.27)
Supervised	TCv8	1.62 (1.6, 1.65)	4.29 (2.26, 6.31)	2.65 (1.38, 3.95)
	Mirai	0.87 (0.84, 0.90)	0.07 (-1.86, 2.03)	0.09 (-2.07, 2.42)

Tempo	TCv8	0.96 (0.93, 0.97)	1.56 (0.52, 2.48)	1.64 (0.54, 2.67)
	Mirai	0.94 (0.91, 0.95)	3.83 (2.70, 5.00)	4.10 (2.85, 5.48)
Density: Dense. 7430 exams from 2839 patients. 116 patients develop cancer.				
Annual	Age	1.0 (1.0, 1.0)	1.621 (0.28, 2.95)	1.621 (0.28, 2.95)
Biennial	Age	0.5 (0.5, 0.5)	-5.07 (-6.50, -3.61)	-10.15 (-12.98, -7.21)
USPSTF	Age	0.79 (0.78, 0.80)	-2.19 (-3.80, -0.69)	-2.78 (-4.77, -0.89)
Supervised	TCv8	1.72 (1.70, 1.74)	4.84 (3.20, 6.80)	2.82 (1.84, 4.02)
	Mirai	1.02 (0.98, 1.05)	1.52 (-0.17, 4.00)	1.49 (-0.15, 4.08)
Tempo	TCv8	0.96 (0.93, 0.98)	2.63 (1.34, 4.54)	2.77 (1.37, 4.89)
	Mirai	0.98 (0.96, 1.0)	4.4 (2.86, 6.08)	4.49 (2.86, 6.35)

Table S6: Results for all screening policies on subgroups of the MGH test set. For each policy, we report the average number of mammograms per year, the early detection benefit in months relative to historical screening (higher positive number means earlier), and the screening efficiency (higher positive number is better). We defined screening efficiency as the early detection benefit divided by the average number of mammograms per year. All metrics are followed by their 95% confidence interval.

Supplementary Figures

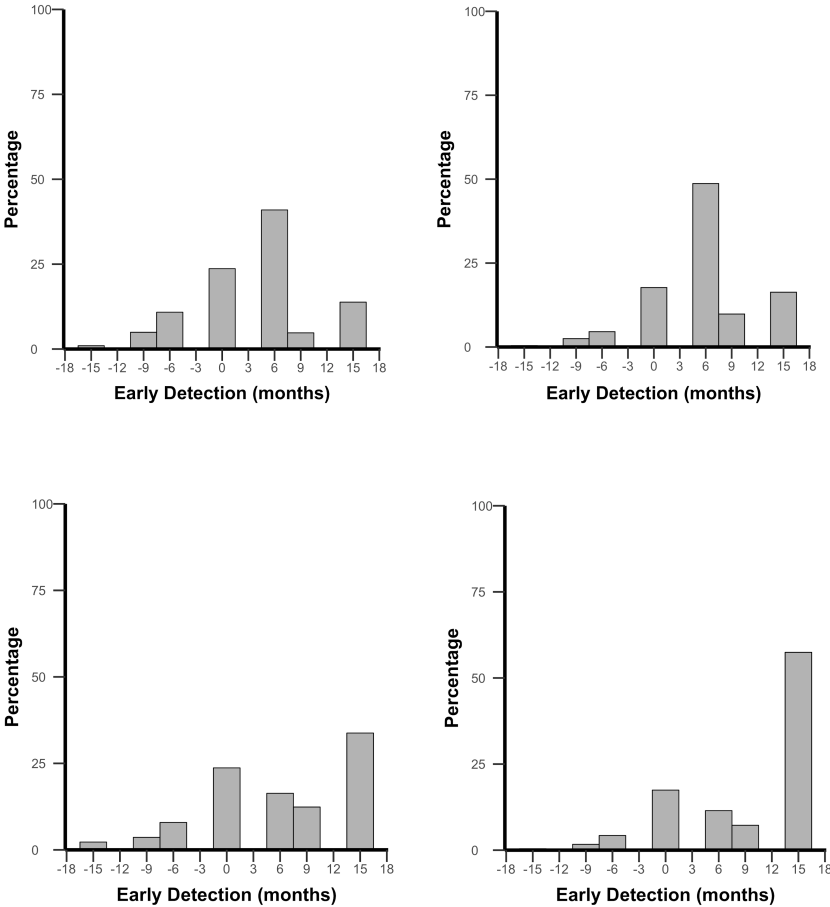


Figure S1. Histogram of early detection benefit in months relative to historical screening for patients who developed cancer in the MGH (top left), Emory (top right), Karolinska (bottom left), and CGMH (bottom right) test sets.

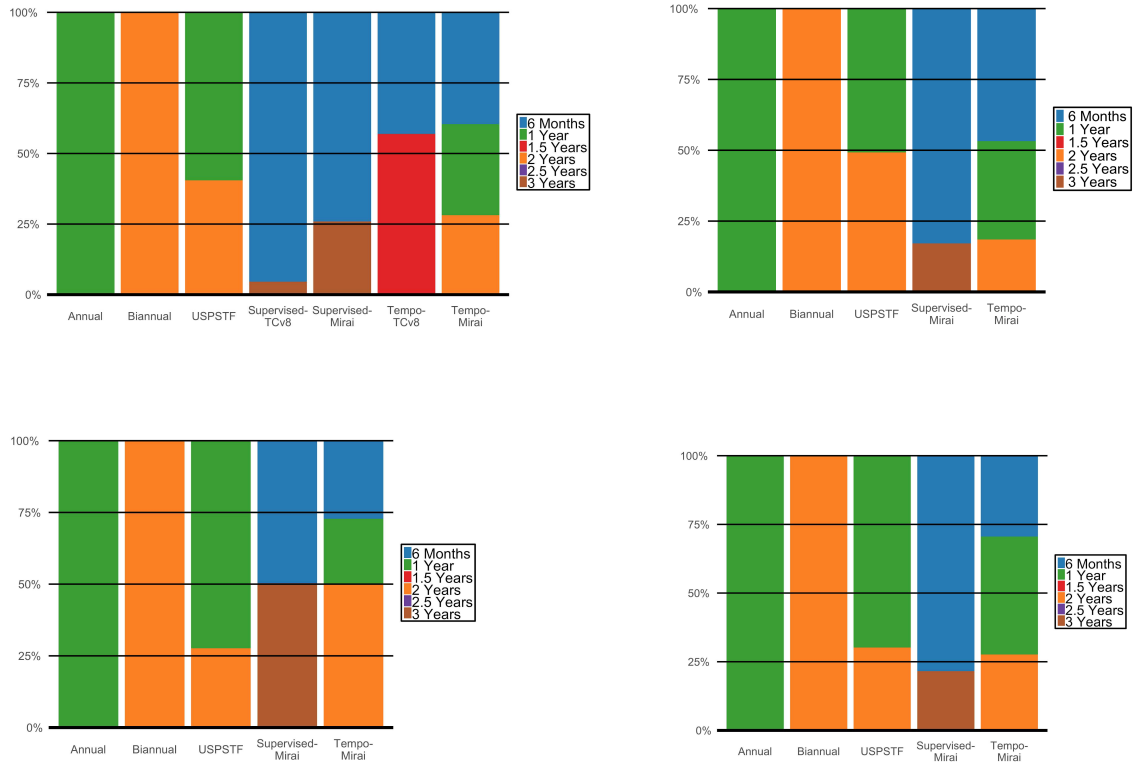


Figure S2: Histogram of screening recommendations for each screening policy. MGH (top left), Emory (top right), Karolinska (bottom left), CGMH (bottom right).

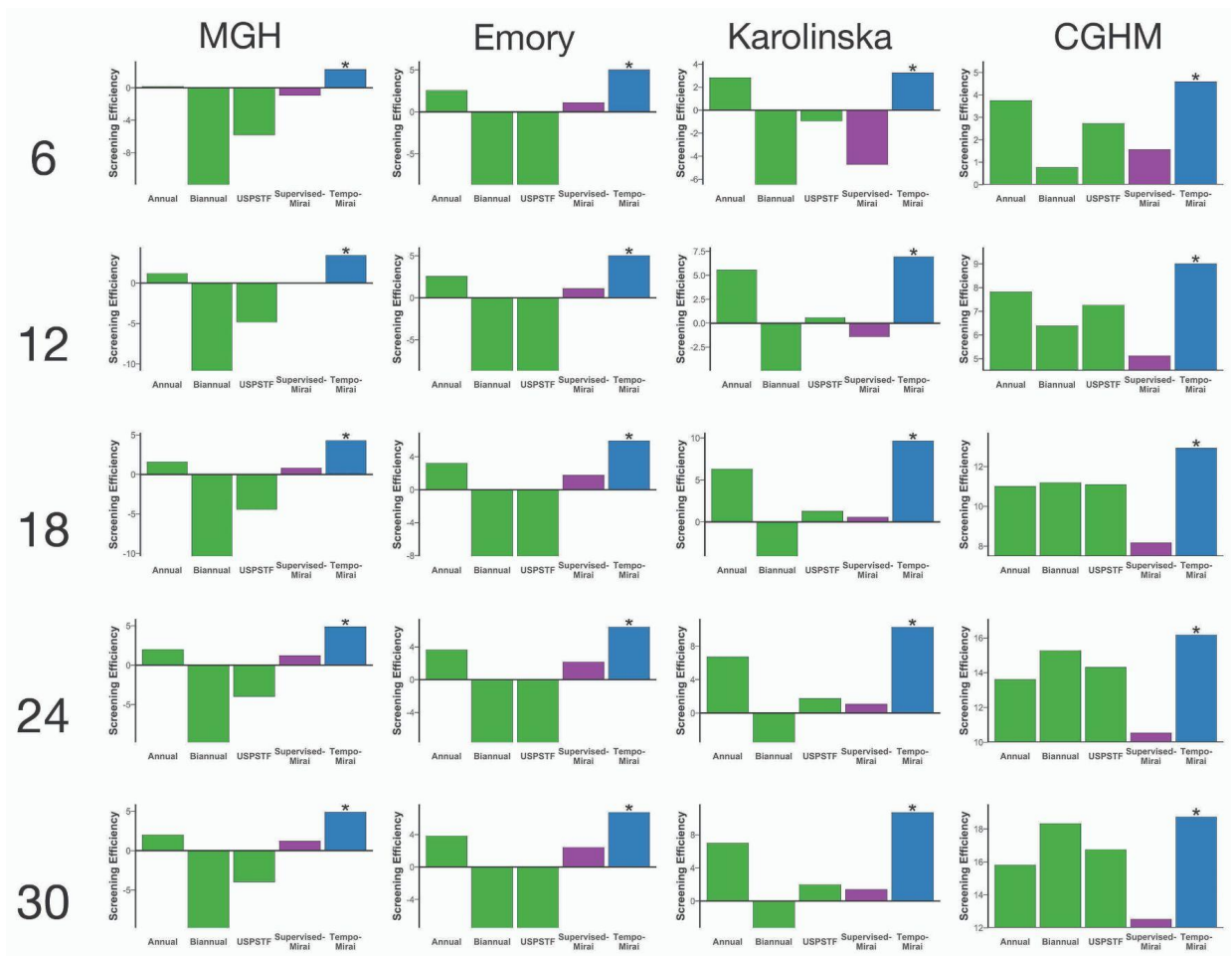


Figure S3. Our early detection metric assumed that a cancer could be caught up to 18 months before diagnosis. To test the robustness of our results to this assumption, we also evaluated our screening policies when changing this assumption to six months, 12 months and 24 months. For each policy, we report its screening efficiency, which is defined as its early detection benefit in months divided by the amount of mammograms it recommends per year. We use a * to denote the policy with the highest screening efficiency.

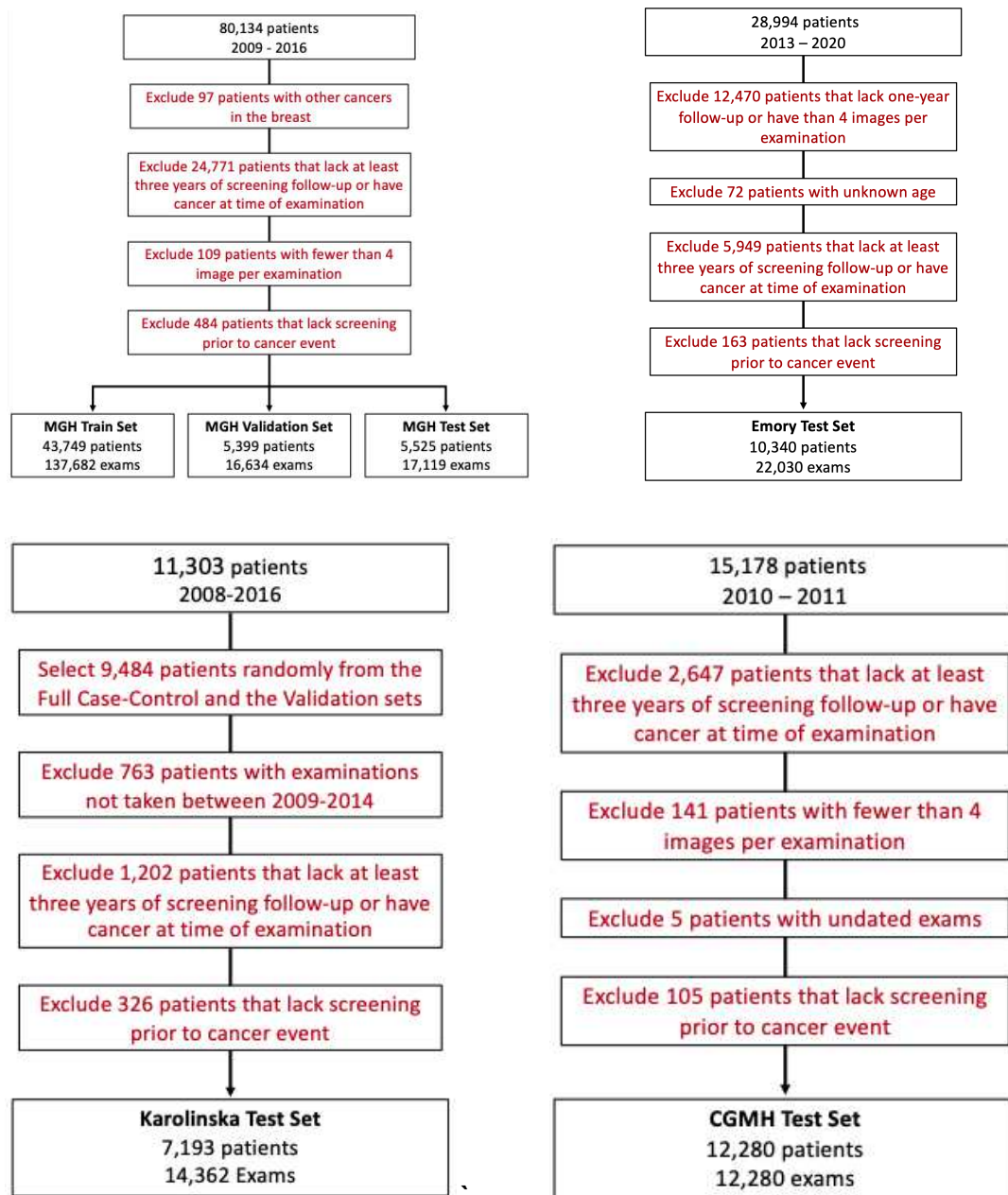


Figure S4: Dataset construction flow chart for the MGH dataset (top left), Emory (top right), Karolinska test set (bottom left), and CGMH test set (bottom right).

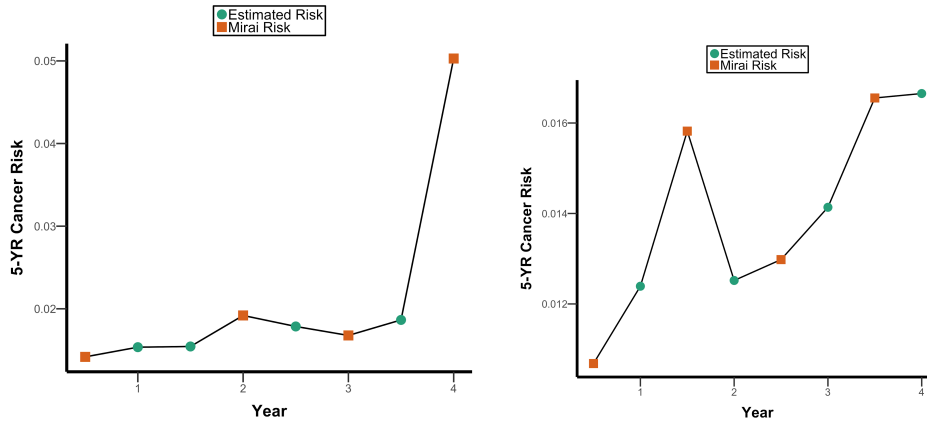


Figure S5: Estimated (circle) and observed (square) Mirai five-year risk for two random patients in the MGH test set. We estimated unobserved risk observations using a recurrent neural network, which was optimized to predict future risk assessments from past risk assessments on the MGH training set.

References

- [1] Robert A Smith, Kimberly S Andrews, Durado Brooks, Stacey A Fedewa, Deana Manassaram-Baptiste, Debbie Saslow, and Richard C Wender. Cancer screening in the United States, 2019: A review of current american cancer society guidelines and current issues in cancer screening. *CA: a cancer journal for clinicians*, 69(3):184–210, 2019.
- [2] Karen J Wernli, Nora B Henrikson, Caitlin C Morrison, Matthew Nguyen, Gaia Pocobelli, and Paula R Blasi. Screening for skin cancer in adults: updated evidence report and systematic review for the US Preventive Services Task Force. *Jama*, 316(4):436–447, 2016.
- [3] Cathy Coleman. Early detection and screening for breast cancer. In *Seminars in oncology nursing*, volume 33, pages 141–155. Elsevier, 2017.
- [4] Susan J Curry, Alex H Krist, Douglas K Owens, Michael J Barry, Aaron B Caughey, Karina W Davidson, Chyke A Doubeni, John W Epling, Alex R Kemper, Martha Kubik, et al. Screening for cervical cancer: Us preventive services task force recommendation statement. *Jama*, 320(7):674–686, 2018.
- [5] Rodrigo A Gier, Krista A Budinich, Niklaus H Evitt, Zhendong Cao, Elizabeth S Freilich, Qingzhou Chen, Jun Qi, Yemin Lan, Rahul M Kohli, and Junwei Shi. High-performance crispr-cas12a genome editing for combinatorial genetic screening. *Nature communications*, 11(1):1–9, 2020.
- [6] Adam Yala, Constance Lehman, Tal Schuster, Tally Portnoi, and Regina Barzilay. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*, 292(1):60–66, 2019.
- [7] Mitchell H Gail, Louise A Brinton, David P Byar, Donald K Corle, Sylvan B Green, Catherine Schairer, and John J Mulvihill. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *JNCI: Journal of the National Cancer Institute*, 81(24):1879–1886, 1989.
- [8] Jonathan Tyrer, Stephen W Duffy, and Jack Cuzick. A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in medicine*, 23(7):1111–1130, 2004.
- [9] Kirsten Bibbins-Domingo, David C Grossman, Susan J Curry, Karina W Davidson, John W Epling, Francisco AR Garcia, Matthew W Gillman, Diane M Harper, Alex R Kemper, Alex H Krist, et al. Screening for colorectal cancer: Us preventive services task force recommendation statement. *Jama*, 315(23):2564–2575, 2016.
- [10] Virginia A Moyer. Screening for lung cancer: Us preventive services task force recommendation statement. *Annals of internal medicine*, 160(5):330–338, 2014.

- [11] Albert L Siu. Screening for breast cancer: Us preventive services task force recommendation statement. *Annals of internal medicine*, 164(4):279–296, 2016.
- [12] Adam Yala, Peter G. Mikhael, Fredrik Strand, Gigin Lin, Kevin Smith, Yung-Liang Wan, Leslie Lamb, Kevin Hughes, Constance Lehman, and Regina Barzilay. "Toward robust mammography-based models for breast cancer risk." *Science Translational Medicine* 13, no. 578 (2021).
- [13] Karin Dembrower, Yue Liu, Hossein Azizpour, Martin Eklund, Kevin Smith, Peter Lindholm, and Fredrik Strand. Comparison of a deep learning risk score and standard mammographic density score for breast cancer risk prediction. *Radiology*, 294(2):265–272, 2020.
- [14] Michael T Lu, Vineet K Raghu, Thomas Mayrhofer, Hugo JWL Aerts, and Udo Hoffmann. Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening computed tomography: development and validation of a prediction model. *Annals of Internal Medicine*, 2020.
- [15] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research*, 48:67–113, 2013.
- [16] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- [17] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.
- [18] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.
- [19] Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, et al. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*, 2018.
- [20] Avi Singh, Larry Yang, Kristian Hartikainen, Chelsea Finn, and Sergey Levine. End-to-end robotic reinforcement learning without reward engineering. *arXiv preprint arXiv:1904.07854*, 2019.
- [21] Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, and Konstantinos Bousmalis. Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks. In

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 12627–12637, 2019.

[22] Gerrit Schoettler, Ashvin Nair, Jianlan Luo, Shikhar Bahl, Juan Aparicio Ojea, Eugen Solowjow, and Sergey Levine. Deep reinforcement learning for industrial insertion tasks with visual inputs and natural reward signals. In ICML 2019 Workshop, 2019.

[23] Dmitry Kalashnikov, Alex Irpan, Peter Pastor, Julian Ibarz, Alexander Herzog, Eric Jang, Deirdre Quillen, Ethan Holly, Mrinal Kalakrishnan, Vincent Vanhoucke, et al. Scalable deep reinforcement learning for vision-based robotic manipulation. In Conference on Robot Learning, pages 651–673, 2018.

[24] Kevin Shen Hoong Ong, Dusit Niyato, and Chau Yuen. Predictive maintenance for edge-based sensor networks: A deep reinforcement learning approach. arXiv e-prints, pages arXiv–2007, 2020.

[25] Nassima Aissani, Bouziane Beldjilali, and Damien Trentesaux. Dynamic scheduling of maintenance tasks in the petroleum industry: A reinforcement approach. Engineering Applications of Artificial Intelligence, 22(7):1089–1103, 2009.

[26] Roberto Rocchetta, L Bellani, M Compare, E Zio, and E Patelli. A reinforcement learning framework for optimal operation and maintenance of power grids. Applied energy, 241:291–301, 2019.

[27] Runzhe Yang, Xingyuan Sun, and Karthik Narasimhan. A generalized algorithm for multi-objective reinforcement learning and policy adaptation. In Advances in Neural Information Processing Systems, pages 14636–14647, 2019.

[28] Kevin C Oeffinger, Elizabeth TH Fontham, Ruth Etzioni, Abbe Herzig, James S Michaelson, Ya-Chen Tina Shih, Louise C Walter, Timothy R Church, Christopher R Flowers, Samuel J LaMonte, et al. Breast cancer screening for women at average risk: 2015 guideline update from the american cancer society. Jama, 314(15):1599–1614, 2015.

[29] Debra L Monticciolo, Mary S Newell, Linda Moy, Bethany Niell, Barbara Monsees, and Edward A Sickles. Breast cancer screening in women at higher-than-average risk: recommendations from the ACR. Journal of the American College of Radiology, 15(3):408–414, 2018.

[30] Yiwey Shieh, Martin Eklund, Lisa Madlensky, Sarah D Sawyer, Carlie K Thompson, Allison Stover Fiscalini, Elad Ziv, Laura J van’t Veer, Laura J Esserman, Jeffrey A Tice, et al. Breast cancer screening in the precision medicine era: risk-based screening in a population-based trial. Journal of the National Cancer Institute, 109(5):djw290, 2017.

[31] Douglas K Owens, Karina W Davidson, Alex H Krist, Michael J Barry, Michael Cabana, Aaron B Caughey, Chyke A Doubeni, John W Epling, Martha Kubik, C Seth Landefeld, et al.

Medication use to reduce risk of breast cancer: Us preventive services task force recommendation statement. *Jama*, 322(9):857–867, 2019.

[32] Kala Visvanathan, Rowan T Chlebowski, Patricia Hurley, Nananda F Col, Mary Ropka, Deborah Collyar, Monica Morrow, Carolyn Runow-icz, Kathleen I Pritchard, Karen Hagerty, et al. American society of clinical oncology clinical practice guideline update on the use of pharmacologic interventions including tamoxifen, raloxifene, and aromatase inhibition for breast cancer risk reduction. *Journal of clinical oncology*, 27(19):3235, 2009.

[33] Marije F Bakker, Stephanie V de Lange, Ruud M Pijnappel, Ritse M Mann, Petra HM Peeters, Evelyn M Monninkhof, Marleen J Emaus, Claudette E Loo, Robertus HC Bisschops, Marc BI Lobbes, et al. Supplemental MRI screening for women with extremely dense breast tissue. *New England Journal of Medicine*, 381(22):2091–2102, 2019.

[34] ClinicalTrials.gov. National Library of Medicine (US). International randomized study comparing personalized, risk-stratified to standard breast cancer screening in women aged 40-70., 2019, Jul 18 -. Identifier NCT03672331. Retrieved from <http://clinicaltrials.gov/ct/show/NCT03672331>.

[35] M Le Boulch, A Bekhouche, E Kermarrec, A Milon, C Abdel Wahab, S Zilberman, N Chabbert-Buffet, and I Thomassin-Naggara. Comparison of breast density assessment between human eye and automated software on digital and synthetic mammography: Impact on breast cancer risk. *Diagnostic and Interventional Imaging*, 2020.

[36] Lisa M Maillart, Julie Simmons Ivy, Scott Ransom, and Kathleen Diehl. Assessing dynamic breast cancer screening policies. *Operations Research*, 56(6):1411–1427, 2008.

[37] Turgay Ayer, Oguzhan Alagoz, and Natasha K Stout. OR forum - a POMDP approach to personalize mammography screening decisions. *Operations Research*, 60(5):1019–1034, 2012.

[38] Wang, F., Zhang, S., & Henderson, L. M. (2018). Adaptive decision-making of breast cancer mammography screening: A heuristic-based regression model. *Omega*, 76, 70-84.

[39] Mandelblatt, J. S., Stout, N. K., Schechter, C. B., Van Den Broek, J. J., Miglioretti, D. L., Krapcho, M., ... & Van Ravesteyn, N. T. (2016). Collaborative modeling of the benefits and harms associated with different US breast cancer screening strategies. *Annals of internal medicine*, 164(4), 215-225.

[40] Trentham-Dietz, A., Kerlikowske, K., Stout, N. K., Miglioretti, D. L., Schechter, C. B., Ergun, M. A., ... & Near, A. M. (2016). Tailoring breast cancer screening intervals by breast density and risk for women aged 50 years or older: collaborative modeling of screening outcomes. *Annals of internal medicine*, 165(10), 700-712.

- [41] John T Schousboe, Karla Kerlikowske, Andrew Loh, and Steven R Cummings. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Annals of internal medicine*, 155(1):10–20, 2011.
- [42] Kartik Ahuja, William Zame, and Mihaela van der Schaar. Dpscreen: Dynamic personalized screening. In *Advances in Neural Information Processing Systems*, pages 1321–1332, 2017.
- [43] Maartje van Seijen, Esther H Lips, Alastair M Thompson, Serena Nik-Zainal, Andrew Futreal, E Shelley Hwang, Ellen Verschuur, Joanna Lane, Jos Jonkers, Daniel W Rea, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. *British journal of cancer*, page 1, 2019.
- [44] Yinlam Chow and Mohammad Ghavamzadeh. Algorithms for cvar optimization in mdps. In *Advances in neural information processing systems*, pages 3509–3517, 2014.
- [45] Yinlam Chow, Aviv Tamar, Shie Mannor, and Marco Pavone. Risk-sensitive and robust decision-making: a cvar optimization approach. In *Advances in Neural Information Processing Systems*, pages 1522–1530, 2015.
- [46] Karin Dembrower, Peter Lindholm, and Fredrik Strand. A multi-million mammography image dataset and population-based screening cohort for the training and evaluation of deep neural networks—the cohort of screen-aged women (CSAW). *Journal of digital imaging*, pages 1–6, 2019.
- [47] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.
- [48] Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, and Wojciech Zaremba. Hindsight experience replay. In *Advances in neural information processing systems*, pages 5048–5058, 2017.