Title
Optimizing cancer screening with POMDPs

Permalink
https://escholarship.org/uc/item/0cz63433

Author
Petousis, Panayiotis

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Optimizing cancer screening with POMDPs

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Bioengineering

by

Panayiotis Petousis

2019

ABSTRACT OF THE DISSERTATION

Optimizing cancer screening with POMDPs

by

Panayiotis Petousis

Doctor of Philosophy in Bioengineering

University of California, Los Angeles, 2018

Professor Alex Anh-Tuan Bui, Co-Chair,

Professor Denise R. Aberle, Co-Chair

Current clinical decision-making relies heavily both upon the experience of a physician and the recommendations of evidence-based practice guidelines, the latter often informed by population-level policies. Yet with the heightened complexity of patient care given newer types of data and longitudinal observations (e.g., from the electronic health record, EHR), as well as the goal of more individually-tailored healthcare, medical decision-making is increasingly complicated. This issue is particularly true in cancer with emergent techniques for early detection and personalized treatment. This research establishes an informatics-based framework to inform optimal cancer screening through sequential decision-making methods. This dissertation develops tools to formulate a partially observable Markov decision process (POMDP) model, enabling each component to be learned from a dataset: dynamic Bayesian networks (DBNs) are embedded in the POMDP learning process to estimate transition and observations probabilities; inverse reinforcement learning is used to learn a reward function from experts' prior decisions; and risk prediction models are employed to compute individualized initial beliefs about disease state. The result is a comprehensive approach to implementing sequential decision making agents. These methods are validated using large datasets from lung and breast cancer screening efforts, demonstrating the potential to help

tailor and improve early cancer prediction while reducing false positive tests.

The dissertation of Panayiotis Petousis is approved.

Ricky Kiyotaka Taira

William Hsu

Denise R. Aberle, Committee Co-Chair

Alex Anh-Tuan Bui, Committee Co-Chair

University of California, Los Angeles

2019

As you set out for Ithaka hope the voyage is a long one, full of adventure, full of discovery. Laistrygonians and Cyclops, angry Poseidon − don't be afraid of them.

—Constantine P. Cavafy

And if you find her poor, Ithaca has not deceived you. Wise as you have become, with so much experience, you must already have understood what these Ithacas mean.

—Constantine P. Cavafy

The measure of who we are is what we do with what we have.

—Vince Lombardi

It always seems impossible until its done.

—nelson mandela

Contents

List of Figures

Panayiotis Petousis

Los Angeles, CA

6 January 2019

CONTRIBUTION OF AUTHORS

Portions of this dissertation have been published in my first-author papers, which are listed below under PUBLICATIONS. Chapter 3 is based on [Petousis, P., Han, S. X., Aberle, D., & Bui, A. A. (2016). Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network. Artificial intelligence in medicine, 72, 42-55.], chapter 4 is based on [Petousis, P., Han, S. X., Hsu, W., & Bui, A. A. (2018, January). Generating Reward Functions using IRL Towards Individualized Cancer Screening. In Joint Workshop on Artificial Intelligence for Health (AIH) in conjunction with ECAI/IJCAI, AAMAS, ICML.], chapter 5 is based on [Petousis, P., Winter, A., Speier, W., Aberle, D., Hsu, W., & Bui, A. A. (2019) Using Sequential Decision Making to Improve Lung Cancer Screening Performance. Nature npj Digital Medicine. In submission.], and chapter 6 is based on [Petousis, P., Naeim, A., Mosleh, A., & Hsu, W. (2018, January). Evaluating the Impact of Uncertainty on Risk Prediction: Towards More Robust Prediction Models. In AMIA... Annual Symposium proceedings.].

<center>VITA</center>

$2007 - 2009$	National Guard Cyprus
$2009 - 2013$	Bachelor and Master of Engineering in Biomedical Engineering with Electrical Engineering streaming, Imperial College London
$2013 - 2018$	Graduate Student Researcher in Bioengineering, UCLA
2019(expected)	Doctor of Philosophy in Bioengineering (Medical Informatics), UCLA

<center>PUBLICATIONS</center>

Shen, S., Han, S. X., **Petousis, P.**, Meng, F., Hsu, W., & Bui, A. A. (2015). A Continuous Markov Model Approach Using Individual Patient Data to Estimate Mean Sojourn Time of Lung Cancer. In AMIA.

Petousis, P., Han, S. X., Aberle, D., & Bui, A. A. (2016). Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network. Artificial intelligence in medicine, 72, 42-55.

Shen, S., Han, S. X., **Petousis, P.**, Weiss, R. E., Meng, F., Bui, A. A., & Hsu, W. (2017). A Bayesian model for estimating multi-state disease progression. Computers in biology and medicine, 81, 111-120.

Petousis, P., Han, S. X., Hsu, W., & Bui, A. A. (2018, January). Generating Reward Functions using IRL Towards Individualized Cancer Screening. In Joint Workshop on Artificial Intelligence for Health (AIH) in conjunction with ECAI/IJCAI, AAMAS, ICML.

Petousis, P., Naeim, A., Mosleh, A., & Hsu, W. (2018, January). Evaluating the Impact of Uncertainty on Risk Prediction: Towards More Robust Prediction Models. In AMIA... Annual Symposium proceedings.

Petousis, P., Winter, A., Speier, W., Aberle, D., Hsu, W., & Bui, A. A. (2019) Using Sequential Decision Making to Improve Lung Cancer Screening Performance. Nature npj Digital Medicine. In submission.

Chapter 1

Introduction

1.1 Overview

Chronic diseases, such as cancer, are the leading cause of death and healthcare costs worldwide. Many of these conditions are preventable and/or treatable if detected sufficiently early. As such, millions of people undergo screening annually for disease prevention and surveillance based on published clinical guidelines. During screening, physicians use patients' past and most current observations to determine a subsequent action (e.g., further diagnostic testing, increased monitoring, following regular screening schedules, etc.) that optimizes identification of health problems while balancing other (pragmatic) concerns (e.g., quality of life, resource utilization, cost). Cancer screening in particular is designed to identify individuals at high-risk for developing the disease, providing recommendations around the frequency of observation to maximize early detection and thereby improving survival outcomes. Decision-making tools that aid physicians with assessing these observations in the context of a patient's personal circumstances are critical to selecting an appropriate management strategy for cancer screening. Yet the design of such tools is challenging for any number of reasons, including the inherent uncertainty around clinical observations and the potential for sub-optimal selection of next steps in a patient's care. Imaging-based lung

cancer screening directly reflects this difficulty, where radiologists must judge the likelihood of an indeterminate pulmonary nodule (IPN) being an indolent cancer, with evolution from benign to malignant status sometimes happening quickly: determining if and when a patient should undergo a diagnostic intervention or continue with watchful waiting is difficult, even for experts. This dissertation focuses on improving low-dose computed tomography (LDCT) lung cancer screening management decisions and diagnosis through novel machine learning and partially-observable Markov decision process (POMDP) models: 1) to individualize the lung cancer screening process; 2) to reduce the false positive rate associated with lung cancer; and 3) to improve the early prediction of lung cancer.

1.2   Background and Motivation

In the United States, approximately 222,500 new cases of lung cancer will occur in 2018, 13.5% of all newly diagnosed cancers [3], and it is estimated to be responsible for 25% of all cancer-related mortality – over 154,000 deaths. Although the five-year survival rate for this disease improves when it is discovered in its nascent stages [4], only 15% of all lung cancers are detected early as symptoms often do not appear until the disease has advanced to a late or terminal stage. In the past decade, key studies demonstrated the effectiveness of imaging in early detection and reduction of lung cancer mortality: [5–7] found that annual computed tomography (CT) screening of Stage 1 cancers prevents upwards of 80% of cancer-related deaths; and the landmark National Lung Screening Trial (NLST) showed that screening with LDCT results in 20% relative mortality reduction from lung cancer as compared to screening with x-ray [5]. Motivated by these findings, national recommendations now mandate LDCT screening for high-risk populations and imaging-based lung cancer screening programs are increasingly common. However, the false positive rate for LDCT screening is disproportionately high [8]. Indeed, in the NLST the overall positive screen rate with LDCT was 24%, yet the positive predictive value of a positive screen was less than 4% [8]. Of the total

number of lung nodules diagnosed in the NLST, only 3-6% were found to be malignant, depending on nodule size. Unfortunately, LDCT detects many benign nodules and non-cancer related pathologies (e.g., inflammation, emphysema, other lesions), resulting in many false positives and the need for further diagnostic evaluation to confirm findings. The negative consequences of overdiagnosis are significant, with unnecessary diagnostic procedures (e.g., biopsies) and undue stress on patients [9]. New strategies facilitating early, accurate, and precise detection of lung cancers within imaging-based screening programs are imperative.

A framework optimizing early detection while reducing false positive rates would be ideal and can be used to support more individually-tailored screening recommendations. Two questions thus arise in lung cancer screening : 1) how to individualize and optimize screening; and 2) what is the individualized benefit of a cancer screening patient over time and how does that change based on a patient's initial risk and future clinical observations? Indeed, tailoring cancer screening is challenging, given its dynamic nature: selecting the best next action must account for a patient's initial and evolving risk factors, future imaging observations, and changing benefit of screening decisions as time goes on.

1.2.1 Using machine learning methods

To address these issues, machine learning (ML) algorithms are being explored to develop predictive disease models that can supplement physicians' knowledge for individualized decision-making. However, building effective models entails tackling several issues. Conventional predictive models frequently make unrealistic assumptions about the nature of routinely collected clinical data. For example, static models such as logistic regression, support vector machines, and naive Bayes introduce two limitations. First, they ignore the time-dependence in data generation as a disease evolves and is observed over time. Second, they are limited to quantifying the effect of past decisions on present observations. In comparison, dynamic models, like dynamic Bayesian networks (DBNs), address some of these temporal limitations but have other constraints (e.g., fixed frequency sampling of data over time). Dynamic

models that overcome these constraints, such as continuous time DBNs or recurrent neural networks (RNNs) have the following limitations when applied in disease screening: 1) they lack the ability to model the hidden disease state space, which is only partially observed through tests or patients' treatment response; and 2) they lack components that model patient benefits, and are hence limited to classification-oriented tasks (i.e., the models do not suggest the next best action to take). The fundamental difference between observational and decision making models is that the former classify patients as high or low risk solely based on disease outcomes whereas decision making models have decision recommendation components that take into account patient preferences, costs, and physicians knowledge in decision recommendation. Sequential decision making methods provide a potential solution. These approaches can integrate and analyze multiple sources of patient data, while handling issues related to temporal credit assignment [10]. For instance, to suggest individualized screening policies, Markov decision processes (MDPs) have been used to determine optimal time points for initiating antiviral therapy [11] and organ transplants [11, 12]. In particular, partially-observable Markov decision processes (POMDPs) have been applied to cancer screening (e.g., breast, colorectal, prostate) to determine policies based on patients' risk factors and prior screening results [11, 13–16]. Unfortunately, the construction of a POMDP is very involved. First, learning transition probabilities and observation probabilities for such models are not straightforward. Second, learning appropriate rewards is imperative to generate meaningful POMDP models and recommendations. Third, defining accurate initial beliefs is challenging. Lastly, scaling POMDP models to observation spaces of multiple inputs over time introduces computational complexities.

1.3 Contributions

This dissertation presents a framework to address the aforementioned challenges, developing data-driven methods that calculate the necessary components of a POMDP. This research

provides a foundation for efficient implementation of POMDPs for cancer screening processes, combining multiple ML-based approaches for optimal decision-making. This work is structured around four key contributions:

1. *Understanding how screening recommendations over time can be individualized.* I explore the issues around the development of dynamic models built from observational data and the systematic evaluation of such models over time. I developed this methodology to use sequential imaging observations for the prediction of a positive biopsy of cancer and evaluated this prediction based on patient's cancer outcome.

2. *The development of a multiple clustering imputation approach to handle missing data, introducing variability in uncertain variables via bootstrapping.* I developed this methodology for use with conventional disease risk models to generate initial beliefs for POMDPs with specified confidence intervals (CI).

3. *The development of a methodology that efficiently learns rewards functions for POMDPs and MDPs using experts' previous decisions.* This methodology uses the maximum entropy inverse reinforcement algorithm with the state and screening decisions of patients over time to learn state rewards. I extended this methodology to efficiently learn optimal rewards functions for disease screening.

4. *Development of a methodology that learns transition and observation probabilities for POMDPs through the integration of dynamic Bayesian networks in the learning process.* The use of this methodology includes the incorporation of past observations and actions into current decisions and modeling the observations of a POMDP model with a methodology that allows scaling into multiple simultaneous observations over time while capturing their interactions. I developed this methodology to compress multiple observations in POMDPs and then learn the transition and observation probabilities for any POMDP model.

1.4   Organization of the Dissertation

This dissertation is organized as follows. Chapter 2 provides a background on the clinical problem of lung cancer, along with current risk and diagnostic tools. I also provide a technical discussion of dynamic Bayesian networks; DBN applications in healthcare; a comparison of dynamic and static models; expert and learned models; POMDPs; and an overview of POMDP applications in medicine.

Chapter 3 covers the majority of research Contribution 1, expanding on the description of how to individualize cancer screening using DBNs and on how to evaluate dynamic models. This chapter includes information on how dynamic models outperform static models and how they perform comparable to experts.

Chapter 4 centers on research Contribution 3, explaining the methodology used to generate rewards for POMDPs. This chapter includes an adaptive maximum entropy inverse reinforcement learning algorithm optimization, experiments, and the multiplicative model used to generate state-action pair rewards. This method is evaluated on lung and breast cancer screening using the NLST and Athena dataset, respectively.

Chapter 5 explores the remaining methodologies used in my approach to learn POMDPs. More specifically, this chapter focuses on research Contributions 4 and 5, expanding on the transformation of the POMDP observation model to include a larger number of observations combined with temporal dynamic models. It includes an explanation on how transition and observation probabilities are learned using a naive Bayes DBN and explains how risk models and patient demographics can be used to generate individualized cancer beliefs.

Chapter 6 explores the utility of a multiple clustering imputation methodology that calculates imputed missing values and introduces a range of possible values for uncertain variables. This methodology combines clustering, bootstrapping, and a similarity algorithm to generate a range of plausible values for missing or uncertain values. It was applied on the Athena breast cancer screening dataset [17] and demonstrated the effect of missing values or uncertainty in patient management for the prescription of tamoxifen by generating

confidence intervals on individualized risk predictions. A variation of this methodology was used to impute missing values for demographic and clinical variables on the NLST dataset.

Finally, Chapter 7 synthesizes the findings of this dissertation, limitations, and future directions for this research.

Chapter 2

Background and Literature Review

2.1 Lung Cancer Risk and Diagnostic Prediction Models

Several studies have focused on developing lung cancer risk prediction models from epidemiological, demographic, and clinical data. Other studies have looked at methods for classifying pulmonary nodules as malignant or benign from pathological and imaging data. This section reviews several key studies in predicting lung cancer, the selection of risk factors by experts and/or automated systems, and the statistical models used to improve performance.

2.1.1 Risk models based on epidemiological and clinical risk factors

In 2003, Bach et al. [18] developed two logistic regression models to calculate the risk of developing lung cancer. The models were two 1-year risk models: one for obtaining the risk of dying from lung cancer, and one for the risk of dying without lung cancer. These models used features such as smoking, sex, and smoking history to calculate over 10-year period the cumulative probability of an individual being diagnosed with lung cancer [19]. Conin et al. [20] validated this model on individuals from the placebo arm of the Alpha-Tocopherol Beta-Carotene Cancer Prevention (ATBC) study. The model underestimated the observed lung cancer risk and the observed non-lung cancer risk individuals that smoked

less than 20 cigarettes/day. To improve lung cancer risk prediction, more individualized risk factors such as occupational exposures, previous respiratory disorders, family history, and other epidemiological factors must be considered. Moreover, validation of risk models in external datasets and datasets derived from diverse populations are necessary to assess the overfitting and usefulness of risk models. A single log-odds model based on logistic regression was developed from the Liverpool Lung Project Cohort (LLPC) [21]. Raji et al. evaluated this model in three independent external datasets from Europe and North America. The model's reciever's operating characteristic (ROC) area under the curve (AUC) in these datasets varied from 0.67-0.82 [22]. Spitz et al. [23] developed log-odds models for never, current, and former smokers. The models achieved an AUC of 0.57, 0.63 and 0.58, respectively. In the COSMOS trial [24], a model based on epidemiologic and clinical risk factors was developed to estimate the probability of an individual being diagnosed with cancer. The model was a multivariate Cox proportional hazard regression with nodule type as a categorical covariate, and was trained and tested on high-risk individuals. The model failed to recognize early cancer, but can be used for identifying lower-risk individuals who are usually over-diagnosed. Lastly, Tammemägi et al. [25] used the Prostate Lung Colorectal Ovarian (PLCO) trial dataset to develop 6-year lung cancer risk models. The utility of this model was evaluated incrementally using AUC as a metric. The models achieved high discrimination and calibration performance on their dataset, and used a wider range of risk factors than most of the aforementioned models.

2.1.2   Disease prediction models based on pathological data

Other risk prediction models focus on pathological data. Yu et al. [26] developed a fast correlation algorithm that performs feature selection on pathological data based on entropy measures to predict lung cancer. On a dataset of 32 cases and 57 attributes their method achieved an accuracy of 87%. On the same dataset [27] achieved an accuracy of 100% through the use of a 10-fold cross validation. Their model involved the use of principal component

analysis (PCA) for dimensionality reduction and a computational algorithm inspired by the immune system. The artificial immune recognition algorithm defines training and testing cases' features as antigens, then based on biological behavior mechanisms of T and B cells it classifies test cases using a k-NN algorithm into cancer or non-cancer. Other rule-based methods have also been applied on this dataset with inferior performance relative to those using dimensionality reduction methods [28, 29].

2.1.3   Disease prediction models based on clinical and imaging data

Thamilselvan et al. [30] used an enhanced k-NN algorithm to classify magnetic resonance imaging (MRI) as malignant or benign. They applied a weighted distance measure between images to classify new test images. Other studies focused on classification of pulmonary nodules as malignant or benign. Feature extraction, selection, and denoising of medical images are some of the steps to achieve malignancy identification. McWilliams et al. [31] developed parsimonious and fuller multivariate logistic regressions based on epidemiologic, clinical, and imaging variables such as nodule size and nodule consistency to predict the probability of developing lung cancer 24 years based only on baseline covariates. Their models achieved high discrimination and calibration on two datasets of 7,008 and 5,021 pulmonary nodules. Several studies [32, 33] developed deep learning systems with high accuracy, performing automatic feature exploration, denoising, and classification of pulmonary nodules. Other studies [34, 35] use expert-defined features such as histogram of oriented gradients (HOG), Tamura textures, and distance descriptors (skeleton measures, triangulation measures, Feret diameters) of nodules with support vector machines (SVMs) and achieve comparable performance.

2.2 Time Series Models for Disease Prediction in Chronic Diseases

The previously cited risk models are static: they only focus on snapshots of certain disease processes over time. With the advent of screening programs more longitudinal clinical and imaging data are routinely collected that can be used for lung cancer diagnosis. Statistical models without temporal components assume that risk factors/features are independent and ignore the inherent time-dependence in data generation. However, static models are easier to train and implement, whereas dynamic models are computationally expensive as they require a larger number of training samples and more parameters to learn [36]. Dynamic models take into account temporal patterns and are better suited at related identification and classification tasks. Various methods have been proposed in time series modelling of chronic diseases and the development of conditions over time. This section covers some of the approaches and challenges in time-based prediction, drawing on different health domains.

2.2.1 Time series analysis models

Time series analysis models are typically based on regression, determining correlations between variables. Generally, the aim of a regression model is to analyze short-term variations of an outcome of interest and the variable driving this outcome. For example, Bhaskaran et al. [37] used a Poisson regression to analyze the correlation between ozone levels and mortality. They demonstrated that by removing seasonal variations and long-term trends from the data the sign of the correlation between death and ozone levels is reversed, suggesting higher death with higher ozone levels. Ali et al. [38] developed a seasonal auto-regressive integrated moving average model (SARIMA) to forecast cholera outbreaks in Bangladesh. Cholera incidence was correlated with climatic and other environmental changes, such as temperature and rainfall. Hu et al. [39] compared a Poisson regression and a SARIMA model for predicting cryptosporidiosis with weather variability.

2.2.2 Conceptual modelling

A common problem in monitoring disease progression is the incorporation of knowledge into the process. Van Gerven et al. [40] introduced a carcinoid model that captures dependencies of pathophysiology in tumor growth. This model contains data on 129 patients diagnosed with a carcinoid tumor and predicts the future health of the individual as a probability curve over time. Nicholson et al. [41] combined state and transition models with dynamic Bayesian networks (DBNs) to model temporal systems in ecological problems. An interesting approach was introduced by Kohda et al. [42] to combine input from multiple time series sensors for safety monitoring. Their method introduces optimal logic by judging a decision, on multiple inputs, in a causal structure.

2.2.3 Data interpolation

Modeling disease progression requires a sufficiently large dataset, which poses a problem for rare diseases. Li et al. [43] used cross-sectional data from clinical trials to explain certain patterns in disease processes. In this work they explored an extension of a temporal bootstrap technique to generate the intermediate stages of a disease over time. They applied their approach on glaucoma, breast cancer, and Parkinson's disease data to explore how diseases progress when time series data are not available. A hidden Markov model (HMM) was used to identify "interesting" states of each disease. For example, in breast cancer they identified stable states that reflect benign and malignant tumors, and an intermediate state that can be characterized by a set of symptoms. For chronic obstructive pulmonary disease (COPD), Van der Heijden et al. [43] used bootstrapping methods to replicate time series data from smaller datasets, subsequently learning a set of predictive models able to predict COPD exacerbation events. All models were validated in terms of structural equivalence with an expert-defined model and evaluated with synthetic and external data.

Regression-based time series models can be used to study and predict discrete and continuous stochastic processes. They can be also used in the analysis of short- and long-term

patterns between variables. However, they cannot accommodate logic in forecasting, such as changes in outcome due to simple rule-based logic (e.g., a BRCA gene test, wherein a positive result increases the risk of breast cancer). Furthermore, such approaches consider variables independent and cannot analyze causal relationships between variables and the outcome. Although future observations are assumed independent of past decisions, this is not the case in healthcare. In contrast, graphical models can provide causal insights and represent associations between variables, but are often limited to discrete time models and are computationally expensive.

2.3 Technical Methods

2.3.1 Dynamic Methods

Dynamic Bayesian networks (DBNs)

A dynamic Bayesian network is a temporal extension of Bayesian networks (BNs). BNs offer complete representation and inference for a joint probability distribution. A BN is a directed acyclic graph (DAG), G, which utilizes the chain rule to represent joint probability distributions in a factorized way. A joint probability distribution consists of random variables that describe the domain of interest. The DAG (G) consists of a set of nodes (ν) that represent the random variables (X) of the joint distribution, and set of edges (ϵ) that represent probabilistic influences between nodes (i.e., $G = (\nu, \epsilon)$). Bayesian networks are composed of factors. Factors are functions or conditional probability tables (CPTs) that represent a conditional probability of the variable given its parent nodes. Thus a joint probability distribution can be factorized according to G if the joint distribution can be expressed as a product of factors [44]:

$$P(X_1, ..., X_n) = \prod_{n=1}^{n} P(X_i | X_1, ..., X_{i-1})$$

DBNs repeat the static interactions of a conventional Bayesian network over time [40]. They represent a joint probability distribution over temporal trajectories that specify the assignment of values to each random variable $X_i^{(t)}$ different time points t. A DBN follows the Markov assumption in which the current state of the system only depends on the past state of the system. Thus, in the case of a DBN, which is an unrolled Bayesian network, the random variable X_i the network will depend only on its parents, $Par(X_i)$

$$P(X_1, ..., X_n) = \prod_{i=1}^{n} P(X_i | Par(X_i)).$$

The structure and the probabilities $P(X^{(t+1)} | X^{(t)})$ can be assumed the same for all t (i.e., time invariant). Such a system is a stationary dynamical system. In this case the model can consist of two parts [40]: 1) a prior model that specifies the initial distribution of the process:

$$P(X^{(0)}) = \prod_{X^{(0)} \in \mathbf{X^{(0)}}} P(X^{(0)} | Par(X^{(0)}))$$

and 2) a transition model that specifies the evolution of the process across time points:

$$P(X^{(t+1)} | X^{(t)}) = \prod_{X^{(t+1)} \in \mathbf{X^{(t+1)}}} P(X^{(t+1)} | Par(X^{(t+1)}))$$

A DBN can be used to estimate conditional distributions through the use of the chain rule for Bayesian networks. The probability $P(X_i^{(t)} | \mathbf{X})$ below represents the conditional probability of variable $X_i^{(t)}$,given evidence about certain random variables $\mathbf{X} = (X_1, ..., X_{n-1}$

in a network's structure:

$$P(X^{(t)}|\mathbf{X}) = \prod_{t \in \mathbf{T}} \prod_{i=1}^{n} P(X_i^{(t)}|Par(X_i^{(t)}))$$

Learning DBNs

The (automated) learning of DBNs can be considered twofold: one involves learning the structure (DAG) of the network and the other learning the parameters of the nodes comprising the network. When learning the structure of DBNs there is a compromise between complexity (e.g., number of edges) and exactness. Including more edges results in a more complex model to parameterize, whereas including fewer edges means capturing less dependencies between variables. [44] observes that sometimes structures with fewer edges generalize better, even if they do not exactly represent the underlying distribution. Methods for structure learning include: constraint-based structure learning, which tries to find the best structure based on constraints that represent dependencies in a domain; score-based methods that select from a variety of structures, built from a given training dataset, choosing the one with the highest score on a test set [45]; and a "backward construction" process that describes the events occurring in the domain of interest (a dependent variable of interest and its associated variables are specified, and a structure representative of the events occurring in the domain is created). Model parameterization can be accomplished through the use of maximum likelihood methods; calculated based on dataset frequencies; or through experts' beliefs. The most popular algorithm for parameterizing DBNs is expectation maximization (EM). One advantage of the EM algorithm is the parameterization of models with and without missing data. The EM algorithm calculates log-likelihood estimates of the network parameters in an iterative manner, filling in missing values with statistical estimates using observed data, and subsequently re-estimates the parameters of the network. The use of frequency-based probabilities, computed from datasets, is usually used in cases where the dependencies between the variables are sufficiently few and there is a significant amount of

data to accurately represent the underlying distribution.

DBNs in Medicine

DBNs and BNs have been used in the domain of nosocomial infections [46], pneumonia [47], cardiac surgery [48], decision support in forensic psychiatry [49], gait analysis [50], osteoporosis [51], oral cancer [52], colon cancer [53], cervical cancer [54], breast cancer [55–57] and lung cancer [58]. In the domain of oral cancer, Exarchos et al. [52] developed a dynamic Bayesian network to predict post-treatment conditions as well as the time of cancer reoccurrence. The model utilizes a personalized genetic signature to analyze gene expressions over time and expression differentiation between patients with and without disease reoccurrence. The DBN achieved complete discrimination; however, the sample size of this study consisted of only six patients. [55] evaluated the performance of seven Bayesian networks as tools for breast cancer diagnosis through a series of four experiments and demonstrated that additional information is required to correctly identify breast cancer from cytological data. Gevaert et al. [57] evaluated three structure learning methods for the creation of a prognostic Bayesian network for breast cancer: a full integration method, in which both clinical and microarray data are provided to the learning algorithm; a decision integration method, in which the clinical and microarray data are used to learn two different structures; and a partial integration method, which involves an initial step of learning the network structure separately for both datasets and then combining the two datasets with the common outcome variable and revising the structure with both datasets. The partial integration method achieved the highest AUC of 0.845. This study demonstrated that the combination of clinical and microarray variables improves performance as the AUCs with only clinical and microarray variables is 0.804 and 0.798, respectively. In lung cancer, [58] proposed a Bayesian network built from both physical and biological data (biomarkers) for the prediction of local failure in non-small cell lung cancer (NSCLC) after radiotherapy. This integrated approach was tested on two different NSCLC datasets with the biological data contributing the most in the model's performance.

16

The BN was compared on one of the two datasets with the naive Bayes (NB) method and the tree augmented naive Bayes (TAN) method. The BN outperformed both the NB and TAN methods, with an accuracy of 80.18%, 74.05%, and 62.86%, respectively.

2.3.2 Dynamic versus static models

Many studies investigate the comparison of static and dynamic models based on performance improvement. In 2009, Charitos et al. developed a dynamic Bayesian network for diagnosing ventilator-associated pneumonia (VAP) in ICU patients. The model consists of two dynamic processes, colonization and pneumonia. The model was evaluated on a group of patients and compared with a static VAP model [59] using the Brier Score. The difference between the two models is the distinction of the aforementioned dynamic processes. The dynamic VAP (dVAP) compared with the original static VAP (sVAP) had a lower Brier score, 0.2376 and 0.3370, respectively, which demonstrates that the dVAP can distinguish VAP and non-VAP patients better. Watt et al. [51] developed a dynamic Bayesian network to predict the likelihood of a patient diagnosis with knee osteoarthritis (OA) as well as the symptoms associated with OA. This model was compared with a static BN and a logistic regression (LR) model using a random subset of 200 patients. The DBN outperformed both the LR and BN. Sensitivity analysis on the static and dynamic BNs identified knee OA and knee pain as general OA predictors.

2.3.3 Expert versus learned models

Cuaya et al. [50] proposed two DBNs to predict the risk of falls in elderly using spatiotemporal gait data obtained every 6 months for a period of 3 years. The first model was constructed using expert selected features whereas the second model used a forward sequential selection (FSS) algorithm. The models achieved an average precision of 70% in predicting imminent falls as well as in 6-month risk of falling. Both the computationally derived model and the expert selected model had comparable performances. Stojadinovic et al. [53] developed

machine-learned Bayesian belief networks (ml-BBN) to provide individualized survival estimates for colon cancer post-treatment at follow-up times of 12-, 24-, 36- and 60-months. AUCs for these ml-BBNs were 0.85. Maskery et al. [60] used breast pathologies to build a BN of breast pathology co-occurrence. The objective of this study was to learn a topology that sufficiently represents the observed data. The topology of the network was learned from a dataset of 1,631 pathology reports. Model verification involved an iterative derivation of the Bayesian network with 25%, 50%, and 75% of the original data randomly removed. The model regenerated 81%, 92%, and 97% of the dataset's structure. A percentage of 95% of the pathology co-occurrences concurred with the literature and expert opinion.

2.3.4 Partially-Observable Markov Decision Processes (POMDPs)

Markov decision processes (MDPs) are extensions of Markov chains that include an agent and decision-making process in an environment where outcomes may be stochastic and based on an agent's decisions over time (see Figure 2.1). A Markov chain has a set of states and a dynamic process that moves between states over time. The period of time over which an MDP is applied is referred to as a horizon. Moves between states are performed at each time step and are based on a transition model (T). MDPs consist of the following components:

- A set of states (S): States describe the world the decision problem is taking place.

- An initial state (s_0): The initial state represents the initial state of the agent.

- A set of actions (A): Represent the actions that can be taken in each state.

- A transition model ($T(s, a, s')$): The transition model describes the dynamics of our environment, the state-action dependent transition probabilities between states. s' is the current state, a is the action, and s is the state the agent transitioned.

- A rewards (R) function: Rewards are a set of scalar values that represent the value received in a state. Rewards embed knowledge about the process and ultimate goals.
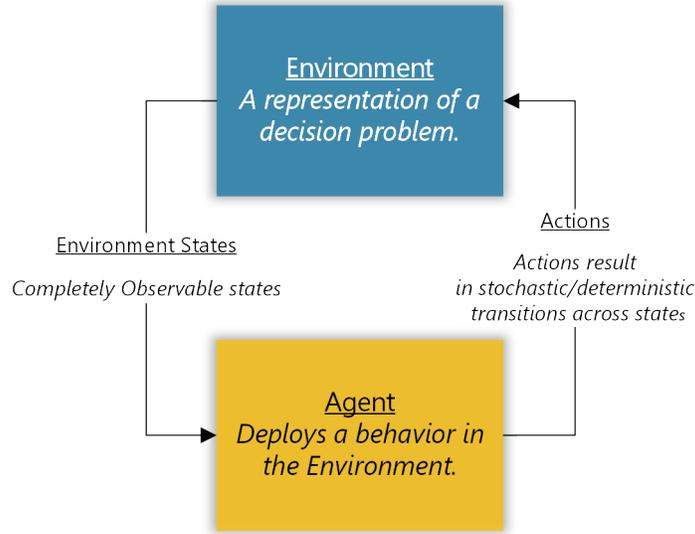
Figure 2.1: MDP model, adapted from [1]

MDPs are Markovian processes, as the next state and the expected reward in that state solely depend on the previous state and action taken (i.e., the Markov property) [1,61]. A solution to an MDP is called a policy. A policy represents a function that takes a state as an input argument and returns an action or a mapping from states to actions, $\pi(s) \rightarrow a$. A stationary policy specifies an action for each state irrespective of time step t. A non-stationary policy is indexed by time (π^t) and summarizes the information to perform the optimal action at each time step t [62]. The optimal policy (π^*) represents the policy that maximizes long-term expected rewards, thus the optimal policy returns an action for each state that is the action that returns the maximum expected reward from that state. This maximum expected reward of an optimal policy is the sum of the product of the transition probability of ending up in a state (s) times the expected utility[1] of that state (see Equation 2.3.1).

$$\pi^*(s) = \mathbf{argmax}_a \sum_{s'} T(s, a, s')U(s') \tag{2.3.1}$$

A utility function of rewards over a sequence of states is the sum of rewards of being in a state over time. Rewards over time are discounted, as getting the same reward in the present

---

[1]Note that utility and reward are not the same thing ($R(s) \neq U^T(s)$). Rewards represent the short-term value of an action whereas utilities represent the long-term value of an action.

Algorithm 1: Value Iteration Algorithm

   Input: S, A, R, ϵ
   Output: Value function $V(s)$
   Randomly initialize $V(s)$;
   repeat
      for $s \in S$ **do**
         for $a \in A$ **do**
            $Q(s,a) = R(s,a) + \gamma \sum_{s' \in S} T(s,a,s')V(s')$;
         end
         $V(s) = \mathbf{max}_a Q(s,a)$;
      end
   until $\epsilon < |V_t - V_{t-1}|$;
   return $V(s)$
---

rather in the future is more beneficial. Thus, utilities can be defined as the discounted sum of rewards over time:

$$U(s_0, s_1, s_2, ...) = \sum_{t=0}^{\infty} \gamma^t R(s_t) \tag{2.3.2}$$

To learn a policy, we need a value function that represents the value we obtain from being in a state over time. The value iteration algorithm can be used to estimate the aforementioned value function over the states of a process with finite or infinite horizons.

Value iteration. The value iteration (VI) algorithm computes the value function of an MDP by finding a sequence of value functions, each one derived from the previous one. Value iteration uses the *Bellman equation* (see Equation 2.3.3) recursively to compute the optimal value function. Interestingly, the Bellman equation accounts for every component of an MDP: the reward of being in a state ($R(s)$), the discount factor (γ) and the expected utility ($U(s')$) of performing an action (a) in state s and ending up in state s'.

$$U(s) = R(s) + \gamma \, \mathbf{max}_a \sum_{s'} T(s,a,s')U(s') \tag{2.3.3}$$

In the VI algorithm, we start with arbitrary utilities and update the utility of each state based on the states we can reach from this state. This process is repeated until convergence. Algorithm 1 shows the pseudocode for value iteration. Once we have a value function (i.e.,

a representation of utilities at each time step of the process for each state) we can derive an optimal policy.

In Figure 2.2, an MDP problem of a 2D maze is depicted [2]. In this "grid world" actions do not always go as planned. There are four possible actions: going north, south, west, or east. There is 0.8 probability of moving in the desired direction and 0.2 probability of moving in the wrong direction. For example, if the agent wants to go north there is 0.1 probability that it travels west and 0.1 probability that it travels east. The agent's goal state is the (1,4) state in which it receives a reward of +1. The (2,4) state represents the death state, this is the state the agent wants to avoid as in this state it experiences a negative reward (-1). In all other states the agent experiences a small negative reward by delaying to enter the goal state. After five iterations of the VI we obtain an optimal policy. The optimal policy represents the optimal actions to perform in each state. Figure 2.2 depicts the utility of each state. Intuitively, the reward of the goal state propagates out as the agent aims to go there and the reward of the death state propagates less as the agent tries to avoid this state.

From MDPs to POMDPs. POMDPs are MDPs with two additional components: 1) a set of observations, which are tokens describing the potential observations in a problem space; and 2) an observation model, which describes the probability of seeing a given observation in different states. In a POMDP, states are partially observable. The agent makes observations by moving across states and performs actions based on the type of observation it observes and its current belief state (see Figure 2.3). In partially observable environments the agent does not know the state it is in; instead, it has a belief (e.g., in a two state problem there might be a 60% probability of being in one state and 40% of being in the other state) of being in each state of the environment. At each time point the agent uses an updated belief state that summarizes previous experiences. The belief state is updated based on the observation and the previous belief state. The belief state[2] represents a probability distribution of being

---

[2]The belief state is a probability distribution over the states of the process. The use of a probability distribution as belief state instead of discrete state makes the overall process non-Markovian as the probability distribution captures more than the previous step of the process.
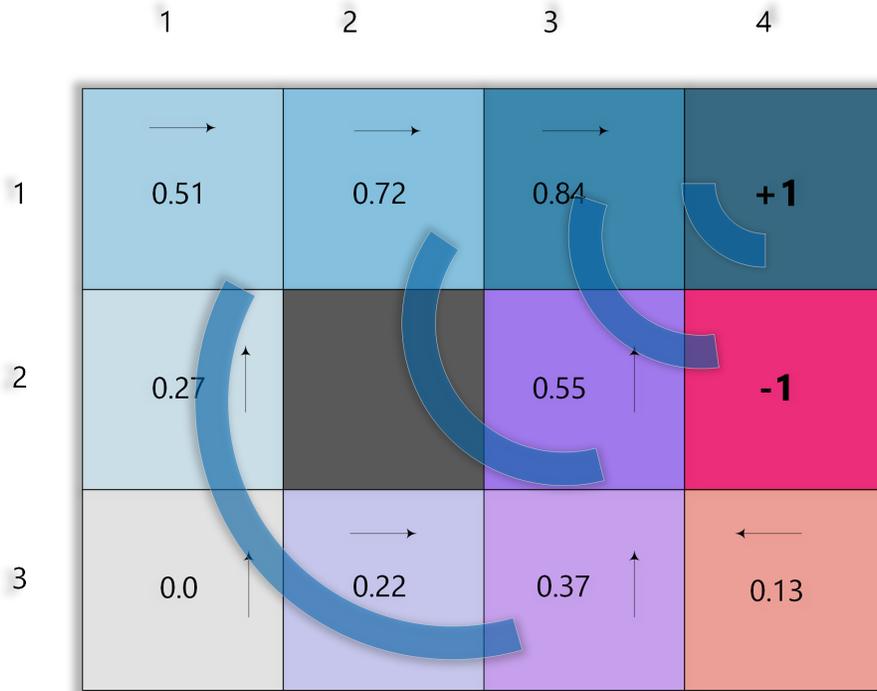
Figure 2.2: Maze MDP example. Reward state (1,4) will propagate out towards the other states whereas the state (2,4) will propagate less as the agent tries to avoid it. Adapted from [2]. The gray block represents a wall, when the agent hits a wall it bounces back to its original state.

in the states of the process.
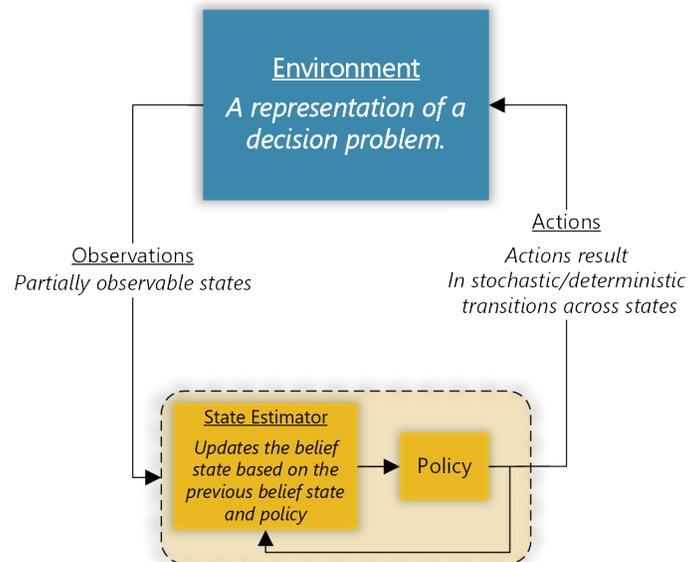
Figure 2.3: POMDP model. Adapted from [1].

For example, Figure 2.4 depicts a maze problem in which the agent has to find its goal state, in this case state 4. The agent does not know its initial state but the agent has an initial belief that represents a probability distribution of the likelihood of being in any of the given states. If the agent is equally likely to be in each of the states initially, the belief state would be: [0.333 0.333 0.333 0.000]. There are only two observations in this problem: the observation of being in the goal state and the observation of not being in the goal state, which is experienced in all other states except state 4. The probability of performing a successful action (the decided action) is 0.9 and the probability of failing to do so is 0.1. Thus, if the agent takes the action east and does not observe the goal state then the belief state would change to represent the experience of the agent. The new updated belief would be: [0.222 0.444 0.333 0.000]. The state estimator computes the belief estimates at each time step. Subsequently, the actions are chosen based on the policy component of the POMDP. The policy of a POMDP represents a mapping from belief states to actions.

To summarize, POMDPs are defined in terms of the same components as an MDP and the following additional components:

- A set of observations (O): Representing the observations sensed in each state.

- An observation model ($O(a, s', z)$): The probability of observing z after taking action a and transitioning to state s'.
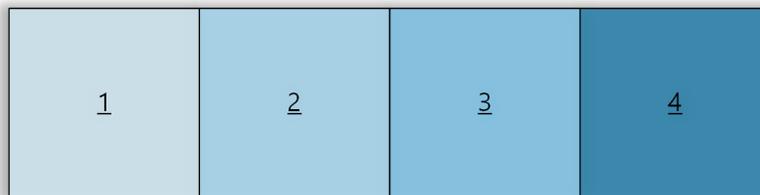
Figure 2.4: Maze POMDP example. Our world has four states, one of which is the goal where the agent receives a positive reward. In all other states the agent experiences time loss (a mild negative reward) as the agent is shortening its time in its target state. Adapted from [1].
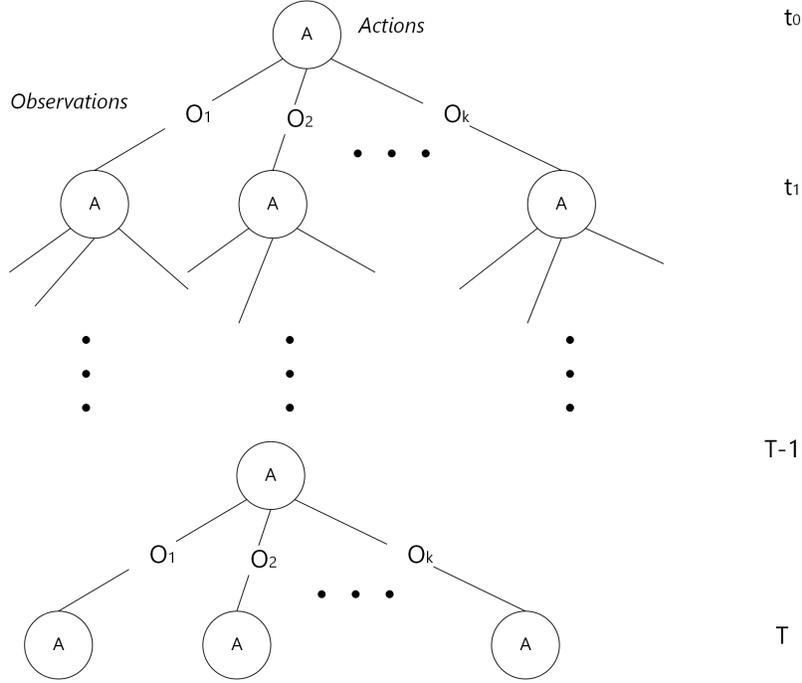
Figure 2.5: t-step policy trees. A: Actions; O_k: Observation; T: horizon total time points; t_i: time point (epoch). Adapted from [1].

The non-stationary policy of a POMDP can be depicted by a very large policy tree (p), which represents every possible combination of observation and action as shown in Figure 2.5. The aim of the agent is to be able to calculate the value of every possible policy tree (see Equation 2.3.4) from a finite set of policy trees P, with every possible belief state:

$$V_p(s) = R(s, a(p)) + \gamma \sum_{s' \in S} T(s, a(p), s') \sum_{o_i \in O} O(s', a(p), o_i) V_{o_i(p)}(s') \qquad (2.3.4)$$

The number of policy trees can be infinite; however, it is finite for a finite horizon process (non-stationary policy). Each policy tree is of finite length. Its length is defined by the number of time steps the process propagates over time. The agent aims to calculate a value for each policy tree. Thus, by selecting the policy tree with the maximum value at each belief point (which corresponds to a time step, as the belief is updated periodically after each action) the agent can obtain an optimal policy. The optimal value function can be obtained using the value iteration algorithm for POMDPs (see below) or other VI approximation
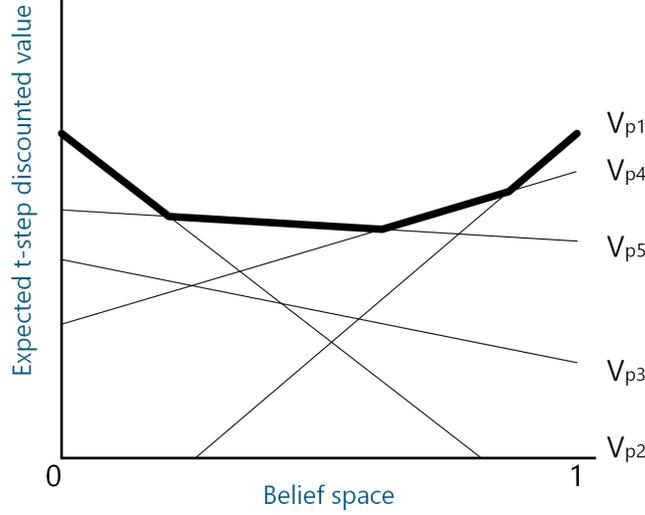
Figure 2.6: Optimal value function of a 2-state problem. Adapted from [1].Each line represents the value function of a policy tree. The bold black line represents the best policy tree at each partition of the belief space.

algorithms. A detailed explanation of POMDP solvers can be found in [61, 63, 64]. Per Sodnik et al. [63], V_t is a piecewise-linear and convex function with respect to the belief state, b (see Equation 2.3.5). The weights of this linear function are called "alpha" vectors and are unique for each policy tree. Alpha vectors represent the utility of each policy tree's state ($\alpha_p = \langle V_p(s_1), V_p(s_2), ..., V_p(s_n) \rangle$). In each region of the belief space there is a policy tree with the maximum value function. The optimal action at each region of the belief space is the action at the root of the best policy tree.

$$V_t(b) = \mathbf{max}_{p \in P} \ b \cdot \alpha_p \tag{2.3.5}$$

Figure 2.6 depicts the value function of a two state POMDP. Belief states in the middle of Figure 2.6 are more uncertain whereas decisions on the edges have a higher value and more confidence. For example, in the two-state problem, being in the middle of the state space implies a $50 - 50$ chance of being in either state (randomness) and a lower utility value. The optimal value function represents the highest value over all value functions for each belief state point (bold line Figure 2.6).

Value iteration for POMDPs. Under infinite horizons the value function can be discontinuous and non-linear. Here we present the value iteration algorithm for finite horizon POMDPs. This algorithm accepts as input the components of the POMDP and the planning horizon of the model and returns a set of alpha vectors each representing a linear function over the belief space (value function). These vectors are computed recursively by executing an action followed by an observation at each time step of the horizon and generating new alpha vectors. The algorithm generates a linear function (alpha vector) for each combination of action, observation, and linear function of the previous value function. For each iteration of the VI algorithm pruning is used to reduce the subsequent iteration's computations of linear functions by removing dominated linear functions (i.e., linear functions that do not constitute the value function see V_{p3} in Figure 2.6). Aglorithm 2 shows the pseudo-code of the value iteration algorithm for finite POMDPs.

Approximation algorithms for solving POMDPs. POMDP models can be solved through the value iteration (VI) algorithm. However, the number of possibilities to be considered is exponential in terms of the number of states, actions, and observations modeled. At each time step, the VI algorithm enumerates k^Ω new policies trees, where k is the previous time step number of policy trees and Ω is the number of observations. Each policy tree represents a linear function. For an infinite horizon process the value function will have infinite linear functions, a key reason why POMDPs are often considered impractical. To solve infinite horizon problems, we can use approximation algorithms [61,63–65], providing significant speed-up. Here, we explain the QMDP approximation algorithm, shown in Algorithm 3. QMDP solves the POMDP problem as an MDP and then generalizes the value function into a POMDP. The computed value function for MDPs is used to compute the Q matrix. The Q matrix as an (n_s, n_a) where n_s is the number of states and n_a is the number of actions. Each column of the Q matrix is an alpha vector that can be used by the POMDP model for decision support. The main disadvantage of the QMDP algorithm is that it dismisses the state uncertainty described in POMDPs but solves the POMDP with MDP computational

26

Algorithm 2: Value Iteration Algorithm for Finite POMDPs. Adapted from [66].

   Input: S, A, R, O, γ, T
   Output: Value function Γ
   $\Upsilon = (0; 0, ..., 0);$
   for $\tau \in T$ **do**
      for $\upsilon \in \Upsilon$ **do**
        $\Upsilon' = \emptyset$
        for $a \in A$ **do**
           for $z \in O$ **do**
              for $j \in S$ **do**
$$\upsilon_{a,z,j} = \sum_{i=1}^{N} \upsilon_i p(z|s_i) p(s_i|a, s_j)$$
              end
              for Finite POMDPs
           end
        end
      end
      for $a \in A$ **do**
        for *each υ combination with O* **do**
           for $i \in S$ **do**
$$\upsilon'_i = \gamma[r(s_i, u) + \sum_z \upsilon^{k(z)}_{u,z,i}]$$
           end
           add $(a; \upsilon'_1, ...\upsilon'_N)$ to Υ'
        end
      end
      prune Υ'
      $\Upsilon = \Upsilon'$
   end
   return Υ
---

time complexity. To select optimal actions that maximize expected utility we use Algorithm 4, which given a belief and the Q matrix computes their dot product to compute the utility of each action when being in a belief (b).

2.3.5 Learning POMDPs

The transition, observation, and reward components of a POMDP encompass knowledge that can be "learned." There are two ways to learn a POMDP: using retrospective data or through the use of reinforcement learning. The transition and observation model can be ascertained from existing retrospective data in a frequency-based way; using the EM algorithm;

Algorithm 3: QMDP Algorithm

Input: S, A, R, ϵ

Output: Q matrix

Computing the Q matrix;

$V(s) = MDP\_VI(S, A, R, \epsilon)$;

for $s_i \in S$ **do**

    for $a \in A$ **do**

        $Q(s_i, a) = R(s, a) + \sum_{s_j \in S} T(s_j, a, s_i) V(s_i)$;

    end

end

return Q

---

using expert's beliefs; or through Bayesian reinforcement learning, which involves planning and learning. The learning of the transition and observation models requires probability updating. An effort by Chrisman et al. to obtain more accurate transition and observation models and to learn the states mostly associated with a decision problem using a variant of the Baum-Welch algorithm for POMDPs is explained in [67]. Mescheder et al. [68] used a variation of McCallum's utile distinction memory algorithm [69], a way of computing a POMDP model from existing data using the Baum-Welch algorithm. This algorithm can be used to identify the suitable number of states in a decision problem. The reward function can be learned from data reflecting the experience (e.g., surveying several agents experiencing the same problem), using inverse reinforcement learning (IRL) and modeling a problem's reward function to only penalize critical mistakes. Ng et al. [70] addressed the problem of IRL in MDPs with three different approaches: the first two approaches try to learn a reward function when the policy is known, while the last learns a reward function with a known policy for a finite set of observations. Interestingly, Choi et al. [71] extended IRL for POMDPs

---

Algorithm 4: Action selection Algorithm

Input: Q, b

Output: a_{opt} optimal action

Given belief b;

$a_{opt} = argmax_a \sum_{s_i \in S} b(s_i) Q(s_i, a)$

return a_{opt}

---

to estimate a reward function of a POMDP in the case that the expert's policy is available and the case in which the observed trajectories (retrospective data) are available.

2.3.6   Learning reward functions

Although Markov decision processes and POMDPs are used in a number of domains, their application in healthcare is limited and few strategies exist for estimating the associated reward functions that drive agent behavior in clinical settings. Taken from the perspective of epidemiological and health services research, different cost and patient benefit metrics are frequently adapted for optimization. Classic examples include: Bennet et al. [72], who proposed a cost-effectiveness metric based on the cost required to obtain one unit of outcome change (CPUC); Hauskrecht et al. [73], who designed a reward model that combines economic cost and patient quality of life measures; and Tusch et al. [12], who predicated rewards on 30-day mortality risk for a surgical procedure. In contrast, in this dissertation I take advantage of growing amounts of longitudinal data, using recorded information and actions from electronic health records (EHRs) and other observational data sources, to learn a POMDP reward function that imitates expert physicians' behavior for desired health outcomes. Specifically, IRL is proposed for this task.

Briefly, IRL addresses the problem of obtaining a reward function given an agent's optimal behavior over time towards a stated goal. A reward function for the environment is unknown and is hence learned through empirical investigation of sensory inputs (i.e., observations) that progressively change the agent's selection of different actions. Two families of IRL algorithms exist: 1) linear programming (LP) methods [70, 74]; and 2) probabilistic IRL algorithms [75, 76]. Examples of LP-based methods include:

- Ng et al. [70] used IRL with MDPs and developed algorithms representing the reward function as a linear combination of basis functions, called features, using linear programming to compute the weights of the basis functions. These algorithms were applied in discrete and continuous navigation problems such as the "mountain-car" problem to

compute a reward function that describes the observed behavior of an agent.

- Abbeel et al. [74] developed IRL algorithms that use Monte Carlo to compute feature expectations; quadratic programming (QP) to compute feature weights; and reinforcement learning (RL) to compute optimal policies. A simpler algorithm was also developed, the projection algorithm, in which the QP step in no longer used to compute feature weights, but instead the experts' feature expectations are computed using orthogonal projection and then used to compute feature weights. These algorithms were applied on car simulation examples with the learned agent achieving expert optimal behavior.

While potentially more computationally complex, probabilistic IRL approaches have two advantages: they guarantee a unique solution for deterministic MDPs; and compared to LP methods, they can handle stochasticity in the data [77]. Vroman et al. [75] developed a maximum likelihood IRL algorithm using clusters of experts' data trajectories to characterize different intentions. Applying the maximum likelihood IRL algorithm to each cluster subsequently derives a reward function representing the experts' behavior. Ziebart et al. [76, 78] describe a probabilistic IRL algorithm that employs the principle of maximum entropy, dealing with noise and imperfect behavior as it normalizes globally over behaviors. This algorithm was applied on route preference modeling using 100,000 miles of collected GPS data of taxi-cab driving. In this approach, demonstrated for modeling routing preferences of vehicle drivers, behaviors with higher rewards are exponentially preferred by the algorithm when learning the reward function. In this dissertation, we build on and adapt this approach to obtain reward functions for cancer screening POMDPs.

2.3.7   MDPs and POMDPs in Medicine

Both MDPs and POMDPs are increasingly being used in clinical screening and treatment decision-making. In particular, POMDPs have been used in the domain of epidemic control

[11], drug infusion [11], organ transplantation [11,12], spherocytosis [11], sepsis [79], diabetes [80], ischemic heart disease [73,81], Parkinson's disease [82], colorectal cancer [13,15], breast cancer [14,83,84] and prostate cancer [85]. Examples are as follows:

1. **Large state space.** POMDPs cannot model a very large state space as the number of policy trees grows exponentially. MDPs, however, have tractable state spaces. Chhatwal et al. [86] developed a finite-horizon MDP model to investigate whether the decision of performing a biopsy to detect breast cancer changes with age. The model consisted of 100 states, each one corresponding to a risk score, and two actions: to perform an immediate biopsy or wait for the next annual mammogram. The state of each breast cancer patient was computed through the use of a BN for breast cancer. The model did not consider the cost of mammograms or biopsy in the reward model. This model was compared with radiologists' decisions. The model had an improved false positive rate and reduced true positive rate in comparison to the radiologists. Similarly, Alagoz et al. [83] built an MDP based on the same BN to define the state space of the MDP. The result of each screening was used in the BN to obtain a risk of cancer (state). To address the problem of missing screens they used interpolation methods to compute the risk at each screen of their dataset. A solution to large state, observation, or action spaces when using POMDPs are approximation POMDP solver algorithms discussed in section 2.3.4, such as the QMDP algorithm.

2. **POMDPs derived from microsimulation models.** Erenay et al. [15] developed a POMDP model for colorectal cancer that considers both static (e.g., gender) and dynamic (e.g., age) features when updating belief states. This addition of dynamic features is achieved through the addition of completely observable risk states (the risk of having adenomatous polyp and cancer). The aim of the model is to maximize quality-of-life (QoL) years while minimizing harm. Transition probabilities of the model are estimated through the use of a microsimulation model, the Surveillance, Epidemiology, and End Results (SEER) registry, and clinical literature. Also in the domain

31

of colorectal cancer, Lesno et al. [13] implemented a POMDP model to monitor the natural history of colorectal cancer. The model included the costs of various screening and treatment procedures in its reward model. Model parameters were obtained from published literature. Specifically, this study demonstrated that a colonoscopy every 10 years was the superior screening technique and that patient compliance rate plays a significant role in the cost-effectiveness analysis, which in turn determines the optimal screening strategy. Maillart et al. [84] formulated a POMDP to model the natural history of breast cancer. This model generates a menu of efficient policies from which the patient can select the policy that better suits its financial capabilities and risk of death from breast cancer. Transition probabilities were obtained from the SEER registry and clinical literature. Turgay and Alagoz et al. [87] developed a modeling framework to obtain individualized mammography-screening decisions with the goal of increasing life-savings in high-risk breast cancer cases. The model accounts for both static and dynamic risk factors; however, it does not consider the financial costs associated with screening. The transition model of this framework was developed through the use of a microsimulation model. Zhang et al. [85] used a POMDP model to formulate optimal policies that balance early detection benefits of prostate cancer with the side effects of early treatment and short-term negative impact of biopsy. In prostate cancer screening each patient receives a prostate-specific antigen (PSA) screening, based on the outcome of this screening test; a patient is referred for a biopsy. The model has the capability of applying a control-limit type policy (e.g., discontinue screening), which is an intuitive management strategy for elderly patients.

2.3.8 Summary and Current Challenges

Disease screening decision-making algorithms have been proposed in different domains. The above literature review uncovered the following challenges in designing and evaluating disease screening decision making frameworks:

- BNs or DBNs are a class of probabilistic algorithms that can be used to infer the probability of disease and then use this probability for decision making. However, such approaches do not readily infer optimal actions based on long-term considerations for the patient or otherwise (e.g., resource utilization, costs). POMDPs and MDPs have the potential to include patient preferences, consider complications, and other factors over time into decision-making processes through reward functions. Nevertheless, computing sensible rewards from existing data rather than QALYs is challenging. Thus far, the most promising computation of rewards functions in the literature is through the use of IRL algorithms. However, IRL algorithms only compute state rewards and not state-action pair rewards. In Chapter 4, I present an adaptive MaxEnt IRL algorithm that learns state and action rewards from data, and combined with a multiplicative model defines reward functions for MDPs and POMDPs.

- Defining the state and observation space for POMDPs is challenging. It is difficult to incorporate all of the temporal observations and their interactions in an efficient way in POMDPs, given the the myriad sources used in patient care. Presently, most POMDP observation models assume that observations are independent. In addition, the estimation of transition and observation probabilities from data simultaneously is problematic, with many studies turning to micro-simulation models to define required probabilities. In Chapter 5, I describe an alternate methodology that uses DBNs to address these aforementioned limitations by compressing observations into the probability space.

- To initiate the individualized decision making process using POMDPs an initial belief of disease is required. Although risk models have been successful in estimating the risk of disease they cannot be immediately expanded into a probability distribution to be used as beliefs over multiple cancer states. This issue is addressed in Chapters 4 and 5 using the Tammemägi and Gail risk model in breast and lung cancer screening,

respectively. Additionally, most established risk prediction models require complete patient information for the estimation of risk. Clinical datasets suffer from missing data, hence a methodology that can impute missing values accurately from existing clinical datasets is required. In Chapter 6, I present a multiple clustering imputation approach that can use clinical datasets with missing values to calculte missing values.

- Finally, a systematic evaluation of POMDPs in healthcare has not yet been presented in literature, with most studies instead of presenting classical evaluation metrics used in medical studies such as true positive/recall (sensitivity), false positive, false negative, true negative (specificity), and precision (positive predictive value) to present an estimation of the cumulative rewards a patient receives following the optimal policy. Arguably, these policies should instead be evaluated on patient outcomes and compared with other decision-making models. This approach to evaluation is presented in Chapter 3 for DBNs and subsequently modified for POMDPs in Chapters 4 and 5.

Chapter 3

Personalizing Lung Cancer Screening using Dynamic Bayesian Models

3.1 Overview

The use of low dose computed tomography (LDCT) in lung cancer screening for high risk individuals was established after the National Lung Screening Trial (NLST), demonstrating that annual screening with LDCT reduced the risk of mortality by 20% for high risk individuals (relative to x-ray screening). Despite this, there is uncertainty amongst practitioners for using CT in the context of screening [88]. Predictive models that can account for individualized risk factors and provide personalized screening recommendations over time, can improve the lung cancer screening process, potentially reducing the number of false positives and possibly detecting cancers earlier, thereby effecting treatment with improved patient outcomes. This chapter describes how recommendations can be individualized over time in the context of lung cancer screening. We explore issues surrounding the development and evaluation of a dynamic Bayesian network (DBN), built from the NLST dataset, to predict the development of lung cancer in high-risk patients. A comparison of DBNs built

35

using the "*backward construction*" method and *"learned"*[1] DBNs is performed. In addition, a comparison of the DBNs' performance versus experts and other predictive models was performed. Relative to existing predictive models, the proposed methodology has several advantages. First, it can make sensible predictions even with missing data, a common occurrence in real-world settings (e.g., a missed screening exam). Second, it is built on top of a lung cancer state-space defined on cancer staging. This state space unites lung cancer risk factors and diagnostic procedures in a meaningful network structure while also enabling the flow of probabilistic influence between these variables. Third, contrary to existing predictive methods for lung cancer screening, this methodology and in particular DBNs can explain and show the contributing factors for its predictions (i.e., factors investigated in lung cancer screening). This chapter includes the results of multiple evaluation strategies and discusses the advantages and limitations of this methodology.

3.2   Methods

We used the NLST dataset to create DBNs for the prediction of lung cancer incidence. The description of the dataset, overall methods, measured outcomes, and statistical evaluation methods used in this study are as follows.

3.2.1   The NLST dataset

The NLST is a randomized, multi-site trial that examined lung cancer-specific mortality among participants in an asymptomatic high-risk cohort. Subjects underwent screening with the use of low-dose CT or a chest x-ray. Over 53,000 participants each underwent three annual screenings from 2002–2007 (approximately 25,500 in the LDCT study arm), with follow-up post-screening through 2009. Lung cancers identified as pulmonary nodules were confirmed by diagnostic procedures (e.g., biopsy, cytology); participants with confirmed lung

---

[1]Here, learned DBNs represent models generated through the use of structure learning methods.

	First Screening	Second Screening	Third Screening	Post-screening	Post-trial	Total Cases
Remaining non-cancer subjects	25,530	25,217	24,842	24,477	24,461	-
Individuals with confirmed cancer	305	174	223	365	16	1,083
Deceased subjects	11	139	152	-	-	302
Total subjects	25,846	25,530	25,217	24,842	24,477	-

Table 3.1: NLST dataset, detailing the determined health state of a subject after each screening exam. Post-trial cancer cases represent the cancer cases that lung cancer was the cause of their death and were not identified as lung cancer cases through the NLST trial. The number of patients shown represent the patients for which we have information about the development of lung cancer. A cancer incidence occurring after the first screening and before the second screening was assumed to be a first screening cancer. A cancer incidence occurring after the second screening and before the third screening was assumed to be a second screening cancer. A cancer incidence occurring after the third screening and before the post screening period was assumed to be a third screening cancer. The above information was computed from the NLST dataset under our possession.

cancer were subsequently removed from the trial for treatment.

The NLST dataset provides a longitudinal perspective on high-risk lung cancer patients in terms of demographics, clinical history, and imaging data. We used subjects from the LDCT arm, across all three screening events and the post-screening period of the trial. Information used in our study includes: demographics (e.g., age, gender, body mass index); smoking history; family history of cancer; personal history of cancer; history of comorbidities related to lung cancer; occupational exposures (e.g., asbestos, coal, chemicals); and LDCT screening outcomes. Table 3.1 summarizes the number of cases determined to have cancer during any of the three imaging points of intervention (and the remaining number of non-cancer patients), as well as post-screening cancer patients (i.e., those individuals who went on to develop lung cancer after the third screening event).

Based on the true state of each patient (i.e., cancer or non-cancer) we designed a simplified state space model representing the "ground truth" disease state of each patient, after each screening time point. Figure 3.1 represents the state-space and the allowed transitions through these states. No-Cancer (NC) is the state in which the individual has no abnormalities or has abnormalities that are not suspicious for lung cancer (e.g., lung nodules smaller than 4 mm). The In Uncertain (U) state captures an individual who has abnormalities suspicious for lung cancer (e.g., findings larger than 4 mm). In terms of lung cancer staging,

the U state captures Stage 0 and occult carcinoma stages [89]. The Invasive-Cancer (LC) state represents individuals with confirmed diagnoses of cancer through the use of additional diagnostic procedures (e.g., biopsy). The LC state captures Stage IA–IV lung cancers. The Treatment state represents the state in which the individual was confirmed with cancer and is receiving treatment. Lastly, the Death state indicates an individual who is deceased, either from the cancer (without treatment) or due to some other cause. From this state model, the three cancer-related states (NC, U, LC) were used to represent discrete characterizations for a given patient's likelihood of cancer following screening observations over time.
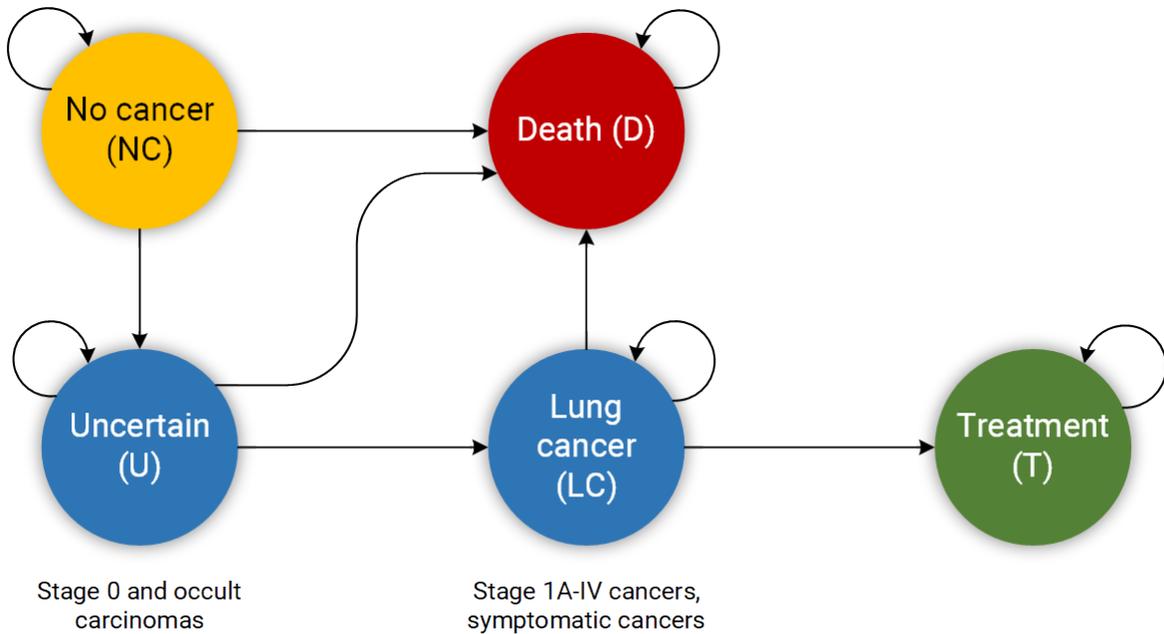
Figure 3.1: The underlying disease state space model for lung cancer used in this study, modeled after the process flow in the NLST. The arrows depict allowed transitions in the state space. In the Non-Cancer state, where everyone starts, the individual has no abnormalities or abnormalities smaller than 4 mm. In the In Uncertain state the individual has abnormalities larger than 4 mm, which are not confirmed to be cancerous. In the Lung Cancer state the individual is confirmed to have cancer through the use of diagnostic procedures, such as biopsy. In the Treatment state the individual is receiving care for the cancer, and is removed from the screening process. Finally, in the Death state the individual is deceased. The process described in this study terminates when an individual enters the Death or the Treatment state. The transition from the Treatment to the Death state is not depicted here as we only focus at the process of identifying an individual with lung cancer (e.g., an individual with lung cancer whose process ends when the individual enters the Death or Treatment state).

3.2.2 Dynamic Bayesian networks

A DBN can be used to estimate conditional distributions through the use of the chain rule for Bayesian networks. This ability was used in our lung cancer screening DBN to obtain the probability of a positive outcome of a biopsy for a given individual. Equation 3.2.1 represents the conditional probability of variable $X_i^{(t)}$ given evidence about certain random variables $\mathbf{X} = \{X_1, ..., X_{n-1}\}$ in the network structure.

$$P(X_i^{(t)}|\mathbf{X}) = \prod_{t \in T} \prod_{i=1}^{n} P(X_i^{(t)}|Par(X_i^{(t)})). \tag{3.2.1}$$

An example of the computation of the probability of the Biopsy outcome on a patient at the second screening ($t = 1$) based on the networks in Figure 3.2 is shown below. The computation of the conditional probability is based on the evidence of the individual on the variables of the model:

$\mathsf{P(Biopsy)}^{(1)}|\mathsf{Gender} = \mathbf{Female}, \mathsf{Family\ History} = \mathbf{Yes}, \mathsf{Body\ Mass\ Index} = \mathbf{Obese},$

$\mathsf{Work\ Exposure} = \mathbf{Yes}, \mathsf{Disease\ History} = \mathbf{Yes}, \mathsf{Age} = \mathbf{64}, \mathsf{Cancer\ History} = \mathbf{No}, \mathsf{Smoking\ Status} =$

$\mathbf{Yes}).$

3.2.3 The lung cancer screening DBNs

Deriving a DBN broadly involves two steps. First, deriving the structure (i.e., a directed acyclic graph) and second, parameterizing the network structure (i.e., estimating the probabilities for the CPTs of the network). In this work, we used the NLST dataset to build five different variations of networks: three expert-driven DBNs ("*backward construction*") and two DBNs derived from structure learning methods. Specifically, the models are as follows:

- The expert-driven DBNs consist of two *Forward-Arrow* DBNs (see Figure 3.2a) and one *Reversed-Arrow* DBN (Model **B**, see Figure 3.2b): 1) a *Forward-Arrow* DBN using a NoisyMax gate (Model **A**) for parameter reduction of the Cancer node, and for comparison, 2) a *Forward-Arrow* DBN without a NoisyMax gate (Model **C**); and 3) a

Reversed-Arrow DBN (Model **B**, see Figure 3.2b), providing an equivalent naïve Bayes classifier in the first time point.

- The learned DBNs consist of two DBNs created through structure learning methods; 4) a learned DBN with "compositional" variables (Model **D**); and 5) a learned DBN without "compositional" variables (i.e., with variables as referenced in the NLST dataset, see Appendix Section A.2) – (Model **E**).

The design process of the models consisted of five steps:

1. **Variable selection**. The structured data captured during the NLST provides a wide array of variables that can be considered in a predictive model. To confine the scope of variables considered, we limited consideration to variables found in previously published studies [22, 23, 25], as well as comorbidities and exposures known to be correlated with lung cancer. Information on family and personal cancer history, and related diseases were represented as "compositional" variables, combining several pieces of evidence into one larger variable. For example, the family history variable is the aggregation of the father, mother, sibling, and child having had cancer. This approach reduces the dimensionality of the associated conditional probability tables (CPTs) in the network. Figure 3.2 depicts all the variables of our models; more information on all the variables used, can be found in Section A.2 of the Appendix. In the case of the learned DBN without "compositional" variables, all the variables shown in Section A.2 of the Appendix are nodes in the network.

2. **Defining the structure (network topology)**.

   - **Defining the structure of the backward construction DBNs**. The *Forward-Arrow* and *Reversed-Arrow* DBNs were constructed using a *backward construction* process, in which we have our variable of interest, in this case lung cancer, and the associated precursors and related contributors to the disease (leftmost part of

40
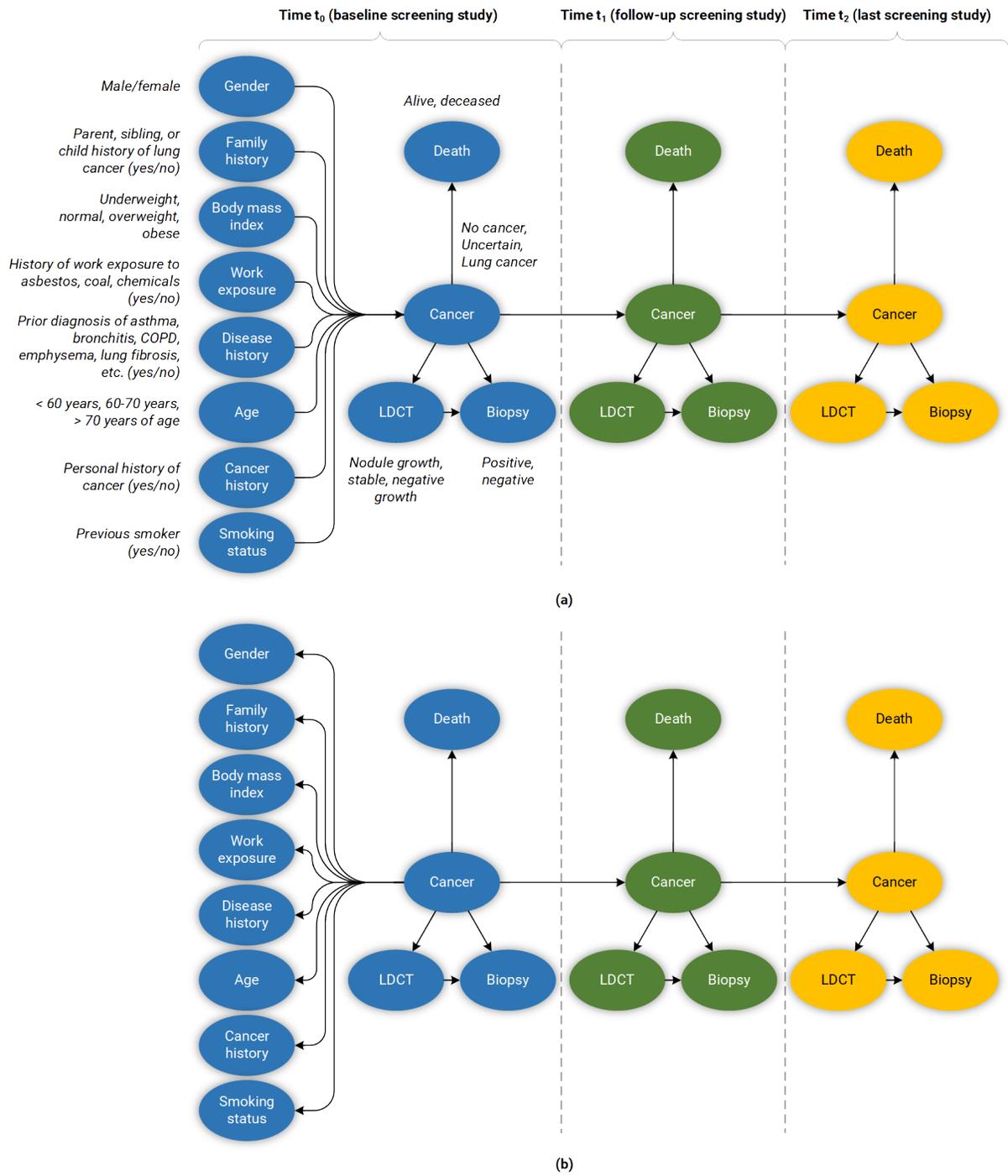
Figure 3.2: The diagram above depicts the structure of the lung cancer screening DBNs. Italicized text indicates the discreted states considered per variable. (a) The *Forward-Arrow* DBNs. (b) The *Reversed-Arrow* DBN. The total number of epochs in both models is 3.

the networks at ($t = 0$), as shown in Figure 3.2 (a)–(b)). The middle and right-most parts of the networks ($t = 1$, $t = 2$) reflect the observations made during screening in the NLST trial. This approach [44] aims to reflect a causal hierarchy for lung cancer screening, in which causes are parents of effects. For example, the evidence of growing abnormalities in an individual's CT screening exam is one of the causes of an individual having a positive biopsy outcome.

- **Defining the structure of the learned DBNs**. The structures of these networks (see Appendix A.2) were learned using the Bayesian search algorithm (see Appendix A.5 Table A.2) provided in Genie [90], enforced with temporal background knowledge. That is to say, we preserved the transition model structure of the DBNs across screenings (e.g., we enforced the fact that the Cancer node at the first screening precedes the Cancer node at the second screening, and that each Cancer node is at least linked to its corresponding LDCT outcome node).

3. **Computing the probabilities**. Given these network topologies, the CPTs and associated probabilities were computed from the observational data of the NLST dataset. The *Forward-Arrow* DBN with a NoisyMax Gate (A), the *Reversed-Arrow* DBN (B) and the learned DBNs (D,E) were parameterized using the expectation maximization (EM) algorithm. The EM algorithm iteratively calculates log-likelihood estimates of the parameters of the network given the data and the structure of the network [91]. For the leftmost part's random variables, such as Gender and BMI, the CPTs represent an estimate of the probability distribution of the variables in the training set. For instance, the CPT for the random variable Gender represents the percentage of females vs. males in the training set. The Cancer node, at baseline, has the most complex CPT table in terms of dimensionality. In the *Forward-Arrow* DBN the number of parameters of the Cancer node at baseline is 2,304. This CPT consists of conditional probabilities that represent the percentage of cases in the training set in one of the three states (NC, U and LC) of the Cancer node and the different combinations of risk factors in

the leftmost part of the network. To deal with this high number of parameters and estimate these parameters from our data, we used a NoisyMax gate to represent the Cancer node. The NoisyMax gate reduced the number of parameters of the Cancer node CPT from 2,304 to 60. NoisyMax, which is a generalization of the NoisyOR gate, can be used to represent more highly connected nodes [92] by taking advantage of the independence of causal interactions to provide a logarithmic reduction in the parameters of a complex CPT. The LDCT CPT represents the percentages of cases in each of the three states NC, U and LC of the Cancer node, with one of the three outcomes (growth, stable, or negative) after their first LDCT screening at baseline. The Biopsy node's probabilities of a positive/negative outcome were abstracted from the literature (i.e., the false negative/positive rate for biopsies) [93]. The Death node represents the death rate of individuals across the whole NLST dataset at the onset of trial. Both the Biopsy and Death nodes in all models were set as fixed nodes (i.e., fixed CPT parameters) during parameterization. The *Forward-Arrow* DBN without a NoisyMax Gate was not parameterized using the EM algorithm. More details regarding the parameterization of this *Forward-Arrow* DBN without a NoisyMax gate can be found in Section A.6.1 of the Appendix.

4. **Computing the probabilities of the transition model**. Our DBN models are not stationary systems. Even though the transition model structure of the networks is repeated over the three time points of the process, the transition models' CPTs change based on the number of cancer cases detected in the NLST dataset annually. For example, the Cancer node at $t = 1$ and $t = 2$ represents the percentage of cases that transitioned from one of the three states at $t = 0$ and $t = 1$ to one of the three states of the Cancer node at $t = 1$ and $t = 2$, respectively. The LDCT nodes' CPTs at $t = 1$ and $t = 2$ represent the percentage of cases in each of the three states NC, U and LC of the Cancer node with one of the three outcomes (growth, stable or negative) after the second and third LDCT screening. The Biopsy and Death node CPTs at $t = 1$ and

$t = 2$ (fixed nodes) are the same as in baseline. Our DBNs were parameterized using the EM algorithm, in a manner akin to a regular Bayesian network (BN) given the way that the growth of nodules were reported in the NLST trial. The reporting of nodule growth in the NLST trial commenced in the second screening period. For example, a suspicious abnormality (>4 mm, considered as a positive finding) that remained stable in size in the second screening was classified as "stable" but if this occurred in the third screening, this abnormality could have been classified as negative. Additionally, during the first screening point all suspicious abnormalities were classified as positive and all non-suspicious abnormalities and negative screenings as negative. There was no reporting of stable cases in the first screening of the trial, as there was no comparison LDCT scan at baseline. This way of abnormality reports was partially continued for a portion of cases in the second screening and eliminated by the third screening of the trial.

5. **Training and testing**. Given a training set with data for each node of our networks, all the models were trained with the Biopsy and Death nodes set as fixed nodes (i.e., fixed CPT parameters). In testing, we had to take into account temporality. We tested each Biopsy node independently and in sequential order. In addition, during testing, instantiating the cancer nodes with evidence would require the individual to undergo additional diagnostic procedures such as a biopsy to confirm their cancer stage. Our classification task was to identify whether individuals should undergo a biopsy given that the positive Biopsy probability is significantly high. This classification was deemed correct if the individual with a high probability of a positive Biopsy had developed cancer and vice versa. Thus, during testing, we did not instantiate any cancer nodes at any screening point of the trial as cancer staging is only validated using additional diagnostic procedures. While this inevitable uncertainty is unfortunate, according to d-separation constraints, it allows the probabilistic influence flow between nodes at any screening point of the trial, for the *Forward-Arrow* DBNs.

3.2.4 Comparison methods

All DBN models were compared with a naïve Bayes model, in which each screening was modeled as independent. Figure A.2 in Section A.5 of the Appendix depicts the structure of the naïve Bayes model. This model was trained using the EM algorithm, and tested in Genie. A logistic regression model (LR) [31] without spiculation, trained and tested on NLST cases at baseline, and a decision tree model were also employed for comparison purposes. The decision tree model was implemented using RapidMiner, which uses a variation of the C4.5 algorithm.

3.3 Evaluation and results

A 10-fold cross-validation was conducted on the complete NLST dataset for each model. The NLST dataset is an imbalanced dataset. The ratio of cancer to non-cancer cases is 1,083:24,461, or around 1 cancer case for every 24 non-cancer cases. As such, imbalance problems arise in classic cross-validation studies: a model trained mainly from negative cases will tend to be inherently biased towards the majority class. Notably, metrics such as the receiver operating characteristic (ROC) curve and the area under the curve (AUC) can be deceiving when training and testing on imbalanced datasets [94]. In our situation, such an evaluation will always have a high accuracy, and thus would not provide insight into whether the model truly identifies cancer cases and how it compares with other models. More informative metrics for imbalanced datasets include precision, recall, and the F-Score [94]. In Section A.8, Figure A.15 of the Appendix we present the F-score over recall curves of the 10-fold corss-validation evaluation of the *Forward-Arrow* DBN model with a NoisyMax gate. The F-score curves improve with additional screenings. However, we note here that we cannot truly evaluate whether our model truly identifies cancer cases, compared with other models over the same dataset, given the large number of non-cancer cases that flatten the F-score curves.
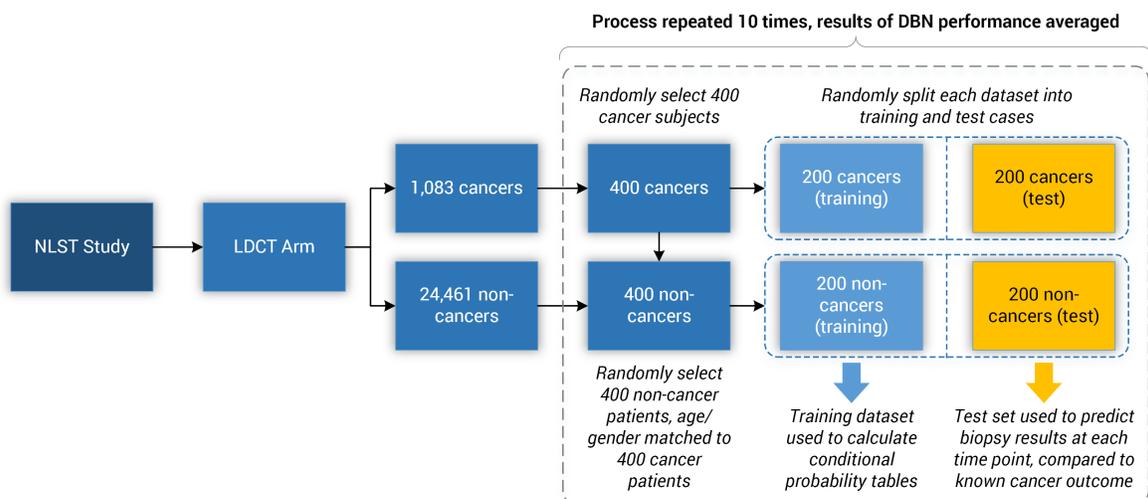
Figure 3.3: The training and testing sets' random selection process of cases from the NLST dataset. The training and test set consist of 200 cancer and 200 non-cancer cases, respectively. Ten random training and test sets, with replacement, were selected for our analysis.

One approach to deal with data imbalance problems is through the use of resampling techniques [95]. In this work, we under-sampled the training and test sets from the majority class (i.e., non-cancer cases) to preserve a 1:1 ratio of the cancer to non-cancer cases. The models were trained and evaluated a total of 10 times. Each time, the training and test sets were randomly selected from the NLST cohort and each consisted of 200 cancer cases and 200 randomly selected non-cancer cases, matched by age and gender. This process was used to assess overfitting and the variability in accuracy of the models, as well as to create a balanced dataset for computing the associated probabilities of a positive Biopsy of an individual. Figure 3.3 illustrates this process. Additionally, the models were tested against the full NLST dataset to assess generalization.

The evaluation of the models was based on the computed probability of the Biopsy variable for a test case, given all prior and current evidence, for each of the three intervention points of the NLST trial. A threshold, θ, was determined for the probability value of Biopsy to indicate a positive biopsy outcome (i.e., probability values below θ were non-cancer cases, values larger or equal to θ were cancer cases). This enabled us to perform a binary classification.
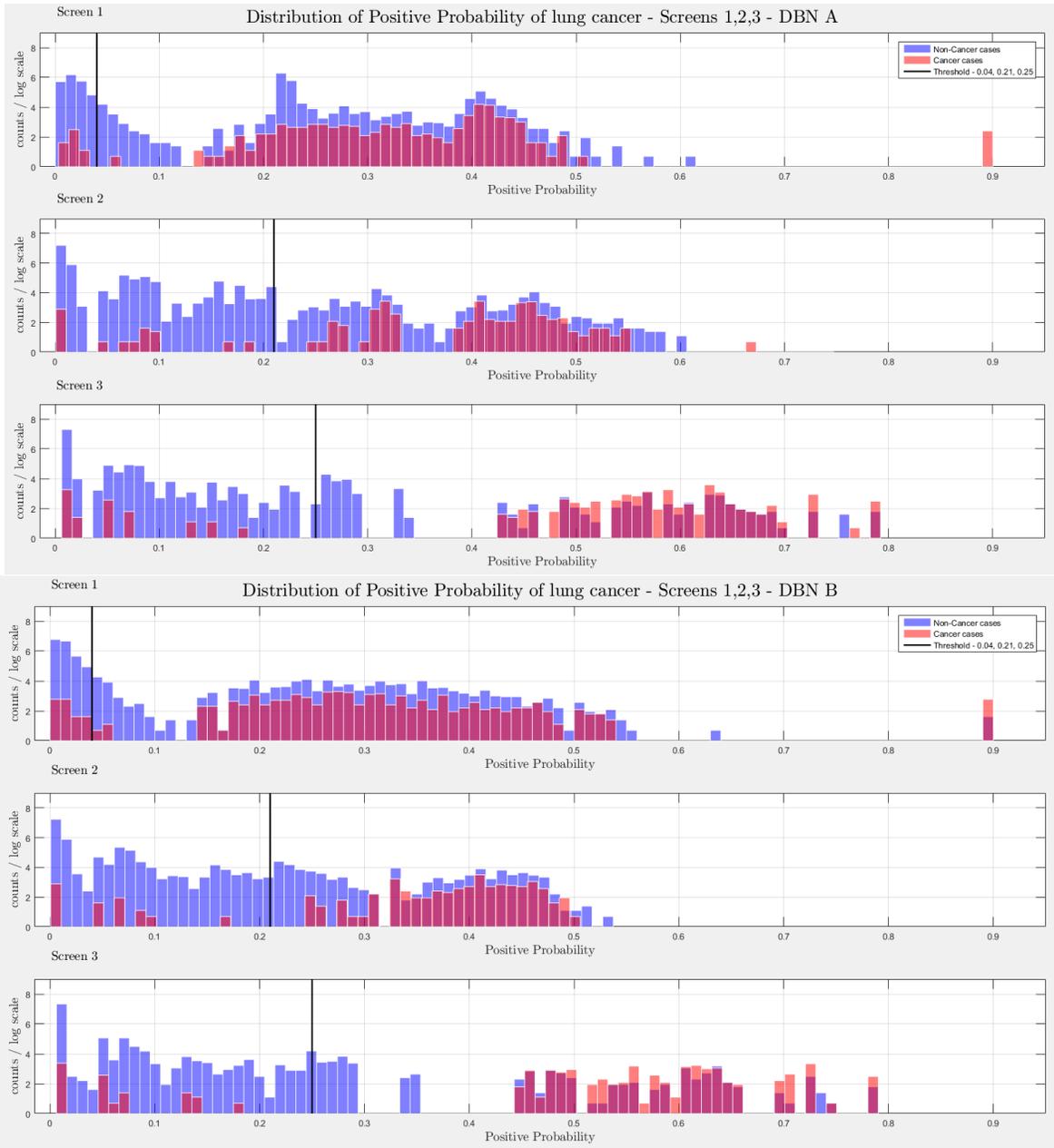
Figure 3.4: The combined probability distributions for a positive biopsy, of DBN A (top) and DBN B (bottom), for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of screening time points. Blue indicates the confirmed non-cancer cases. The three subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.
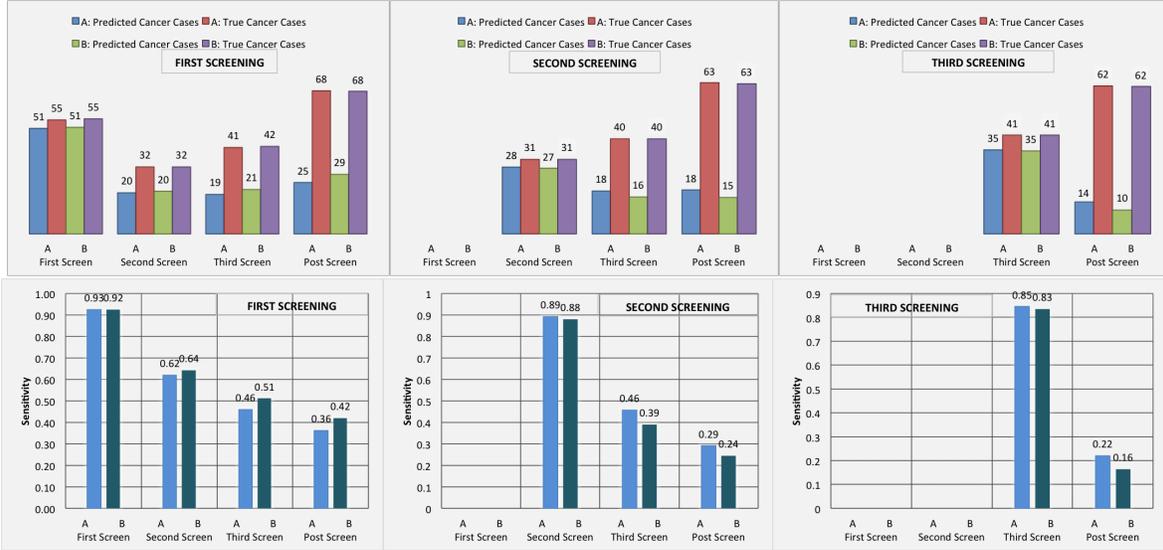
Figure 3.5: **Top**: The diagrams represent the true number of cancer cases in each screening point of the trial and the number of cancer cases predicted by the models in each screening. **Bottom**: (Left) The sensitivity of the lung cancer screening DBNs for the first, second, third, and post-screening cases after the first screening event (baseline). The sensitivities at the second, third, and post-screening cases represent the true positive rate achieved from the pool of false positive cases in the first screen. (Middle) Sensitivity of the DBN for the second, third and post-screening events after the second screening exam. The sensitivities at the third and post-screening cases represent the true positive rate achieved from the pool of false positive cases in the second screen. (Right) The sensitivity of the DBN for the third and post-screening cases after the last screening exam. The sensitivities at the post-screening cases represent the true positive rate achieved from the pool of false positive cases in the third screen.

A positive case prediction by a physician represents any case that resulted in ordering an additional diagnostic procedure. Subsequently, we present for each screen the sensitivity and counts of cancer cases detected by our models at specific thresholds for θ, which were determined based on the distribution of the positive Biopsy probability values (see Figure 3.4), as well as the receiver operating characteristic (ROC) curve.

For discussion purposes we focus our models' comparisons with models A and B. The higher and lower complexity models, in terms of model parameters, respectively. Figure 3.4 depicts the probability of a positive biopsy, as predicted by the models in each screening, of confirmed cancer (red) and non-cancer (blue) cases in the trial. Both DBN A and B tend to discriminate cancer and non-cancer cases better with increasing number of screenings. The

thresholds for θ were chosen in a way that favors recall. For example, each threshold aims to minimize the number of cancer cases missed while preserving an acceptable rate of falsely predicted cancer cases. The results for each of the 10 randomization test sets and resultant models as well as the physicians' predictions were averaged for visualization purposes.
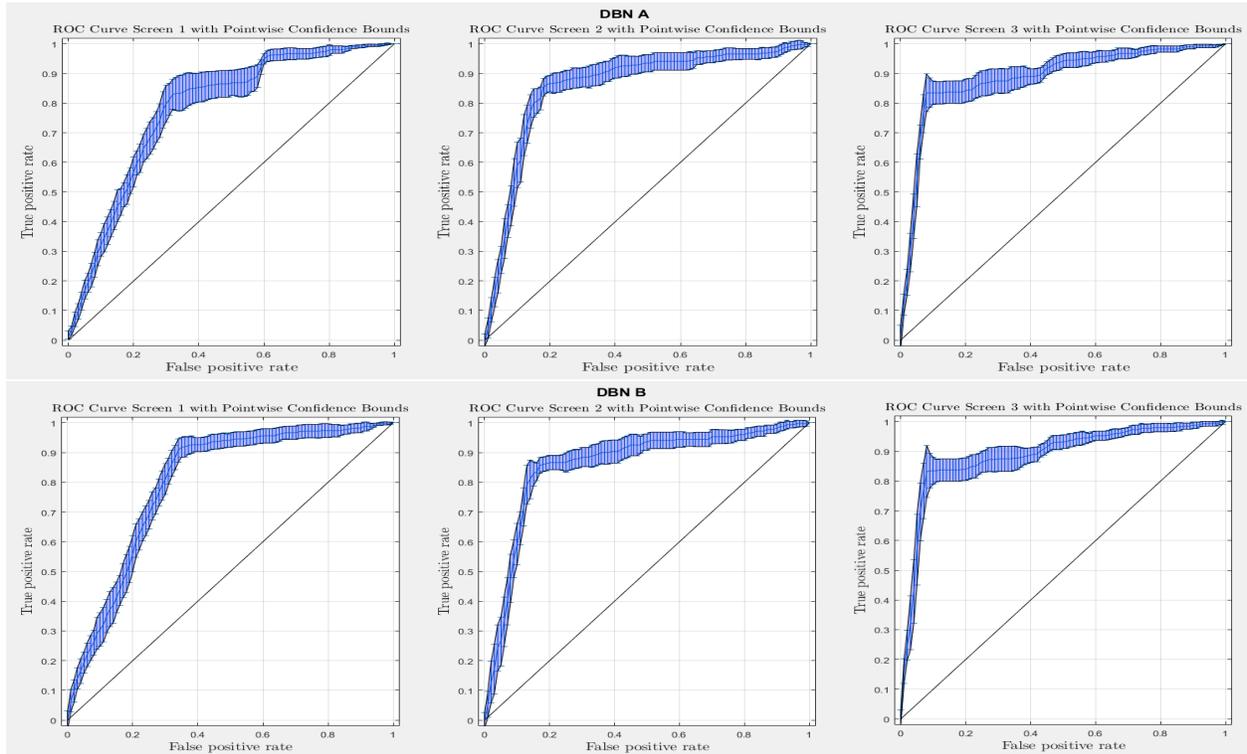
Figure 3.6: The ROC curve of three intervention points of the NLST trial with point-wise 95% confidence bounds.

Comparison with experts. Concordance between the models' positive prediction for a

Model	A		B		C		D		E		F	
	AUC	C.I.	AUC	C.I.	AUC	C.I.	AUC	C.I.	AUC	C.I.	AUC	C.I.
First Screening	0.778	0.757 - 0.800	0.798	0.776 - 0.821	0.789	0.774 - 0.804	0.790	0.769 - 0.810	0.751	0.654 - 0.849	0.799	0.777 - 0.821
Second Screening	0.857	0.834 - 0.880	0.858	0.832 - 0.884	0.844	0.819 - 0.869	0.862	0.839 - 0.886	0.853	0.832 - 0.875	0.865	0.844 - 0.885
Third Screening	0.887	0.869 - 0.905	0.887	0.866 - 0.907	0.884	0.863 - 0.906	0.877	0.858 - 0.896	0.878	0.859 - 0.897	0.886	0.866 - 0.907

Table 3.2: The AUC and the 95% confidence interval for the first, second and third screening. A: The *Forward-Arrow* DBN with a NoisyMax gate; B: The *Reversed-Arrow* DBN; C: The *Forward-Arrow* DBN without a NoisyMax gate; D: The learned DBN with "compositional" variables; E: The learned DBN without "compositional" variables; F: The naïve Bayes Model.

	DBN A				DBN B			
	DBN Predictions By 3rd Screening		DBN Predictions After 3rd Screening		DBN Predictions By 3rd Screening		DBN Predictions After 3rd Screening	
Cases that missed the second screening	8 (tp)	3 (fn)	4 (tp)	3 (fn)	6 (tp)	5 (fn)	4 (tp)	3 (fn)
	91 (fp)	315 (tn)	88 (fp)	311 (tn)	71 (fp)	335 (tn)	67 (fp)	332 (tn)

Table 3.3: **DBN Predictions By 3rd Screening**: Contingency table for the individuals that missed the second screen of the trial and by the third screen were diagnosed with cancer at the $t = 1$ epoch of the DBN. **DBN Predictions After 3rd Screening**: Contingency table for the individuals that missed the second screen of the trial and after the third screen were diagnosed with cancer (i.e., third screening cancer).

biopsy (i.e., $\geq \theta$) and a NLST clinician's recommendation for biopsy and confirmation of lung cancer was determined. The identification of cancer cases was comparable across the three intervention points of the trial for our lung cancer screening DBNs. In terms of the number of predicted cases and discrimination of the same cases, to physicians' performance during the NLST, as shown in Table 3.4. After each screening point, cases that were confirmed as positive lung cancers or deceased were removed in the subsequent screening evaluation. The McNemar's test for each of the contingency tables of similar cases was significant ($p < 0.01$), in each of the three intervention points of the trial, indicating asymmetry. This means that the contingency tables of similar cases are asymmetric and suggests that the models minimize the false negative (fn) rate of cancer cases while maintaining an acceptable false positive (fp) rate. The null hypothesis assesses the equality of proportions of the cases the DBNs and the physicians disagree in classifying as cancer or non-cancer cases. Additionally, the 95% C.I. of the type I and II errors ($b - c$) and of the test of proportions ($p_2 - p_1$) demonstrate that the direction of this asymmetry is toward the fp cases.

Moreover, we examined whether models A and B can predict the majority of cancer cases at a specific screening point of the NLST trial and assessed whether these models could identify cancer cases *before* their occurrence. We evaluated how many of our false positive cases in each screening of the trial turned out to be cancer cases later in the trial. Figure 3.5 illustrates the sensitivity of the lung cancer screening DBNs in each screening, as well as the counts of the predicted number of cancer cases by the models with the total

number of true cancer cases in the trial. Figure 3.5 also illustrates how many false positive cases at a particular screening point of the trial end up being cancer cases in future screening points. For example, in the leftmost histogram for the first screening, DBN A predicted 51 out of 55 cancer cases. From the same screening we examined the false positive cases and identified how many of those cases were cancer cases in subsequent screenings. In the second screening of the trial there were 32 cancer cases. 20 out of those 32 cancer cases were found to be false positive cases in the first screening of the trial. Similarly, in the third screening, 19 out of 41 cancer cases were false positive cases in the first screening. In the post screening 25 out of 68 cancer cases were false positive cases in the first screening of the trial. The middle diagram represents how many cancer cases were identified in the second screening and how many false positive cancer cases in the second screening are cancer cases in the third and post screening cancer cases. The diagram on the right represents how many cancer cases were identified in the third screening and how many false positive cancer cases in the third screening are cancer cases in the post screening cancer cases.

Interestingly, a significant portion of false positive cases are cancer cases in subsequent screenings. Note that confirmed cancer cases from the trial first received a LDCT screening exam, and were then subsequently confirmed through the use of additional diagnostic procedures. In comparison, the DBN models infer that these cases are likely cancer without the diagnostic procedure (i.e., the outcome of a biopsy will likely be positive).

ROC curves with 95% confidence intervals for the first, second, and third screens are shown in Figure 3.6. Table 3.2 summarizes the area under the curve (AUC) for each screen's evaluation and the corresponding confidence interval. The AUC increased with increasing number of screens, which suggests that the models' predictive power improves with time. The AUCs of the *Forward-Arrow* DBN without a NoisyMax gate, the two learned DBNs and the naïve Bayes model are similar to DBN A and B and can be found in Table 3.2. More details on the results of the evaluation of each model are provided in Section A.6 of the Appendix. Overall, all models have similar AUCs and confidence interval (C.I.) of the

51

AUC for each screening. The learned DBNs have similar performance to all models except the AUC and C.I. of the AUC for the first screening of model E, which is lower and higher, respectively, compared with the other models. In addition, as shown by the NLST and the models themselves, performance is improved with consecutive screens. This is evident both from Table 3.4 as well as the precision/recall (PR) and F-score curves (see Appendix A.9 and A.8) computed for each screening time point. The desirable performance of PR and F-score curves is to be in the upper-right-hand corner. The PR and F-score curves in Appendix A.9 and A.8 tend to move towards the upper-right-hand corner with increasing number of screenings. Models A-E achieved the best PR curves across screenings with PR curves improving with increasing number of screenings. The worst PR curves, which are in the bottom-left-hand corner, are the naïve Bayes model (see Appendix A.9). The naïve Bayes PR curves get worse with increasing number of screenings, indicating overfitting to specific features, such as the LDCT outcome. We have also tested the performance of a decision tree on the dataset, using a variation of the C4.5 algorithm. The decision tree performance was extremely low compared to the other models and is not reported.

The models' predictive power was also assessed by investigating the number of *future* cancer cases predicted by the models using only observations from one screening. For example, if we were testing for cancer cases at $t = 0$ (first screening) we assumed that all cancer cases at $t > 0$ were cancer cases at $t = 0$ (i.e., ignored time). In this way, we can evaluate how many cancer cases are predicted before incidence. Out of the 121 true positive cases detected by DBN B on the first screening (see Table A.1 of the Appendix), given that the DBN predicted 51 cancer cases that were cancer cases of the first screening (see Figure 3.5 - top left), the DBN predicted 70 additional cancer cases that were diagnosed with cancer later in the trial (see Table A.1 in Appendix A.3).

	DBN A Predictions		DBN B Predictions		Physicians' Predictions		DBN A & Physicians' Concurrence		DBN B & Physicians' Concurrence		McNemar's Test A	McNemar's Test B	95% C.I. $(b-c)$ A	95% C.I. $(b-c)$ B	95% C.I. $(p_2 - p_1)$ A	95% C.I. $(p_2 - p_1)$ B
First Screening	53 (tp)	2 (fn)	51 (tp)	4 (fn)	49 (tp)	6 (fn)	70.4% (tp)	35.2% (fn)	71.4% (tp)	48.7% (fn)	$x^2 = 91.03$ $p < 1.0e^{-10}$	$x^2 = 77.39$ $p < 1.0e^{-10}$	b c b - c = 94.5 (77.5,110.57)	b c b - c = 82.1 (64.87,98.69)	$p_1 = \frac{a+c}{N}$ $p_2 = \frac{a+b}{N}$ $p_2 - p_1 = 0.3778$ (0.3099,0.4421)	$p_1 = \frac{a+c}{N}$ $p_2 = \frac{a+b}{N}$ $p_2 - p_1 = 0.2655$ (0.2098,0.3192)
	221 (fp)	121 (tn)	134 (fp)	208 (tn)	108 (fp)	235 (tn)	49.3% (fp)	46.0% (tn)	59.1% (fp)	73.3% (tn)						
Second Screening	27 (tp)	4 (fn)	27 (tp)	4 (fn)	29 (tp)	2 (fn)	69.0% (tp)	30.5% (fn)	68.7% (tp)	29.0% (fn)	$x^2 = 26.05$ $p = 3.28e^{-7}$	$x^2 = 25.95$ $p = 3.5e^{-7}$	b c b - c = 28.6 (16.45,39.52)	b c b - c = 28.5 (16.42,39.52)	$p_1 = \frac{a+c}{N}$ $p_2 = \frac{a+b}{N}$ $p_2 - p_1 = 0.1092$ (0.0628,0.1509)	$p_1 = \frac{a+c}{N}$ $p_2 = \frac{a+b}{N}$ $p_2 - p_1 = 0.1088$ (0.0627,0.1509)
	50 (fp)	244 (tn)	50 (fp)	244 (tn)	61 (fp)	233 (tn)	39.3% (fp)	78.9% (tn)	39.3% (fp)	79.0% (tn)						
Third Screening	35 (tp)	7 (fn)	35 (tp)	7 (fn)	37 (tp)	4 (fn)	71.0% (tp)	46.0% (fn)	71.0% (tp)	46.0% (fn)	$x^2 = 8.45$ $p = 0.0036$	$x^2 = 8.45$ $p = 0.0036$	b c b - c = 12.3 (2.83,21.18)	b c b - c =12.7 (3.61,22.34)	$p_1 = \frac{a+c}{N}$ $p_2 = \frac{a+b}{N}$ $p_2 - p_1 = 0.0492$ (0.0113,0.0847)	$p_1 = \frac{a+c}{N}$ $p_2 = \frac{a+b}{N}$ $p_2 - p_1 = 0.0507$ (0.0144,0.0892)
	24 (fp)	227 (tn)	24 (fp)	227 (tn)	32 (fp)	219 (tn)	41.6% (fp)	85.8% (tn)	41.6% (fp)	85.8% (tn)						

Table 3.4: The results of the lung screening DBNs A and B and the Physicians of the NLST Trial for the first, second and third screening, as well as the percentages of equivalent predictions of true positive (tps), false negatives (fns), false positives (fps) and true negatives (tns). N is the total count of cases $(a + b + c + d)$. b and c represent the type I and type II errors of the contingency matrix. Deceased and already identified cancer cases before each intervention point of the trial were excluded from the evaluation of the DBNs as well as the evaluation of the physicians predictions. **DBN A and B Predictions**: The contingency table that depicts the predictions of the lung screening DBNs for each screening at a threshold of 0.04, 0.21 and 0.25, respectively. **Physicians' Predictions**: The contingency table that depicts the predictions of the Physicians in the trial. **DBNs A and B and Physicians Concurrence**: The percentage of equivalent predictions of tps, fns, fps and tns. For example, the percentage of tps, fns, fps and tns represents how many number of cases where equivalently predicted (i.e., if the same cases are predicted by both the DBN and physicians) as tp, fn, fp and tn by the DBN and the Physicians over the total number of tps, fns, fps and tns, respectively. **McNemar's Test**: The chi-square and p-value of the McNemar's test for the contingency matrix of similar cases identified by the models and the physicians. **95% C.I. (b - c)**: Confidence Interval of the difference of type I and type II errors of the concordance matrix. **95% C.I. ($\mathbf{p_2 - p_1}$)**: Confidence Interval of the difference of proportions of the contingency matrix of similar cases.

Assessing model performance given missing data. We grouped all cases in the study that missed the second screen in the NLST, but underwent the first and third screens. There were 417 such cases in the complete NLST dataset, which we used to evaluate whether the models could predict the cancer status (e.g., cancer or non-cancer) of an individual that missed the second LDCT screening exam and was subsequently screened at the third screen. Table 3.3 provides the contingency tables for these cases that went on to develop cancer by the third screening or after the third screening. DBN A and DBN B managed to predict 8 and 6 out of the 11 cases, respectively, that developed lung cancer by the third screening, and both the DBNs predicted 4 out of 7 cases that developed cancer after the third screening.

The NLST dataset is complete in terms of patient information (i.e, parent nodes). To evaluate the effect of missing data on the parent nodes in the training set and the end performance of the *Forward-Arrow* DBN without a NoisyMax gate we randomly selected parent nodes and assigned missing data to each one to simulate a "missing at random" scenario. For example, we selected one random parent node and set 50 random cases with missing values for that node. We repeated this in incremental steps of 50 cases up to 350 (our training set consisted of 400 cases). We then reiterated the process with two random parents, increasing up to all parent nodes. Our results showed that the AUC and the confidence interval of the AUC remained relatively stable. Changing the distribution of these priors does not significantly affect performance. The highest impact on performance of the AUC, which was of the order of -0.01, was on the first screening. This subtle change may be attributed to the fact that biopsy and cancer nodes of the first screening are conditionally dependent on the priors. A strength of influence diagram of each structure depicting the influence amongst variables in each network is provided in Appendix E.

Generalization and comparison to other models. We assessed the generalizability/overfitting of the models on the whole NLST dataset. Table 3.5 depicts that the true positive (tp), false negative (fn), false potive (fp) and true negative (tn) rates of the model over the whole dataset and the random balanced test sets appear stable. The number of

	First Screening		Second Screening		Third Screening	
DBN A, whole dataset	92.6% (tp)	7.40% (fn)	87.3% (tp)	12.7% (fn)	83.9% (tp)	16.1% (fn)
	30.2% (fp)	69.8% (tn)	9.40% (fp)	90.6% (tn)	6.70% (fp)	93.3% (tn)
DBN A, random test sets	96.4% (tp)	3.60% (fn)	87.1% (tp)	12.9% (fn)	83.3% (tp)	16.7% (fn)
	65.6% (fp)	35.4% (tn)	17.0% (fp)	83.0% (tn)	9.60% (fp)	90.4% (tn)
DBN B, whole dataset	92.6% (tp)	7.40% (fn)	87.3% (tp)	12.7% (fn)	83.9% (tp)	16.1% (fn)
	30.2% (fp)	69.8% (tn)	9.40% (fp)	90.6% (tn)	6.9% (fp)	93.1% (tn)
DBN B, random test sets	92.7% (tp)	7.30% (fn)	87.1% (tp)	12.9% (fn)	83.3% (tp)	16.7% (fn)
	39.2% (fp)	60.8% (tn)	17.0% (fp)	83.0% (tn)	9.60% (fp)	90.4% (tn)
Physicians, whole dataset	89.7% (tp)	10.3% (fn)	93.7% (tp)	6.30% (fn)	90.1% (tp)	9.90% (fn)
	23.2% (fp)	76.8% (tn)	15.0% (fp)	85.0% (tn)	9.60% (fp)	90.4% (tn)
Physicians, random test sets	89.1% (tp)	10.9% (fn)	93.6% (tp)	6.40% (fn)	90.2% (tp)	9.80% (fn)
	31.5% (fp)	68.5% (tn)	20.7% (fp)	79.3% (tn)	12.8% (fp)	87.3% (tn)
Logistic Regression [31]	11.0% (tp)	89.0% (fn)	-	-	-	-
	0.50% (fp)	99.5% (tn)	-	-	-	-
Logistic Regression, retrained	16.3% (tp)	83.7% (fn)	-	-	-	-
	0.8% (fp)	99.2% (tn)	-	-	-	-

Table 3.5: The true positive (tp), false negative (fn), false potive (fp) and true negative (tn) rates of the DBN A and DBN B trained on 400 cases and evaluated on the remaining NLST dataset of $25,446$ cases and 400 random balanced test cases. The tp, fn, fp and tn rates of the physicians classifications on the entire NLST dataset of $25,446$ cases and the random sets of 400 cases. **LR model**: The full LR model without spiculation [31] externally validated on $5,353$ cases with nodule information of the NLST at baseline ($t = 0$) and the same model trained on $2,663$ cases and tested on $2,690$ cases.

test cases in the whole dataset and in the random balanced test set are 25446 and 400, respectively. Lastly, we compared how the full logistic regression model (LR) of [31] without spiculation performs on the NLST cases at baseline. We first evaluated how the LR performs on the NLST cases when trained with NLST cases and we also evaluated how the parameterized model, with parameters published in [31], performs on the NLST cases. In both cases, compared to the DBN, the LR maintains a high true negative rate, a high false negative rate, and a significantly lower true positive rate (see Table 3.5). The LR models were evaluated only on baseline as they were trained and evaluated in [31].

3.4 Discussion

In this work we built and tested five different DBNs for lung cancer screening prediction using backward construction and structure learning methods. Given the uncertain nature of lung cancer and the necessity to perform a biopsy to confirm the underlying disease we used a three-state cancer state-space model to represent the cancer status of an individual along the screening process. Such a representation offers the following advantages. First, it represents the cancer state of an individual in terms of cancer staging that captures concepts like disease dynamics and nodule growth, instead of the standard binary "yes" and "no" states. Second, the fact that the cancer nodes are never instantiated with evidence due to the uncertainty of the disease during testing (i.e., cancer staging is only validated using additional diagnostic procedures) allows the flow of probabilistic influence of demographic characteristics as well as previous screening outcomes on any screening point of the trial (i.e., via d-separation and sequential configuration). The performance of the learned DBNs is similar to that of the *Forward-Arrow* and *Reversed-Arrow* DBNs. The results of the learned structures demonstrate similar relationships to those in the expert-driven *Forward-Arrow* models with respect to the imaging assessment over time (see Figure A.2 in Appendix A.5); additional relationships were inferred, but without significant change in model performance. Qualitatively, the expert-driven models provide a more straightforward understanding of the relationship between variables over time. Markedly, the NLST trial patient information (e.g., demographics) was captured only at the start of the trial. While some measures are typically invariant over time (e.g., gender), various measures do change over time (e.g., age, body mass index). The underlying dataset did not have these latter variables reflected in subsequent time points in the screening process. In our opinion, it would be inaccurate to model them as such (and the imaging interpretations were also not informed by any such additional information). Nevertheless, given such data at different time points, the performance of the DBNs could improve with additional modeling.

Based on the results of our evaluation, DBN A and B provide results comparable to

the radiologists who participated in and read the NLST LDCT imaging studies. We also tested other models on this dataset such as decision trees and a naïve Bayes model, but their performance was suboptimal compared to the DBNs. The use of a DBN for our analysis rather than a BN network as in [57,58] takes into account the temporal evolution of a cancer, with improved performance in the discriminative ability of the model in future screenings. A standard 10-fold cross-validation method on the entire dataset would be ideal to assess overfitting. But given the class imbalance present in the dataset (1:24 cancer to non-cancer cases), we would not gain insight into the models' ability for the more important predictive classification of cancer. [95] used similar methods to deal with imbalance in their dataset, but instead chose to oversample the minority class until a 1:1 ratio was achieved in their training set. They also reported metrics such as precision, recall, and F-score to compare performance against imbalanced datasets. The AUC for all networks remained higher than 0.75 in the balanced test sets across the three screening points of the trial, and the AUC curves improve over time. The use of balanced test sets allows the effective comparison of each model in the ROC and PR space over the cancer class. We can see that all models' performance were comparable in the ROC space (AUC of the ROC). However, in the PR space we also see that all models have a clear advantage over the naïve Bayes model (see Appendix A.9 Figures A.17 – A.23). This model adjusts to very specific features, such as the LDCT nodes, and thus overfits its predictions on these features. It can accurately discriminate negative cases (comparable AUC to other models); but when asked for the probability of a real cancer case given that this cancer case is predicted by the model (PR curve), its performance is lower.

Models A and B were also able to identify a significant number of cases at each intervention point of the trial that were future cancer cases (see Appendix A.3). The Brier score as well as the calibration curves of DBNs A and B improve with the increasing number of screenings (see Appendix A.4), demonstrating the ability of the models to perform calibrated cancer incidence predictions over time. Interestingly, the lung screening DBNs A and B only require a small training set, on the order of 50 times smaller than the original dataset, to

make predictions on a large number of cases they have never encountered before. The models demonstrate good discrimination when evaluated on the whole NLST dataset. In addition, the tp, fp, fn and tn rates over the whole dataset compared to the random balance test sets are consistent and in some cases better. Still, it is important to note that in this study the DBNs were developed and trained using data from a randomized controlled trial, where information was gathered in structured case report forms and a large degree of standardization took place. Despite the performance over the entire NLST dataset, real-world application of these DBNs will require adaptation to handle observations made from routine clinical screening processes (i.e., adjusting for "noise" and variance). Ultimately, external validation of the DBN is required.

DBNs present certain advantages regarding lung cancer incidence prediction, including their ability to utilize datasets with missing data. Although the NLST dataset is from a controlled trial, and thus is largely complete with only some missing data (e.g., due to individuals missing a screening exam), our models appear to be robust against missing values and still make reasonable predictions in light of missing data. In our investigation of the cancer status of cases that missed only the second NLST screening, both DBN A and B predicted the majority of cases that were cancer cases by the third screening or after the third screening of the trial. Suggesting that certain lung cancer risk factors and the outcome of the first LDCT are sufficient for an accurate future prediction of cancer. This short-term predictive ability may be applicable in cases where missing a screening exam would result in symptomatic cancer. Cases with missing data were also used in the training phase of the DBN without affecting the models' predictive ability. We can improve the parameterization of a model from cases with incomplete data by only using the information we do have for each case, with incomplete data, for the computation of the corresponding CPT tables of the DBN network. For example, cases that developed lung cancer at the baseline of the trial before they received their first screening exam, even though we do not have information about them after baseline, were still used in the computation of the

baseline CPTs (e.g., Gender, Age). To match a real lung cancer screening setting we included all of the aforementioned cases in our evaluation. We used the EM algorithm to train the *Forward-Arrow* DBN with a NoisyMax gate, the *Reversed-Arrow* DBN, and both the learned DBNs. One advantage of the EM algorithm is its ability to estimate the parameters of a network using the observed data. In particular, it iteratively fills in missing values with estimated values and subsequently re-estimates the parameters from this complete dataset. We believe it would be inappropriate to estimate the disease status of a deceased individual in subsequent screenings as individuals who died during the course of the trial, or who were diagnosed with cancer, were removed from the screening process of the trial. Thus, in the *Forward-Arrow* DBN without a NoisyMax gate, we estimated the parameters of this network empirically from observations in the dataset. Interestingly, both techniques provide similar results (see Appendix A.6). As such, EM would be a more appropriate algorithm in cases that missed a screening exam but is unsuitable with participants who were diagnosed with cancer or who died during the course of the trial. A method that takes into account both types of missing data would be more appropriate in eliminating bias during training.

When compared with the full logistic regression model without spiculation [31] the Lung Screening DBNs had better tp, fp and fn rates. This suggests a superior discriminatory power on the NLST dataset. Nevertheless, the LR model's results in Table 3.5 are trained and tested on a specific portion of the dataset: individuals with reported nodule abnormalities and nodule consistency. The DBN models, in contrast, were trained on a balanced set of cancer cases and non-cancer cases, with the majority of non-cancer cases without abnormalities. Also, the classification task of each model is somewhat different. For example, our DBN models identify lung cancer individuals whereas the LR model identifies cancerous nodules. Further investigation and standardization of the dataset and the classification task of the different types of models would be more appropriate for such a comparison. But similar to other models, baseline information on smoking status, demographics, health status, history of cancer, and exposure risk factors were employed as inputs. However, we did not use

quantitative imaging information. McWilliams *et al.* [31] utilized the maximum nodule size, the type of nodule, and the number of nodules per CT scan, resulting in a parsimonious multivariate logistic regression model. Their models achieved an AUC higher than 0.90. In this study, we did not explicitly use nodule characteristics in our analysis, but rather included the interpretation of the LDCT by the radiologists, which was based on nodules' overall growth between consecutive screening exams. We speculate that a nodule's rate of growth is a significant predictor of lung cancer, as all our models and physicians' predictions improve given the progression of information. An exploration of how much "history" is needed in terms of interpretation and predictive power is also required: it may be that in this domain, only the past n years of observation are required (rather than the entire longitudinal history). The NLST only provided three time points, so it is not possible to ascertain what amount of information would be optimal for temporal analysis of lung cancer screening data. The use of nodule features such as consistency, location, and size would be strong predictors of lung cancer [96] and will be included in subsequent biterations of our model in combination with automated segmentation methods [97] to automatically provide additional evidence for predicting diagnoses.

We recognize that there are some limitations to this work. For example, the screenings received by the individuals in the NLST were not exactly at the same three discrete time points; (on the contrary they had a continuous nature as individuals received their screenings at different days). Given the nature of real-world implementation of lung screening programs, it is unlikely that a fixed time frequency of observation will occur, for any number of reasons. As such, a DBN may ultimately not be well-suited to handle longer sequences of observation and clinical decision-making. Alternative continuous time temporal models will be explored as part of our future work. Also, the thresholds used in this work were selected to favor recall, providing a conservative prediction that would err on the side of detecting a cancer, rather than missing a cancer case. Thus, the optimal threshold was considered to be one that minimized the number of cancer cases while having an acceptable false positive rate.

The use of threshold-determining methods that take into consideration factors such as utility of life and monetary costs will be looked at in the future.

3.5 Summary of findings

In this chapter, we explored five DBNs for lung cancer screening constructed using the NLST dataset. We demonstrated the challenges in providing screening recommendations using a DBN. We dealt with data imbalance and introduced a training and testing procedure for DBNs in uncertain diseases, such as cancer that uses a hidden cancer node, during testing, built on a cancer staging state-space model. Parameter reduction methods and the EM algorithm for parameterization with missing data were also explored. The DBNs aim to identify individuals who will go on to develop lung cancer based on data collected at baseline and radiologist interpretation in sequential (annual) imaging exams. All models achieved high AUC scores across all three screening points of the NLST, demonstrating comparable performance to the experts. As may be expected, the DBNs performance improved over time, as more information about the history of the patient unfolded. Additionally, the models ability to predict future cancer cases in advance was also examined, finding that they were able to identify some cases before the expert (i.e., cases that were deemed false positives by a radiologist, but that in later studies, proved to be cancer). This work is the first step in understanding how we may subsequently tailor the lung cancer screening process to optimize early detection while minimizing false positive findings.

Chapter 4

Generating Reward Functions using IRL Towards Individualized Cancer Screening

4.1 Overview

In this chapter we explore ways to learn one of the fundamental components of a POMDP model from data, the reward function. The reward function embeds knowledge about the problem and the desired goal; in the context of an MDP, it induces the desired behavior for an agent (i.e., optimal policy). Markedly, POMDP models used in medicine typically use a reward function adopted from cost-effectiveness studies [11, 13, 15, 98] or are posed in terms of quality-adjusted life years (QALYs). While such functions are informative about general populations, they do not necessarily reflect how an experienced clinician would make a decision, especially given a specific individual's medical history and preferences. Indeed, little work has been done in designing reward functions that emulate experts' decision processes.

We use Maximum Entropy Inverse Reinforcement Learning (MaxEnt IRL) algorithm [76] to establish reward functions from retrospective screening data, to learn how an expert physician may select a given action based on observed test results. An adaptive step size to expedite the convergence rate of MaxEnt IRL is also demonstrated. Importantly, we present

how to use the MaxEnt IRL learned rewards to generate state-action pair rewards that can be used in POMDPs. This methodology is applied on two real-world clinical datasets for lung and breast cancer screening, mimicking how clinicians made decisions regarding patients. A comparison of the resultant POMDP policies using the MaxEnt IRL reward functions against experts' actions is presented. This chapter's main highlight is the use of the MaxEnt IRL algorithm as an efficient and accurate method in estimating sensible reward functions for cancer screening MDPs or POMDPs.

4.2   Materials and Methods

4.2.1   NLST Dataset

For this work, we used data from the NLST's LDCT arm, comprising approximately 25,500 participants that underwent three annual screenings and follow-up post screening. We further filter this dataset to those subjects who had a reported pulmonary nodule based on imaging. Unfortunately, preprocessing of the NLST data is not straightforward, as longitudinal tracking of the nodules was not considered at the time of the study. Thus, to use imaging-related information, we made the assumption that an imaging finding in individuals with only one reported nodule and in the same anatomical location over time is the same nodule across the three screening points of the trial. This criterion further constrained our dataset to 5,402 LDCT subjects. Figure 5.4 shows the number of patients and total number of nodules reported. Unfortunately, NLST annotation data does not provide a means for tracking individual nodules in participants with multiple nodules. Given this dataset artifact and to ensure proper characterization of changes in imaging-based features over time (e.g., nodule size, consistency), we further constrained our data to individuals with only one IPN reported in the same anatomical lung lobe during the study, assuming that the same nodule was observed over time. This selection criteria and preprocessing to remove inconsistent cases (see below) resulted in a total of 5,402 cases, which we used to train and test

	Lung screening		Breast screening	
	Cancer	Non-cancer	Cancer	Non-cancer
First screening	162	5,240	370	3,636
Second screening	85	5,155	85	3,551
Third screening	107	5,048	34	3,517
Post/Fourth screening	98	4,950	6	3,511

Table 4.1: The number of cancer and non-cancer cases in the lung and breast cancer dataset for each screening point.

our POMDP model. From this subgroup, we learned a reward function, then trained and tested a POMDP. Note that for the reward function we made use of the recorded diagnostic follow-up variables (e.g., recommendation for other procedures) to inform actions. A detailed description of the NLST dataset as well as pre-processing steps are described in Tables 4.1, 4.2 and in section 4.2.2.

4.2.2 Data preprocessing

Table 4.2 summarizes the NLST variables used in our analysis. We considered the same demographic and clinical variables selected in the Tammemägi model [99] and replicated its preprocessing steps. We converted two variables into binary representations: family history of lung cancer (if any first degree relative had a history of lung cancer) and personal lung cancer history (if the individual had any prior history of lung cancer). Missing values for the variables used with the Tammemägi model were imputed using a variation of a multiple clustering imputation approach [100]. In addition to the radiologist's overall interpretation of the LDCT scan, we employed several imaging features describing the nodule, discretizing continuous variables: location; nodule size (Bin 1: ≥ 0 mm and ≤ 3 mm; Bins 2-9: 1 mm bins from $3-11$ mm; Bin 10: > 11 mm and ≤ 27 mm; and Bin 11: > 27 mm); predominant attenuation, and margins. Given the sparsity of cases with nodules of size > 11 mm, we created larger bins by identifying discretizations maximizing POMDP performance using the training data. We removed inconsistent cases with a perpendicular measurement greater than the reported longest nodule diameter and any cases with missing measurements. Cases

Demographic, clinical, outcome variables (% missing)	Variable type	Value	Mean (SD)/Category proportions (%)
Age (0%)	Continuous		61.64 (5.05) years
Education (0.31%)	Categorical	8th grade or less	1.34%
		9-11th grade	4.66%
		High school graduate/GED	24.29%
		Post-high school training, excluding college	13.72%
		Associate degree/some college	23.53%
		Bachelors degree	16.47%
		Graduate/professional school	14.09%
		Other	1.89%
Race (1.48%)	Categorical	White	93%
		Black	4.28%
		Asian	2.05%
		American Indian or Alaskan Native	0.23%
		Native Hawaiian or Other Pacific Islander	0.26%
Body mass index (0.03%)	Continuous		27.61 (4.92)
Chronic obstructive pulmonary disease (COPD) (0.26%)	Binary	No	94.38%
		Yes	5.62%
Family history of lung cancer (0%)	Binary	No	77.32%
		Yes	22.68%
Personal history of lung cancer (0%)	Binary	No	95.59%
		Yes	4.41%
Smoking status (0%)	Binary	No	50.19%
		Yes	49.81%
Smoking intensity (0%)	Continuous		28.71 (11.43)
Duration of smoking (0%)	Continuous		40.2 (7.27) years
Smoking quit time (0.48%)	Continuous		3.67 (4.95) years
Confirmed lung cancer diagnosis (0%)	Binary	No	91.65%
		Yes	8.35%

Study variable	Variable type	Value	t_0	t_1	t_2
Screening outcome (radiologist interpretation)	Categorical	Negative screen, no significant abnormalities	8.83%	4.59%	4.15%
		Negative screen, minor abnormalities not suspicious for lung cancer	24.07%	25.84%	49.09%
		Negative screen, significant abnormalities not suspicious for lung cancer	6.42%	3.74%	4.41%
		Positive, change unspecified, nodule(s) $\geq$ 4 mm or enlarging nodule(s)	60.16%	7.09%	0%
		Positive, no significant change, stable abnormalities	0%	36.82%	17.66%
		Positive, other	0%	13.44%	13.35%
		Not compliant, left study	0%	0.39%	0.68%
		Not compliant, refused a screen	0.35%	4.46%	4.92%
		Not compliant, wrong screen	0.15%	0%	0.02%
		Not compliant, erroneous report of lung cancer before screen (LSS only)	0%	0%	0.04%
		Not compliant, form not submitted, window closed	0%	0.07%	0.11%
		Not expected, cancer before screening window	0%	2.78%	4.09%
		Not expected, cancer in screening window	0%	0.04%	0.15%
		Not expected, death before screening window	0%	0.46%	1.07%
		Not expected, death in screening window	0%	0.28%	0.26%

Nodule variables (% missing)	Variable type	Value	t_0	t_1	t_2
Location (t_0 =40.00%,t_1 =41.58%,t_2 =40.93%)	Categorical	Right upper lobe	24.44%	23.92%	22.78%
		Right middle lobe	13.21%	13.97%	13.22%
		Right lower lobe	23.60%	23.89%	24.16%
		Left upper lobe	13.79%	12.77%	13.51%
		Lingula	4.01%	3.64%	3.79%
		Left lower lobe	20.58%	21.1%	21.53%
		Other	0.37%	0.70%	1.00%
Margins (t_0 =40.00%,t_1 =40.30%,t_2 =40.93%)	Categorical	Spiculated (stellate)	12.5%	9.05%	7.62%
		Smooth	62.97%	67.97%	70.79%
		Poorly defined	18.54%	19.01%	18.8%
		Unable to determine	5.99%	3.97%	2.79%
Longest diameter (t_0 =40.00%,t_1 =41.58%,t_2 =40.93%)	Continuous		7.97 (7.00) mm	7.23 (5.19) mm	7.07 (5.50) mm

Diagnostic intervention variables	Variable type	Value	t_0	t_1	t_2
Biopsy	Binary	No	95.17%	97.6%	97.28%
		Yes	4.83%	2.37%	2.72%
Invasive procedure	Binary	No	94.95%	97.57%	97.17%
		Yes	5.05%	2.43%	2.83%
Non-invasive procedure	Binary	No	46.91%	68.72%	80.91%
		Yes	53.09%	31.28%	19.09%

Table 4.2: Variables used for the development and evaluation of the lung cancer screening POMDP model. After applying selection criteria and removing subjects with inconsistent values, a total of 5,402 LDCT screening cases were used from the NLST. Percentages of missing data are provided, alongside categorical breakdowns and mean values.

without screening abnormalities at an annual screening for reasons other than death, cancer, or missed screening, a nodule size between 0-3 mm was assumed. Cases without nodule size

abnormalities across the three annual screenings were excluded from the analysis. Nodule size was then interpolated between annual screenings using the average value between time points, with nodule consistency, margins, and follow-up decisions unchanged relative to the earlier annual observation. Other variables used in the model include the total number of screening days, occurrence of diagnostic procedures (biopsy, thoracotomy, diagnostic CT exam), and confirmed diagnoses of lung cancer.

4.2.3   Athena Dataset

The Athena Breast Health Network [17] is a University of California (UC)-wide initiative around breast cancer screening and treatment. The effort started in 2009 and includes women who underwent breast screening at five academic medical centers. The portion available at our institution (UCLA) consists of 49,244 patients, with follow-ups of up to 4.8 years; this subset represents 96,515 screening and diagnostic mammograms (MGs), and 2,713 diagnostic biopsies. MG results are reported as Breast Imaging Reporting and Data System (BI-RADS) scores [101]. We selected patients with initial risk (Gail) scores, four consecutive screenings, valid BI-RADS scores, and biopsies results per breast side (i.e., left, right). 2,095 patients with left breast MGs and 2,036 patients with right breast MGs (4,131 total cases, 4,099 after pre-processing) were used in this study. A description of the number of cancer cases over time for the Athena dataset is shown in Table 4.1. At the time of their breast screening exam, women were asked to complete a questionnaire that collected basic demographics and risk factors related to the Gail model. The purpose of the survey was to obtain all the information necessary to provide a risk estimate using the model and to provide radiologists contextual information about the patient's history. Surveys were typically completed by the patient, largely without assistance from a physician. Figure 4.1 depicts the distribution of each variable in the entire cohort. Table 4.3 summarizes these variables as well as the number of inherent missing values.
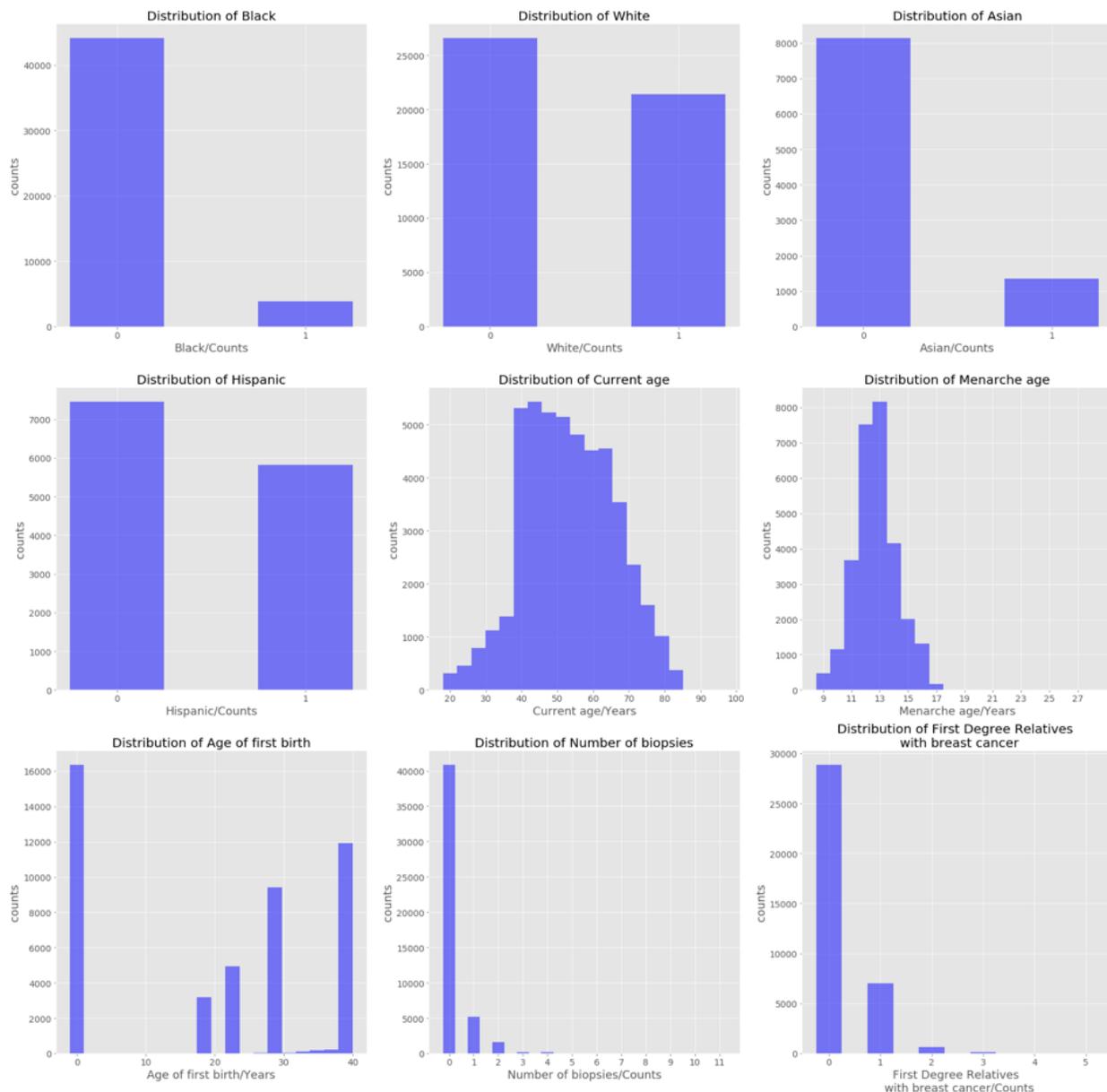
Figure 4.1: Distributions of each variable for the entire cohort. Race is represented as binary indicator variables.

4.2.4 Partially Observable Markov Decision Processes

We designed and evaluated two separate POMDPs for lung and breast cancer screening. Each model consists of three states and two actions. The observations of each POMDP are domain based: in the lung model, they represent findings obtained from LDCT imaging studies, including nodule size, consistency, location, and margins; in the breast model,

Variable Name	Variable Type	# Missing (%)
Current age	Continuous	0 (0)
Age of menarche	Continuous	19,572 (40.8)
Age at first child birth	Continuous	1,630 (3.4)
# of 1st degree relatives with breast cancer	Ordinal	11,459 (23.9)
Number of biopsies	Ordinal	1 ($<$0.1)
White	Indicator variable	0 (0)
Black	Indicator variable	0 (0)
Asian	Indicator variable	38,939 (81.2)
Hispanic	Indicator variable	35,070 (73.1)

Table 4.3: Description of the variables considered by the Gail breast risk model including variable type and percentage of values that are inherently missing in the dataset. The implemented risk model did not adjust for Native Americans; those individuals were excluded from our analysis.

they represent BI-RADS scores derived from MG interpretations. Given the nature of each dataset, both the lung and breast models have a horizon of three and four years, respectively, with 6-month and 1-year epochs. Each epoch represents time points for which we have information on the cancer status of patient (diagnosed with cancer or not). Transition and observation probabilities for each POMDP model are learned using the expectation maximization (EM) algorithm, for learning dynamic Bayesian networks, from each dataset. Both models were solved using the QMDP approximation solver [66]. The computation of transition and observation probabilities is described in chapter 5 sections 5.2.2. An explanation of the QMDP algorithm is described in Chapter 2 section 2.3.4.

Lung cancer screening POMDP

Figure 4.2 (left) depicts the lung POMDP, illustrating the state space and allowed transitions between states, as well as the observations of each state. The state space consists of three states: the no-cancer (NC) state that represents any case with no suspicious abnormalities (i.e., no pulmonary nodules > 4 mm). The uncertain (U) state that represents any case with a noted finding (i.e., nodules 4mm or larger) but not yet a lung cancer. Lastly, the lung-cancer (LC) state is any case with a confirmed lung cancer diagnosis through the use of additional
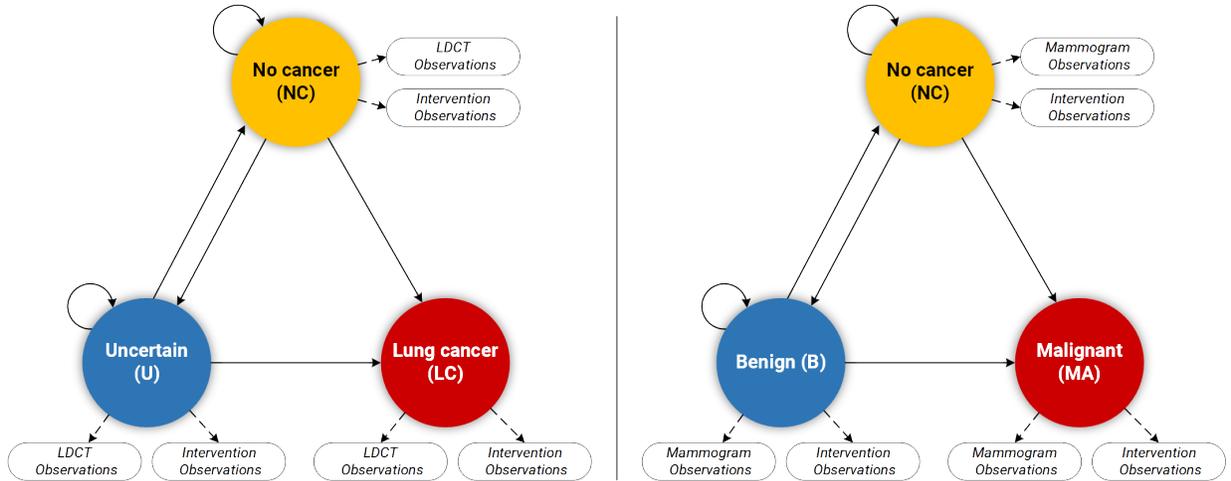
Figure 4.2: **Left**. The lung POMDP; NC: no-cancer state; U: uncertain state; LC: lung cancer state. LDCT and intervention observations can be observed in each state. **Right**. The breast POMDP; NC: non-cancer state; B: benign state; MA: malignant cancer state. MG and intervention observations can be observed in each state.

diagnostic tests. The LC state is terminal such that any individual who enters it leaves the screening process for treatment. An LDCT action implies continuation of screening, whereas an intervention action refers to any diagnostic procedure (e.g., thoracotomy, biopsies, diagnostic CT, positron emissions tomography (PET) scan). Observations represent LDCT findings (nodule size, consistency, margins, and anatomic location) and the occurrence of an intervention. To generate initial belief states for each individual in our dataset we used the Tammemägi PLCO$_{M2012}$ model with demographic and clinical features at baseline to predict the risk of cancer. Demographic features used include age, education, race, and body mass index. Clinical features used were COPD, family history of lung cancer, personal history of cancer, smoking status, smoking intensity, and duration of smoking.

Breast cancer screening POMDP

The breast POMDP model also consists of three states: the no-cancer (NC) state in which no abnormalities are seen, the benign (B) state in which benign breast disease diagnosis follows the MG, and the malignant (MA) cancer state in which the disease is confirmed through biopsy. MA is similarly a terminal state in which the patient leaves the screening process

for treatment. Figure 4.2 (right) shows the breast cancer screening POMDP, transitions, observations (BI-RADS scores 1, 2, 3, 4A, 4B, 4C, 5), and actions. Though an intervention (biopsy in the breast cancer context) is possible after each MG, in practice biopsies are only performed after an MG of BI-RADS 4 or higher. For an initial belief, we used the patient's Gail score. The Gail score is an absolute risk estimate derived using age, age at menarche, age at first birth, the number of first-degree relatives with breast cancer, the number of previous breast biopsies, and race. A detailed description on how to compute individualized can cancer beliefs is presented in 5 section 5.2.2.

4.2.5 Maximum Entropy IRL

In IRL, the reward function, r, is assumed to be a linear combination of feature vectors f_s and weights θ (θ^T is the transpose of θ):

$$r(\tau; \theta) = \theta^T f_\tau = \sum_{s \in \tau} \theta^T f_s \qquad (4.2.1)$$

A feature count, (f_τ), is the sum of feature vectors of the states visited along a trajectory, where f_s represents binary vectors indicating state values. Inputs to the MaxEnt IRL algorithm are an MDP and a set of trajectories (D) [102]. A path or a trajectory (τ) represents the sequence of states (s) and ensuing actions followed by an agent in an MDP. For example, in the NLST dataset, a trajectory comprises three epochs (i.e., the three annual screening exams) with state-action pairs describing the lung cancer states and the actions taken (e.g., NC-LDCT, U-LDCT, and IC-I$_{\text{Biopsy}}$). The probability of a trajectory occurring in our set of trajectories is proportional to the exponential of the reward/cost of the trajectory [103]:

$$p(\tau; \theta) \propto \exp\left(r(\tau; \theta)\right) \qquad (4.2.2)$$

As such, trajectories of equal reward are equally likely to be executed by the expert, whereas trajectories of less reward are less likely. The probability distribution over paths with max-

imum information entropy is parameterized over θ. $Z(\theta)$ is the partition function, where $Z(\theta) = \sum_{\tau \in D} \exp r(\tau; \theta)$.

$$p(\tau; \theta) = \frac{1}{Z(\theta)} \exp \left(r(\tau; \theta) \right) \tag{4.2.3}$$

The log likelihood of the trajectories (loss function) is shown in Equation 4.2.4, M is the number of trajectories:

$$L = \frac{1}{M} \sum_{\tau \in D} r(\tau; \theta) - \log \sum_{\tau \in D} \exp \left(r(\tau; \theta) \right) \tag{4.2.4}$$

This loss function is convex for a linear reward function and a deterministic MDP. To update θ we use a gradient descent function, where η represents the learning rate:

$$\theta_{i+1} = \theta_i + \eta \nabla_\theta L \tag{4.2.5}$$

The gradient $\nabla_\theta L$ represents the difference of feature expectations and sum over state visitation frequencies multiplied with feature vectors:

$$\nabla_\theta L = \tilde{f} - \sum_{s_i} D_{s_i} f_{s_i} \tag{4.2.6}$$

A feature expectation, $(\tilde{f})$, is defined as the average of all feature counts across all trajectories. The frequency of state visitation, D_{s_i}, can be computed using a dynamic programming algorithm; see [102, 103] for more information regarding this algorithm. The pseudocode of the MaxEnt IRL algorithm can be found in Algorithm 5 [103].

4.2.6 Adaptive step size

To improve the convergence of the MaxEnt IRL algorithm, we introduce an adaptive learning rate approach for the update rule of the gradient descent. The idea behind making the step size adaptive is to calculate the inner product of $\nabla_\theta L$, the gradient, in the current step, i.e.,

---
Algorithm 5: MaxEnt Inverse Reinforcement Learning

Input: MDP, and trajectories D
Output: State rewards
Randomly initialize θ;
repeat
    Solve for the optimal stochastic policy using $r(\tau) = \theta^T f_\tau$ with value iteration;
    Use a dynamic algorithm to compute $p(s|\theta, T) = D_s$ the state visitation
    frequencies;
    Compute the gradient $\nabla_\theta L$;
    Update θ;
until *convergence*;

$\nabla_\theta L_i$ with $\nabla_\theta L_{i-1}$, its value from the previous step. If the two are in the same direction then the step size can be increased, otherwise it is decreased. Following [104] we define the learning rate $\eta = \frac{\alpha}{(t+A)^\alpha}$, where t is dependent on the gradient inner product (which becomes the dot product in higher dimensions); α and A are constants. The role of t is to regulate the learning rate:

$$t_{i+1} = \max(t_i + f(\langle -\nabla_\theta L_i, \nabla_\theta L_{i-1}\rangle), 0) \tag{4.2.7}$$

In this definition, $f(\cdot)$ represents the following sigmoidal function where $f(x) = f_{min} + \frac{f_{max} - f_{min}}{1 - \frac{f_{max}}{f_{min}} \exp{-\frac{x}{\omega}}}$. In the above expressions, α, A, f_{min}, f_{max}, and ω are user-defined constants obtained from [104]. With $f_{min} < 0$, $f_{max} > 0$, and $\omega > 0$.

4.2.7 Computation of rewards

We assumed that given the outcome of a known cancer diagnosis for each individual over time, partial observability was no longer a problem while training, so learning the rewards of state-action pairs of an MDP instead of a POMDP was sufficient and computationally more efficient. However, the MaxEnt IRL algorithm computes the rewards of each state of an MDP, not state-action pair rewards ($r(s, a)$). To estimate rewards for each state-action pair combination, we designed two MDPs:
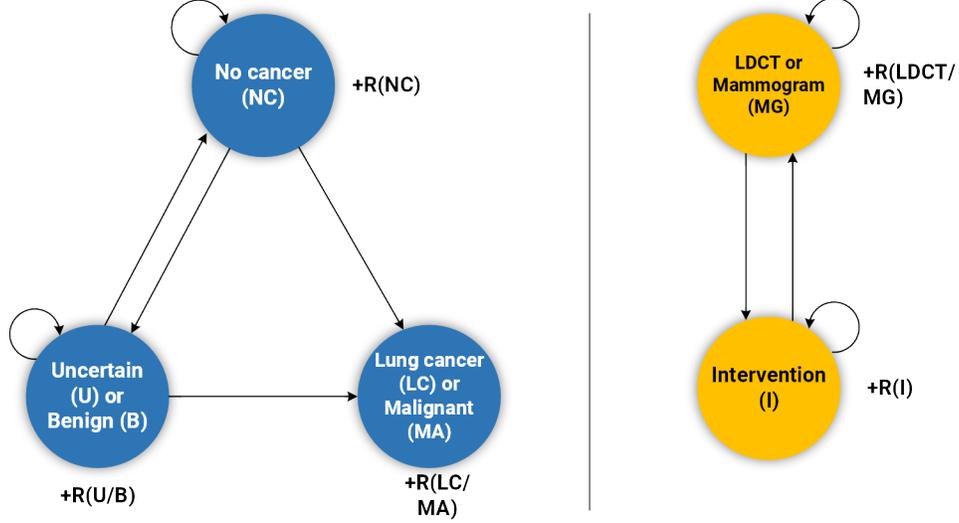
Figure 4.3: **Left**. The state MDP; NC: non-cancer state; U/B: uncertain or benign state; I/MA: lung or malignant cancer state, respectively for the lung and breast models. **Right**. The action MDP; LDCT/MG: state after a LDCT or MG; I: state after an intervention (e.g., biopsy); +R(·): rewards experienced by the agent in each state.

1. *A state MDP model.* The states of this MDP are the states depicted in Figure 4.3, for the lung and breast models. The transition matrix of the state MDP is the same transition matrix used in its respective POMDP model.

2. *An action MDP model.* In the action MDP, the states are defined by the previous action of the agent. These states model the options for screening (e.g., continue annual screening) and intervention (e.g., biopsy), in which the agent enters after performing each action. The action MDP transition model represents the probability of transitioning from the LDCT/MG state to the I state.

Figure 4.3 demonstrates the two MDPs. A combinatorial design decision inspired by [105] was used to learn state-action pair rewards. State-action pair rewards are computed using a multiplicative model shown in Equation 4.2.8:

$$R(s, a) = R(s) \cdot R(a) \tag{4.2.8}$$

73

4.3 Evaluation and Results

A stratified 5-fold cross validation study design was used to evaluate the POMDP models built from the NLST and the Athena datasets. The training set of each fold is used to learn the transition and observation matrices of the POMDPs, as well as the rewards using the MaxEnt IRL algorithm.

4.3.1 Comparison of MaxEnt IRL with & without adaptive step size

Table 4.3 shows the reward value of each state and action as well as different normalizations of these rewards computed using the MaxEnt IRL algorithm with an adaptive step size. We compare the MaxEnt IRL with and without the adaptive step size and assess the speed of convergence. Figure 4.4 depicts the computed rewards for states and actions for the lung POMDP over the number of iterations of gradient descent in the MaxEnt IRL algorithm, with and without an adaptive step size. A similar convergence trend is observed with the breast POMDP. As shown, the adaptive step size method converges to the correct solution more quickly than the standard MaxEnt IRL implementation. For the evaluation of the two models we use a reward function derived from rewards normalized in the [-1,1] range.

Normalization	R(NC)	R(U/B)	R(LC)	R(LDCT/M)	R(I)
Lung cancer					
None	83.530	127.410	-835.730	497.610	-427.530
By range	0.080	0.120	-0.800	0.540	-0.460
[0,1]	0.950	1.000	0.000	1.000	0.000
[-1,1]	0.910	1.000	-1.000	1.000	-1.000
Breast cancer					
None	-37.930	103.950	-571.420	-0.840	-1179.820
By range	-0.050	0.150	-0.800	-0.001	-0.999
[0,1]	0.790	1.000	0.000	1.000	0.000
[-1,1]	0.580	1.000	-1.000	1.000	-1.000

Table 4.4: The rewards for each state (R(NC), R(U/B), R(LC/MA)) and action (R(LDCT/MG), R(I)) computed using the MaxEnt IRL algorithm, for one of the folds of the 5-fold cross validation, with an adaptive step size.

4.3.2 Lung and breast POMDP results

We used the longitudinal observations from the NLST and Athena datasets as input to POMDPs such that each sequential observation updates the belief state of the agent. The belief state of the POMDP, at each epoch, is then used to select the next (optimal) action, with the objective of early detection of cancer. The POMDP models can suggest to continue screening (i.e., MG, LDCT) or to perform an intervention (i.e., biopsy or diagnostic imaging). If an intervention is performed, the individual is removed from further consideration. Evaluation of the POMDP is posed as a binary problem: if the POMDP suggests continued screening (LDCT/MG) then the patient is classified as a *negative* cancer; if it suggests an intervention, then the patient is classified as a *positive* cancer. Based on this definition, if the model suggests a LDCT/MG and the patient did not have a confirmed diagnosis of cancer in a given epoch, it is considered a true negative (TN); if the patient had a confirmed diagnosis of cancer then it is a false negative (FN). Conversely, if the model suggests an intervention and the patient did not have cancer in a given epoch, then it is considered a false positive (FP); if the patient had a diagnosis of cancer then it is considered a true positive (TP). Performance metrics were estimated for each epoch of the screening process. Any subject diagnosed with cancer is removed from the subsequent epoch. The POMDP models are compared against the equivalent physician decisions (recommendations) at each epoch, applying a similar framework for TN/FN/FP/TP to the experts, given the known cancer outcomes from each dataset (e.g., if the physicians suggested an LDCT/MG and the patient did not have a confirmed diagnosis of cancer, it is considered a true negative, etc.).

Table 4.3.2 shows the performance of the lung and breast POMDPs and the corresponding performance of physicians on the same dataset. Notably, both POMDP models show performance comparable to experts. The lung cancer screening model has worse performance in terms of recall in the first and third screening epochs, but an improved performance in terms of recall and false positive rate in the second screening and post-screening. The breast cancer screening model demonstrates excellent recall (as do the expert physicians)

but slightly worse false positive rate. The Cohen's kappa coefficient of agreement was used to assess the concordance between the POMDP models and physicians. The kappa score of the lung POMDP and physicians decreases over time due to the large number of false positives. A large portion of different cases are classified as false positives between the lung POMDP and physicians. The breast POMDP has a high kappa score demonstrating strong agreement with physicians in terms of false positives and true positives. For both lung and breast models, the variance of kappa per screening is less than 0.03.
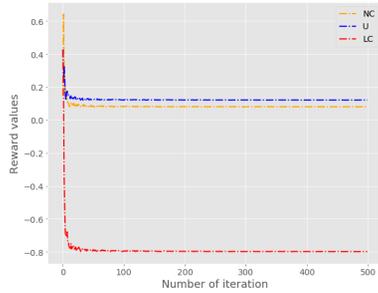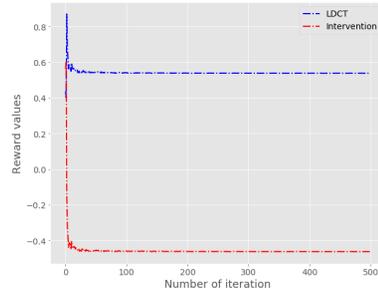
(a) Lung cancer states' rewards.   (b) Lung cancer actions' rewards.   (c) Breast cancer states' rewards.   (d) Breast cancer actions' rewards.
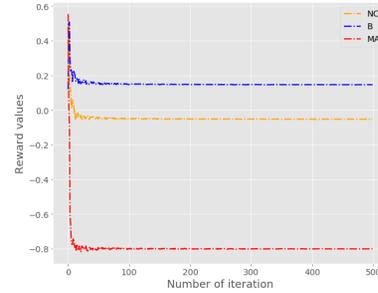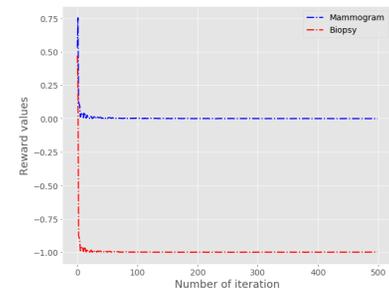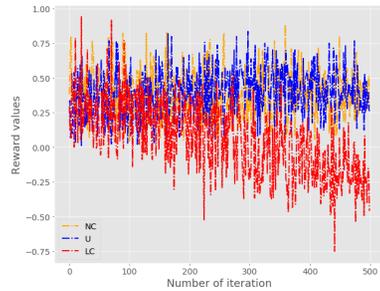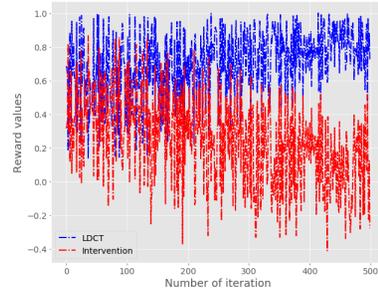
(e) Lung cancer states' rewards.   (f) Lung cancer actions' rewards.   (g) Breast cancer states' rewards.   (h) Breast cancer actions' rewards.
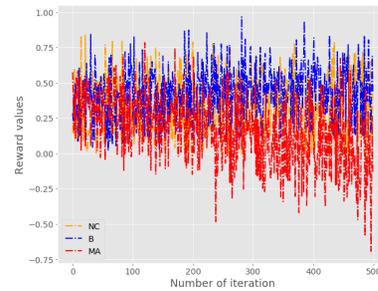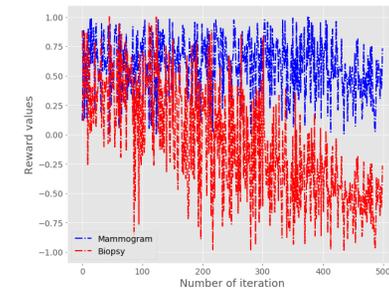
Figure 4.4: State and action rewards computed using the MaxEnt IRL and normalized by range. **Top**: Using an adaptive step size. **Bottom**: Without using an adaptive step size. The adaptive step size MaxEnt IRL algorithm converges to a solution significantly faster than the MaxEnt IRL without an adaptive step size.

	POMDP						Physicians						Kappa
	TN rate	FP rate	FN rate	TP rate	Precision	Recall	TN rate	FP rate	FN rate	TP rate	Precision	Recall	
	Lung cancer												
Training	NCs: 3960 , Cs: Scr1, 2, 3 = 130, 68, 86 ; Pst-Scr = 78												
Scr 1	0.48	0.52	0.02	0.98	0.05	0.98	0.48	0.52	0.00	1.00	0.06	1.00	0.42
Scr 2	0.34	0.66	0.02	0.98	0.02	0.98	0.34	0.67	0.05	0.95	0.02	0.95	0.29
Scr 3	0.24	0.76	0.01	0.99	0.03	0.99	0.21	0.79	0.00	1.00	0.03	1.00	0.05
Pst-Scr	0.25	0.75	0.07	0.93	0.02	0.93	0.22	0.78	0.14	0.86	0.02	0.86	0.05
Testing	NCs: 990 , Cs: Scr1, 2, 3 = 32, 17, 21 ; Pst-Scr = 20												
Scr 1	0.48	0.52	0.04	0.96	0.05	0.96	0.48	0.52	0.00	1.00	0.06	1.00	0.42
Scr 2	0.35	0.65	0.02	0.98	0.02	0.98	0.33	0.67	0.05	0.95	0.02	0.95	0.30
Scr 3	0.25	0.75	0.05	0.95	0.03	0.97	0.21	0.79	0.00	1.00	0.03	1.00	0.07
Pst-Scr	0.25	0.75	0.07	0.93	0.02	0.93	0.22	0.78	0.14	0.86	0.02	0.86	0.06
	Breast cancer												
Training	NCs: 2808 , Cs: Scr1, 2, 3, 4 = 370, 68, 27, 5												
Scr 1	0.99	0.01	0.01	0.99	0.96	0.99	0.99	0.01	0.01	0.99	0.95	0.99	1.00
Scr 2	0.99	0.01	0.01	0.99	0.70	0.99	0.99	0.01	0.01	0.99	0.73	0.99	0.97
Scr 3	0.98	0.02	0.03	0.97	0.40	0.97	0.99	0.01	0.03	0.97	0.43	0.97	0.95
Scr 4	0.98	0.02	0.00	1.00	0.09	1.00	0.98	0.02	0.00	1.00	0.10	1.00	0.92
Testing	NCs: 703 , Cs: Scr1, 2, 3, 4 = 93, 17, 7, 1												
Scr 1	0.99	0.01	0.01	0.99	0.96	0.99	0.99	0.01	0.01	0.99	0.99	0.99	1.00
Scr 2	0.99	0.01	0.01	0.99	0.70	0.99	0.99	0.01	0.01	0.99	0.74	0.99	0.97
Scr 3	0.99	0.01	0.03	0.97	0.40	0.97	0.99	0.01	0.03	0.97	0.44	0.97	0.95
Scr 4	0.98	0.02	0.00	1.00	0.09	1.00	0.98	0.02	0.00	1.00	0.10	1.00	0.91

Table 4.5: **Left**: The lung and breast POMDPs performance per epoch. **Right**: The physicians performance at each epoch. Metrics used for this evaluation are the true positive rate (TP), false negative rate (FN), false positive rate (FP) true negative rate (TN), precision (P), and recall (R). NCs: no-cancer cases. Cs: cancer cases. Kappa: Cohen's kappa score (coefficient of agreement), variance of kappa for all scores: < 0.03.

4.4 Discussion

POMDPs, through the use of beliefs and a hidden state space, can overcome some of the limitations seen in other sequential decision making models used in cancer screening. For instance, given the uncertainty in diagnosing lung and breast cancer from imaging studies, we modeled a hidden cancer state space in three parts [106]: no-cancer, benign/indeterminate, and malignant/lung cancer. Modeling the cancer state space with an additional state rather than a binary state space allows the distinction of lower risk individuals (i.e., no abnormalities) – who constitute a large portion of screening cases and thus result in highly imbalanced datasets – over medium (i.e., benign growth) and high risk individuals (i.e., malignant abnormality).

Driven by the need to define the reward function in these screening POMDPs, we explored the use of the MaxEnt IRL algorithm towards generation of state-action reward pairs. As noted earlier, cost and utility estimation are frequently adopted as reward functions in healthcare models. [80] uses the National Statistical services' costs of procedures to define

reward functions, while QALYs and a lifetime mortality risk model [84] are common alternative approaches. However, cost has certain limitations as it does not generalize to the whole population equally, and does not reflect the importance of quality outcomes. Additionally, QALY data are scarce, and arguably expensive to collect [84]. In contrast, a reward function learned using the MaxEnt IRL algorithm aims to maximize the objective of state-action trajectories.

In this chapter, we used the MaxEnt IRL algorithm to generate reward functions for lung and breast cancer screening POMDP models using experts retrospective decisions. We improved the speed and accuracy of convergence of the gradient descent optimization of the MaxEnt IRL algorithm using an adaptive step size. Moreover, we introduced a multiplicative model for representing state-action pairs as products of state rewards and action rewards. The multiplicative model has the advantage to clearly demonstrate the difference in utility between rewards of different actions, which is what drives decision recommendation. Rewards are thus learned based on the state-visitation frequency of each trajectory. In this context, states with fewer visitations across each trajectory earn the lowest reward (e.g., lung or malignant cancer state), which is why only cancer and non-cancer cases with a complete trajectory are used to learn rewards in our framework. Modeling the expert's decisions with the MaxEnt IRL algorithm resulted in reward functions for the POMDP models with performance comparable to experts. We noticed that when using aggressive reward functions (i.e., identifying all cancer cases), the true positive rate exceeded physicians' true positive rate but at the expense of a higher false positive rate, which in clinical practice can translate into higher costs and unnecessary psychological burden on the patient. Including more observational variables, derived from medical images, in the screening process can overcome this trade-off between true positive and false positive rate. The overall true positive rate and false positive rate using our learned reward functions in the POMDPs is comparable to experts. Nonetheless, in some cases the experts had false negative cases, which is also captured by our approach. When compared with other machine learning algorithms at

the baseline of the lung and breast paradigms the POMDP models demonstrate improved performance.

The kappa coefficient of agreement between the POMDP models and physicians is constantly high for the breast POMDP model, illustrating the discriminatory capability of BI-RADS score as an imaging observation. In our lung cancer screening model, kappa gradually decreased over ensuing epochs, suggesting variability in the interpretation of LDCT imaging observations between the POMDP and the physicians. The lung POMDP is not fully replicating physicians' decision making patterns despite its overall performance being comparable to experts. When it comes to early cancer prediction (e.g., predicting screening 3 cancer from screening 1) the lung POMDP outperforms physicians, suggesting that the model and reward function are discriminating in a different way between positive and negative cases. Error analysis of the lung POMDP false positives shows a different subset from the physicians.

MaxEnt IRL also handles partial trajectories, making it suitable for screening processes in which individuals diagnosed with the disease exit the screening process for treatment. Relative to other IRL methods, MaxEnt IRL has the advantage of handling ambiguity by using a probabilistic model of behavior that exponentially prefers trajectories of higher reward [76, 103]. MaxEnt IRL can also be used to transfer knowledge between datasets, tasks or domains by reusing learned weights (i.e., transfer learning). The only "partial" trajectory cases employed, in this analysis, are individuals diagnosed with cancer across the horizon of the screening process.

The first limitation of using MaxEnt IRL in this study is the fact that more than one combination of rewards can define the same problem. To overcome this, a policy iteration algorithm can be used rather than value iteration algorithm to learn optimal policies, as the policy space is finite in comparison to the rewards space (hence the policy iteration algorithm is guaranteed to optimally converge). A second limitation is the assumption that reward functions are only based on state visitation frequencies. The utility of screening

recommendations is subjective and defined by different factors such as cost, quality of life, and patient satisfaction. To assess the quality of these reward functions a comparison of suggested recommendations with patient satisfaction could be used.

Other limitations are around assumptions about the nature of our datasets. While lung and breast cancer screening tests occurred roughly at one year intervals, we assumed that screening occurs annually (i.e., at fixed frequency). Moreover, data imbalance is a function of time, as at each screening point the number of cancer and non-cancer cases changes (i.e., at the outset of a screening period, more cancers are found at the beginning of a dataset). We did not account for this dynamic nature of the dataset during training. Given the small number of cancer cases across each screening point of both datasets, we utilized a stratified 5-fold cross-validation to obtain an unbiased estimate of model performance. Similarly, other temporal studies have used a k-fold cross validation to assess model performance [50, 51, 106–109]. To simplify modeling, our lung POMDP model considered only cases reporting a single pulmonary nodule over the course of the trial; this represents only a subset of the screened individuals, as many subjects have more than one such finding. A more concrete analysis would include cases with multiple nodules over time. However, it was not possible to ascertain the history of individual nodules in patients with multiple nodules as tracking of the nodules was not considered at the time of the study. Lastly, for the Athena dataset, in breast cancer screening, patients with BI-RADS 1, 2, or 3 rarely undergo biopsy, thus the true FN rate is likely underestimated. Future work involves the exploration of MaxEnt IRL in transfer learning between other datasets and domains, by reusing learned weights.

4.5   Summary of findings

In this chapter, we uncovered the role of learning accurate reward functions for POMDPs from data. More specifically, we used the MaxEnt IRL algorithm to generate reward functions

for lung and breast cancer screening POMDP models. We improved the speed and accuracy of convergence of the gradient descent optimization of the MaxEnt IRL algorithm using an adaptive step size. We introduced a multiplicative model for representing state-action pairs as products of state rewards and action rewards. We evaluated the decisions of the lung and breast cancer POMDP models and compared it against expert decisions. The learned reward functions resulted in POMDP decisions comparable to experts' decisions.

Chapter 5

Using Sequential Decision Making to Improve Lung Cancer Screening Performance

5.1 Overview

Building from the learned rewards function in the Chapter 4, here we explain how to learn a POMDP's remaining components (transition and observation models, initial beliefs). More specifically, we develop a predictive model informing personalized lung cancer screening policies using machine learning and sequential decision making methods. The established framework, for learning POMDPs, progressively optimizes the choice of screening actions given prior observations. We leveraged different techniques to learn a POMDP from NLST data: we integrated a dynamic Bayesian network (DBN) into the POMDP to predict the chance of developing lung cancer and to determine the POMDP's observation and transition probabilities, and we applied inverse reinforcement learning (IRL) to formulate a rewards model [110], mimicking experts' decisions.

We trained and tested our POMDP using a dataset of 5,402 single nodule unique trajec-

tories of lung cancer screening patients from the NLST LDCT trial arm. We compared our model's decisions with experts' decisions over time, and found that: 1) our POMDP lowered the false positive rate for most screenings in the NLST, while maintaining true positive detection rates; and 2) our POMDP improves early prediction of cancer cases with indeterminate pulmonary nodules (IPNs, nodules having some risk of developing into cancer [111]) as compared to radiologists' interpretation.

5.2 Methods

5.2.1 NLST dataset

In this work, we used data gathered from NLST's LDCT arm. Of this population, only 10,231 cases had one or more solitary IPNs over the study period. We further constrained our data to individuals with only one IPN reported in the same anatomical lung lobe during the study, assuming that the same nodule was observed over time. A detailed description of this dataset and pre-processing steps is described in Chapter 4 sections 4.2.1 and 4.2.2.

To perform a five-fold stratified cross validation (80 : 20% training:test ratio) with this data, we randomly generated each fold while maintaining the relative proportion of cancer to non-cancer cases seen at each screening time point of the NLST study.

5.2.2 Defining and learning the POMDP components

States (s) **and actions** (a)

Figure 5.1(a) illustrates the lung cancer screening POMDP state space, observations, and potential state transitions. We adopted a state space used in our earlier work [106]. This state space consists of three states defined around the true cancer state of each subject after each screening. No-cancer (NC) is the state in which the individual has no remarkable findings for lung cancer (e.g., lung nodules < 4 mm). The Uncertain (U) state is an intermediate state

in which an individual exhibits suspicious abnormalities (e.g., lung nodules ≥ 4 mm) but no confirmed diagnosis of lung cancer. The Lung Cancer (LC) state represents any case with a confirmed lung cancer diagnosis through the use of additional procedures. LC is a terminal state in which an individual enters and simultaneously leaves the screening process (as NLST participants diagnosed with lung cancer were removed from the clinical trial for treatment). We simplified the set of possible actions into two types, embodying the core decisions made by experts: to continue screening with a follow-up LDCT or to recommend an intervention (i.e., any procedure performed in relation to diagnosing lung cancer).

Observations (z)

Following from the NLST's screening paradigm, two types of observations are possible: those coming from annual screens (LDCT findings) and interpretation and those arising from a diagnostic intervention. To capture the interactions between the nodule size, consistency, and margins we used a model to combine the observations into a single representation as a probability. Specifically, we used a DBN to infer the probability of cancer over time from these observations. Alternative models were considered, including logistic regression, and an exhaustive search of all combinations of observations, with the DBN and the exhaustive search demonstrating the best performance in conjunction with the POMDP (see Appendix Table A.30). The DBN topology was learned from the data: we learned the intra-slice structure of the DBN (i.e., conditional dependencies between variables in the same time step) using t_0 observations from the K2 algorithm in the Bayes Net Toolbox (BNT) [112]; and inter-slice structure (i.e., dependencies over time) was learned using cases that had a complete trajectory of screening over the NLST screening period (i.e., no missing observations) using the batch Expectation-Maximization (EM) algorithm also in BNT. Figure 5.1(b) shows the intra- and inter-slice structure of the learned model, which we then parameterized using training data. In the POMDP, we then used this DBN with observations of a given patient to infer a probability of cancer over time as our new observation. These probabilistic obser-

(a) Lung cancer screening POMDP.

(b) DBN-based observation model.

(c) Transition model.
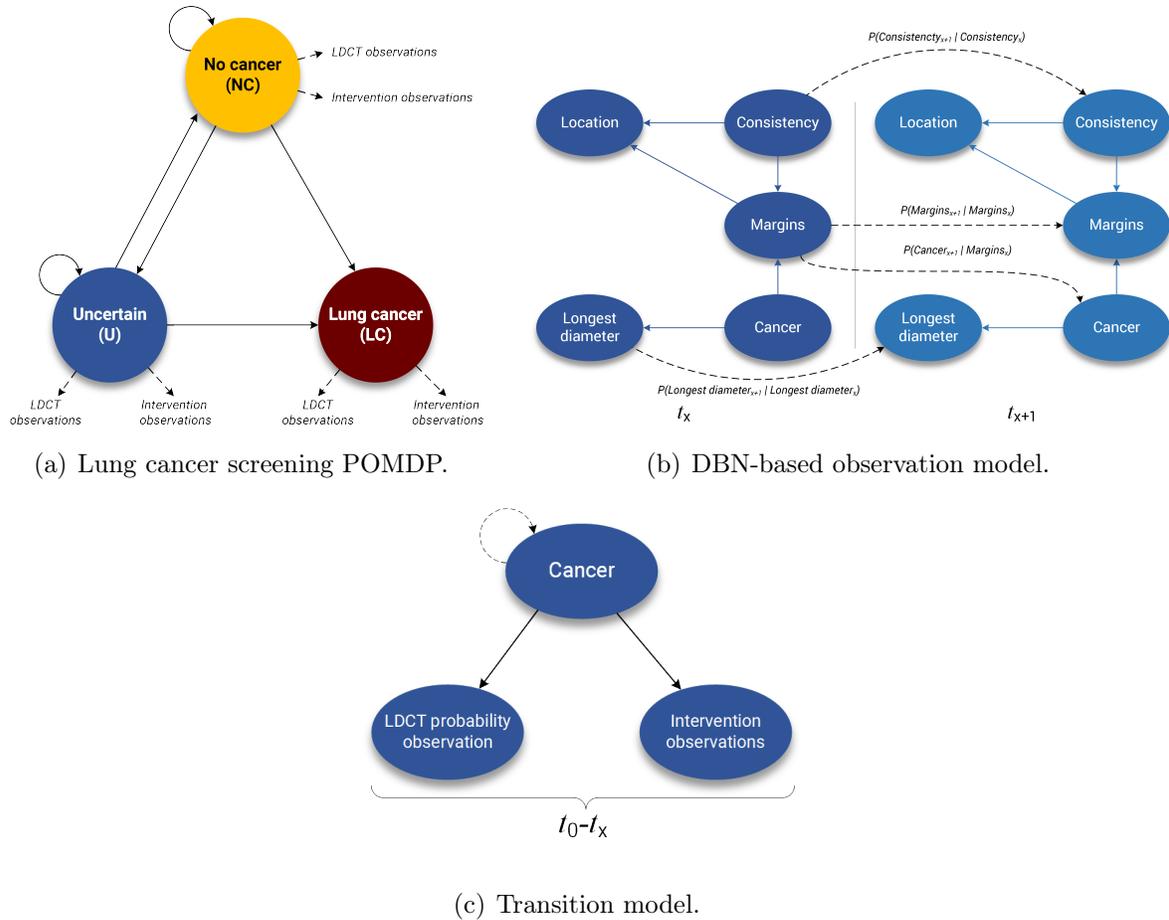
Figure 5.1: **(a)** NC represents a non-cancer state, U is an intermediate uncertain cancer state, and LC is the lung cancer state. Arrows indicate allowed transitions between states. LDCT and intervention observations represent the possible observations of the model in each state. **(b)** The nodule size node represents the possible categories of nodule size. The consistency node represents the categories of nodule consistency and the margin node the categories of nodule margins. The Cancer node represents the categories of cancer or no cancer. t_0 represents the intra-slice structure of the model. Solid line arrows represent the intra-slice interactions between nodes. The inter-slice structure is depicted between the t_0 and t_1 time slices. Dashed arrows represent inter-slice interactions between variables over time. This DBN is recurring for 5-time steps ($x = 4$). **(c)** The LDCT probability observations represent the 100 bins of probabilities as categories. The Intervention observations node consists of two categories the observation of cancer or not, from diagnostic procedures. The Cancer node consists of three states the NC, U and LC cancer states. Solid arrows represent the intra-slice structure interactions between variables. Dashed arrows represent the inter-slice connections between variables.This structure is repeated over a horizon of 5 ($x = 4$) time-steps.

vations were discretized in 100 equal sized bins, from 0-1. For intervention observations, we determined if an individual undergo an intervention and was diagnosed with cancer or did

not undergo an intervention.

Transition and observation probabilities

Transition and observation probabilities were computed using a temporal naïve Bayes DBN model, per Figure 5.1(c): the LDCT node represents a conditional probability table (CPT) of 100 categories corresponding to each discretized probability; the Intervention node represents a CPT table of two observations, cancer after an intervention or no cancer with or without an intervention; and the Cancer node represents a CPT table of three categories per our state model. Usually, the transition probabilities of a POMDP are different based on the choice of action in a given state $(T(s_j, s_i, a))$. The transition matrix used for the lung cancer POMDP model is assumed to be invariant of action. But the observation matrix $(O(z, s, a))$ is state and action dependent. We modeled the observations of Intervention as being impossible (i.e., probability of zero) when the action of LDCT is performed and the observation of an LDCT as impossible when the action of Intervention is performed. An important implementation note is in regards to sparsity, as some LDCT probabilities will be calculated as zero given no instances in the dataset (although they are feasible in real-world settings). Thus, to deal with sparsity we replaced all zero probabilities with a very small probability (0.0001) and normalized over the matrix to improve overall inference [113].

Rewards

A POMDP's reward function defines the behavior of the agent as it aims to optimize based on returned values. In our POMDP, we define rewards in terms of a state-action pair (R(s,a)). We learned a reward function using the recommendations of experts from the NLST dataset. Using inverse reinforcement learning (IRL), we learned state and action rewards via an adaptive maximum entropy IRL algorithm [110] (see Chapter 4). A multiplicative model was then employed to learn each combination of state-action pair rewards.

Initial beliefs

In a POMDP, the belief state is a probability distribution over the states of the process. The initial belief is the initial probability distribution over the states at time t_0. To generate initial beliefs for each individual we used the $PLCO_{M2012}$ model [99] with demographic and clinical features at baseline to predict the risk of cancer. Tammemägi et al. [114] used the Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) Screening Trial to develop 6-year lung cancer risk models. The models achieved high discrimination and calibration performance. The $PLCO_{M2012}$ is an updated version of the original model trained and validated on the PLCO dataset and externally validated on the NLST cohort. The variables and weights of the logistic regression model used are the same as reported in the $PLCO_{M2012}$ model [99]. Demographic features include age, education, race, and body mass index (BMI). Clinical features encompassed the presence of chronic obstructive pulmonary disease (COPD), family history of lung cancer, personal history of cancer, smoking status, smoking intensity, and duration of smoking. To generate an initial belief of cancer over the three states of our state space, we used the following rule: the probability of the LC state is the risk of cancer times two computed by the $PLCO_{M2012}$ model; the probability of the U state is assumed to be zero and the probability of the NC state is the complement of LC. To update beliefs we follow the basic recursive filtering rule [65], given by Equation 5.2.1 where α is a normalization constant such that $\alpha = \frac{1}{\sum_{s_j} P(o|s_j) \sum_{s_i} P(s_j|s_i,a)b(s_i)}$.

$$b'(s_j) = \alpha \cdot P(o|s_j) \sum_{s_j} P(s_j|s_i, a)b(s_i) \tag{5.2.1}$$

5.2.3 Solving the POMDP model

Our proposed POMDP model has three states, two actions and 102 observations. To solve infinite horizon problems, we can use approximation algorithms [61,63–65], providing significant speed-up. We opted to use the QMDP approximation algorithm, shown in Algorithm 3. To select optimal actions that maximize expected utility we use Algorithm 4.

5.3 Results

NLST participants underwent three annual screenings with follow-up over six years to identify subsequent lung cancers. At each screening time point (t_0, t_1, t_2), a radiologist read the imaging study and made a decision to refer patients for a diagnostic procedure (e.g., early repeat LDCT, diagnostic CT, PET-CT, or biopsy/tissue sampling) or to continue annual LDCT screening. Our POMDP suggests actions at these three screening time points as well as between the screenings using imputation, resulting in five recommendations in 6-month intervals (Figure 5.2(a)). Observations used by the POMDP include imaging features about nodule size, margins, location, and consistency. Our evaluation examines the POMDP's recommended actions over all five points $(a_{POMDP_0}, a_{POMDP_{0.5}}, a_{POMDP_1}, a_{POMDP_{1.5}}, a_{POMDP_2})$ and directly compares against the physicians' performance at the annual screenings.

5.3.1 POMDP versus physician performance

To compare the performance of the POMDP model against physicians we calculated the precision (positive predictive value, PPV), recall/true positive (TP) rate (sensitivity), and true

(a) NLST timeline.   (b) Definition of evaluation metrics.
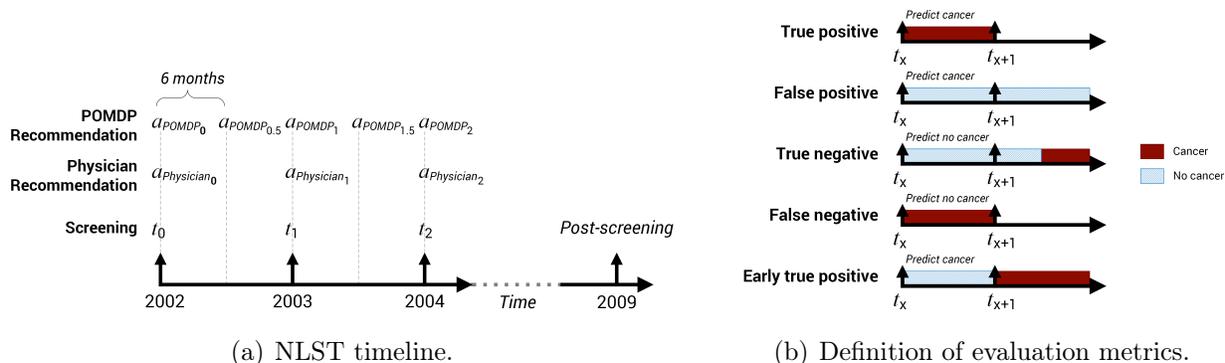
Figure 5.2: **(a)** Screenings represent annual LDCT imaging observations with information about the subject's cancer status. In contrast, our POMDP model suggests screening recommendations every six months. **(b)** Illustration depicting true positive/negative and false positive/negative cases for the POMDP's performance over time. The colored bars indicate truth based on the NLST observations and subjects' known outcomes. We also demonstrate how early true positives are defined in this study.

negative (TN) rate (specificity) for recommended actions at each screening point. We used the following criteria to assess our model: if the POMDP suggests a diagnostic intervention and the individual is subsequently diagnosed with cancer in the following time period, it is counted as a true positive, otherwise it is considered a false positive; if the POMDP suggests no diagnostic intervention, but an annual LDCT screen, and the individual is diagnosed with cancer in the following screen, it is a false negative (FN), otherwise it is a true negative (Figure 5.2(b)).

We assessed our POMDP's performance based on a five-fold cross validation design. To match physicians' TP rates (who had a lower threshold for positive screens) and obtain comparable results, we adjusted the POMDP rewards function (using the training data) to be more conservative. We then evaluated this updated POMDP on our testing data. Table 5.1 shows the results of the POMDP model with tuned rewards against physicians' performance. Our model reduces the FP rate in most screenings (t_1, t_2, and post-screening) compared to the experts while maintaining a high TP rate for screening: at t_0, TN and TP rates are 2% lower and 3% lower than the physicians'; at t_1, TN and TP rate are 1% higher and 3% higher; at t_2, TN and TP rate are 4% higher and 4% lower; and in the post-screening period the POMDP's TN and TP rate are 3% higher and 8% higher than the experts', respectively. We also analyzed the performance of the POMDP model for earlier cancer detection (i.e., detection of a t_2 cancer at t_1). The detection of early TPs is also improved with earlier diagnostic recommendations (e.g., the TP rate for action $a_{POMDP_0}, a_{POMDP_{0.5}}, a_{POMDP_1}, a_{POMDP_{1.5}}$ for t_2 and post-screening) compared to physicians' recommendations. The POMDP TP rate is higher than the physician's over time for post-screening, as depicted in Figure 5.5 and discussed in the following section.

5.3.2   Understanding POMDP and physician differences

We calculated a kappa score to test the level of agreement between physicians and the POMDP. Notably, kappa values trended lower, implying that the POMDP and experts

classify different cases positively over time, which influences the FP rate. To elucidate this difference, we grouped subjects predicted to have lung cancer by the POMDP vs. physicians, analyzing cases where they had different predictions. The preponderance of subjects different between the groups were individuals classified as FPs or early TP cases (i.e., cases predicted as positives earlier by the POMDP relative to their cancer diagnosis in the NLST trial). Figure 5.3 depicts these two cohorts. We explored the feature distributions of each group to assess similarity. We used chi-squared or Fisher's tests for categorical variables and the Student or Wilcoxon-Mann-Whitney tests for continuous variables. Additionally, to assess the effect size of the computed p-value we used the Cramer's V and the r^2 or Cohen's r^2 effect size, correspondingly, for each test [115–117]. Tests with p-values < 0.05 were considered significant. The false positive analysis showed that smoking years, age, largest nodule size at t_0, and smoking quit time had significantly different distributions and the largest effect size between the groups of post-screening cases (see Table 5.2). The additional early prediction TP cases predicted by the POMDP model in comparison with the physicians showed that nodule size at t_0 (largest nodule diameter) and smoking years were significantly different between the groups. The nodule size was shorter and years of smoking less than the early TPs predicted by both the physicians and POMDPs (see Table 5.2 and Figure 5.3). A full analysis comparing these groups is presented in the Appendix.

5.3.3 POMDP stability

In the NLST, a minimum threshold of 4 mm was used to classify findings as nodules. A later analysis [118, 119] showed that changing this threshold to 6 mm significantly reduced the FP rate while maintaining the same TP rate [120, 121]. As such, we stratified our cases into nodules < 6 and ≥ 6 mm at baseline and tested the POMDP. To assess the robustness and performance distribution of the POMDP model we performed a bootstrap evaluation, randomly sampling from our NLST dataset 240 times to define our training and testing sets. Subsequently, all performance measures for each seed were used to calculate the median, the

	Cancers	Non-cancers		POMDP			Physicians			Kappa
				TN rate	Recall	Precision	TN	Recall	Precision	
Screening t_0	32	1,047	a_0	0.46	0.97	0.05	0.48	1.00	0.06	0.41
Screening t_1	17	1,030	a_0	0.47	0.67	0.02	0.48	0.39	0.01	0.40
			$a_{0.5}$	0.46	0.67	0.02				
			a_1	0.34	0.98	0.02	0.33	0.95	0.02	0.28
Screening t_2	21	1,009	a_0	0.47	0.56	0.02	0.48	0.28	0.01	0.40
			$a_{0.5}$	0.47	0.56	0.02				
			a_1	0.35	0.70	0.02	0.32	0.46	0.01	0.27
			$a_{1.5}$	0.34	0.72	0.02				
			a_2	0.25	0.96	0.03	0.21	1.00	0.03	0.06
Post-screening	19	900	a_0	0.47	0.71	0.03	0.48	0.46	0.02	0.40
			$a_{0.5}$	0.47	0.71	0.03				
			a_1	0.35	0.82	0.02	0.32	0.71	0.02	0.25
			$a_{1.5}$	0.34	0.82	0.02				
			a_2	0.25	0.94	0.02	0.22	0.86	0.02	0.05

Table 5.1: POMDP vs. physician performance, 5-fold cross validation using test data partition (average across runs presented). A kappa score was also calculated to compare the level of agreement between the model and experts.

interquartile range (IQR), and the range for each metric. This analysis is summarized in Figure 5.5, where the box plots depict the median and IQR of each action. Significance tests were performed using the Wilcoxon signed rank test or paired t-test as appropriate. Tests with p-values < 0.05 were deemed significant.

Interestingly, the POMDP model's CIs become narrower over time, suggesting that it

(a) False positive cases between POMDP and physicians.

(b) Early predicted positive cases between POMDP and physicians.
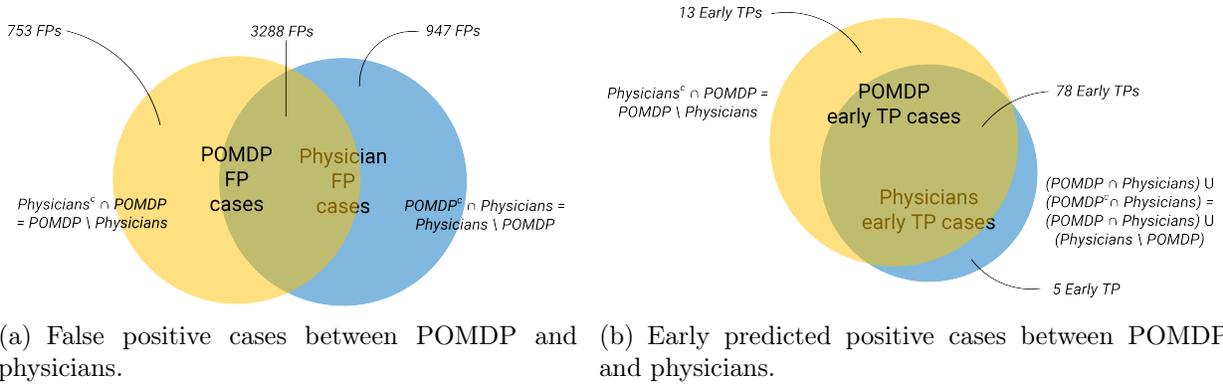
Figure 5.3: Comparison of case agreement between the POMDP and experts. The numbers in each subset represent the total number of FPs or early TPs grouped from every testing set for each fold of the five folds. **(a)** Yellow: Cases predicted as false positives by the POMDP model. Blue: Cases predicted as false positives by the physicians. The union of these groups are all cases predicted by the POMDP or physicians as false positives. $POMDP^c$ represents the complement of the POMDP set. **(b)** Yellow: Cases predicted as early true positives by the POMDP model. Blue: Cases predicted as early true positives by the physicians.

Variables	False positives analysis		Early true positives analysis	
	POMDP	Physicians	POMDP	Physicians
Nodule size t_2 (mm)	**4.73**[+]	**3.35**[+]	5.17	5.24
Nodule size t_1 (mm)	3.8	3.61	3.17	5.79
Nodule size t_0 (mm)	**2.47**[++]	**3.86**[++]	**1.54**[+]	**4.82**[+]
Years of smoking	**41.64**[+++]	**35.52**[+++]	**40.31**[+]	**45.33**[+]
Years since quitting smoking	**2.79**[++]	**5.21**[++]	4.77	2.29
Age at baseline	**61.8**[++]	**58.68**[++]	62.92	64.89
Smoking status at baseline (% smokers)	**60.29**[+]	**32.52**[+]	46.15	65.38

Table 5.2: Feature analysis of cases different between the POMDP and physicians, comparing false positives and early true positives. Reported values represent the post screening average values per variable. Bold values represent features with statistically significantly different distributions. The magnitude of the effect size of the p-value computed using the Cramer's r^2, the r^2, and the Cohen's r^2 are color-coded as: orange, small effect size ([+]); blue, medium effect size ([++]); and black, large effect size ([+++]). The Cramer's r^2, the r^2, and the Cohen's r^2 ranges for small, medium, and large are given in the Appendix.

stabilizes with longer trajectories of observations. When only testing the POMDP model on a cohort of cases with nodules larger than 6 mm at baseline, the POMDP model improves the true negative rate (i.e., reduces the FP rate) while maintaining a TP rate comparable to the physicians. Markedly, precision is significantly improved using the POMDP model in this scenario. When testing on the cases with nodules smaller than 6 mm at baseline,
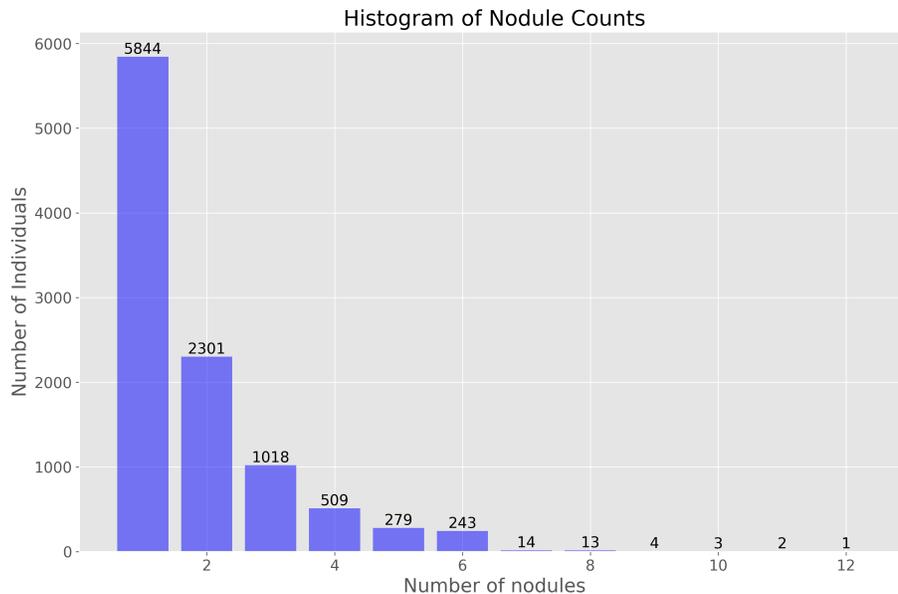
Figure 5.4: Histogram of nodule counts per NLST subject.

initially the POMDP TN rate is lower than that of physicians' but improves over time. The TP rate and early prediction of cancer is significantly improved compared with physicians in post-screening. Precision is also significantly improved for all screenings. This comparison of cases that are typically easier to classify as cancerous due to lager nodule size (i.e., ≥ 6 mm) demonstrates how our approach reduces FPs associated with lung cancer screening. Additionally, in the situation where IPNs are smaller (< 6 mm), our model still improves early prediction and overall precision. Box plots with the smaller than 6 mm and larger than 6 mm cohorts combined is presented in Figure A.24 in the Appendix.

5.4 Discussion

The majority of individuals diagnosed with lung cancer have a low 5-year survival rate of 18% [122]. In sharp contrast, earlier detection of this cancer improves this statistic threefold to 56% [122]. While LDCT lung cancer screening aims to reduce mortality through earlier detection, the FP rate associated with IPNs remains high, with concomitant concerns of increased healthcare costs and unnecessary psychological burden for patients. To address this concern, we developed a POMDP for lung cancer screening, demonstrating simultaneous reduction in FPs and earlier cancer detection when compared to experts' performance. Maintaining a high TP rate while minimizing the FP rate is challenging given the correlation of nodule malignancy and size: larger nodules tend to be malignant; and conversely, nodules smaller than 6 mm are less likely to be cancerous. We improved the TN rate for nodules larger than 6 mm at baseline while maintaining a true positive rate on par with experts. When comparing our POMDP against physicians' predictions for cases with nodules smaller than 6 mm, improved true positive rate and precision overall were seen, while progressively increasing the TN rate (see Figure 5.5).

Our POMDP uses a DBN to generate observations about a patient over time that are used to update its belief about lung cancer. We tested three variations of the POMDP,
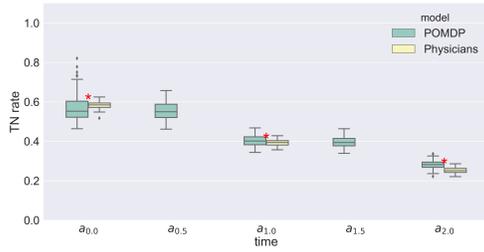
(a) TN rate on individuals with nodules smaller than 6 mm at baseline.
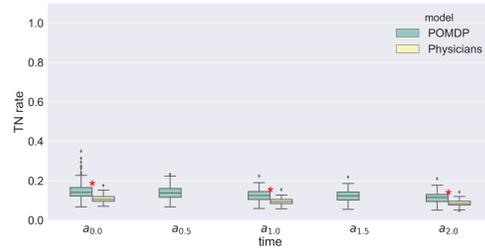
(b) TN rate on individuals with nodules larger than 6 mm at baseline.
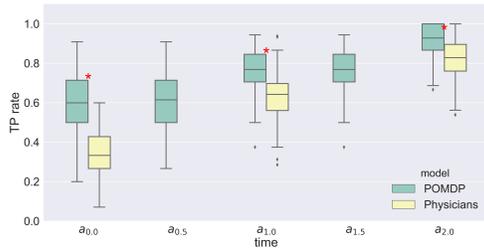
(c) Recall/TP rate on individuals with nodules smaller than 6 mm at baseline.
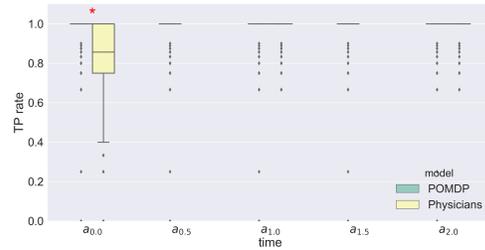
(d) Recall/TP rate on individuals with nodules larger than 6 mm at baseline.
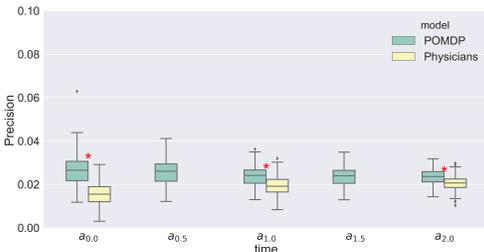
(e) Precision on individuals with nodules smaller than 6 mm at baseline.
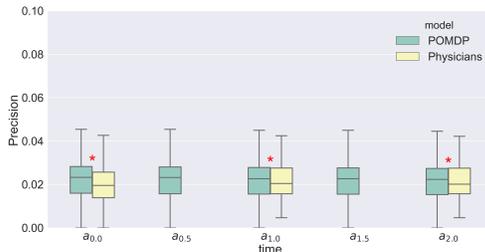
(f) Precision on individuals with nodules larger than 6 mm at baseline.

Figure 5.5: Box plots of the performance (TN, TP, precision) of the POMDP and physicians on cases from the NLST testing set, from the start of the trial through to last screening. **Left column**: Cases with nodules smaller than $6mm$ at baseline. **Right column**: Cases with nodules larger than $6mm$ at baseline. Blue and yellow represent the POMDP and experts, respectively. Red stars depict instances where the performance measure between the physicians and model are significantly different. The TN, Recall/TP rate, and Precision for the two cohorts combined is shown in the Appendix in Figure A.24.

considering observations as being independent over time (i.e., an exhaustive search of every combination of observations), as probabilities of a static regression model, and as probabilities derived from a DBN. Representing these variables dynamically via the DBN improved

model performance in comparison with the logistic regression model and performed similarly as exhaustive search. This analysis is presented in the Appendix in Table A.29). Modeling observations using a dynamic model has two main advantages: first, a dynamic model can capture changes over time in these features, which in our opinion are potential indicators of lung cancer; and second, it allows effective scaling of the observation space with the incorporation of multiple temporal inputs. While considering temporal change is intuitive, many lung cancer risk models are "static" and use only the most current observations when calculating the likelihood of disease. Still, such risk models are useful in baseline assessment: the initial belief for each case in our POMDP uses the Tammemägi model [114], instantiated using the subject's own demographic and clinical variables at baseline, updated with subsequent imaging observations.

The POMDP we designed makes use of a reward function learned through analysis of physicians' past decisions. We recently presented an adaptive maximum entropy inverse reinforcement learning (MaxEnt IRL) algorithm to inform a reward function in different cancers [110]. Using MaxEnt IRL, we established an optimization function explicitly modeling experts' actions. This strategy is different from other health-related POMDP applications [13, 15, 16] that typically employ cost functions based on quality-adjusted life years (QALYs), resource utilization, or other abstract metrics reflecting broader policy considerations. Building atop experts' prior actions, we take advantage of their experience and insights to integrate and weigh disparate information about a given individual; and by learning from multiple physicians and patients, we overcome potential biases. Yet curiously, per the diverging kappa score analysis, the POMDP is not fully replicating physicians' decisions. When it comes to early cancer prediction (e.g., predicting screening t_2 cancer from screening t_0), the POMDP outperforms experts, indicating that the model and associated reward function are discriminating between positive and negative cases in a different way. This difference may be attributed to the dynamic observation model used with this POMDP; when independent observations are instead assumed, we have found kappa scores to 1 in other domains, indicat-

ing high correlation between the model and experts' decisions [110]. Indeed, error analysis of the POMDP's FPs shows a different subset from the physicians: cases with smaller nodule sizes but more years of smoking and older baseline age are predicted as false positives by the POMDP. Early true positive cases share the same distributions, however, suggesting that a portion of POMDP false positives are early true positives. Table 5.1 illustrates this point in screening t_0 and post-screening for action a_0: 71% of TPs are being predicted from a_0 for post-screening cases – but if compared with screening t_0 cancer cases, they would have been classified as FPs.

Our previous work on predicting lung cancer in the LDCT screening setting showed encouraging results with earlier detection [106]. We showed that using a DBN trained on the NLST dataset we can match physicians' performance in predicting lung cancer, and in a portion of cases, in advance of the expert. But that method suffered from two limitations: first, the need to set an acceptable threshold for predicting lung cancer; and second, a decision-making process based solely on immediate outcomes without regard for longer-term benefits to the patient. We compared our current POMDP with our DBN [106], reproducing it on the same cohort of subjects used in this paper (i.e., using identical training and test sets and the same stratified five-fold cross-validation analysis). Even when setting different probability thresholds to generate performance metrics ($7 \cdot 10^{-6}$, 0.01, and 0.01 for each screening time point of the NLST study), our new POMDP-based approach outperformed the earlier model in terms of reducing the FP rate and improving early lung cancer prediction (see Table A.29 in Appendix).

Limitations of this work are around the real-world nature of cancer surveillance. It is unlikely that patients are screened at fixed one-year time intervals, for any number of reasons. As such, a discrete time model may not be well-suited for instances of imaging observations at irregular frequencies. Alternatively, a continuous time model may address this issue more accurately. We also used a simplified, expert-defined three-state cancer state space (e.g., no cancer, uncertain cancer, lung cancer); a more sophisticated approach would

97

involve learning this state space from the data, which we plan to explore in the future. Likewise, the observation space of our POMDP model is discrete, whereas a continuous value space might yield further improvements. This method can be explored through the use of linear Gaussian conditional probability tables (CPTs) instead of discrete observational CPTs. Lastly, the number of cancer and non-cancer cases changes as a function of time (i.e., more cancer cases are found at baseline). We did not account for this imbalance during training other than performing a stratified five-fold cross-validation to obtain an unbiased estimate of the model. Similarly, other temporal studies have used a k-fold cross validation to assess model performance [50, 51, 107–109]. This data imbalance over time occurred as a result of simplifying our lung POMDP model to consider only cases reporting a single pulmonary nodule over the course of the trial. A more concrete analysis would include cases with multiple nodules over time. However, it was not possible to ascertain the history of individual nodules in patients with multiple nodules as the NLST dataset does not contain sufficient tracking information on nodules. Moreover, the imputation of observations by our DBN observational model at six month intervals, even though it reduces over-screening, is inferred rather than based on true screening observations.

Future work includes conducting an external validation study of this NLST-based POMDP using data curated from our institution, expanding our observational model to consider multiple IPNs, as well as incorporating a richer set of imaging features derived from deep learning, which have demonstrated high classification performance in detecting malignant pulmonary nodules [123, 124].

5.5   Summary of findings

In this chapter, we described how to compute transition probabilities, observation probabilities, and initial cancer beliefs. Transition and observation probabilities were computed simultaneously using a DBN model. In addition, we described a way of using DBNs to

incorporate multiple imaging observations and their interactions in a single observation. We also described a way to define initial cancer beliefs using cancer risk models. We introduced a way to incorporate all technical contributions described in Chapters 4 and 6 to design a POMDP framework for optimal lung cancer screening decision making. The learned POMDP model was evaluated against physicians' decisions. This comparison demonstrated that the POMDP model can reduce the FP rate associated with lung cancer screening while improving early lung cancer prediction.

Chapter 6

Evaluating the Impact of Uncertainty on Risk Prediction: Towards More Robust Prediction Models

6.1 Overview

This chapter describes a new approach for imputing missing values. It is particularly burdensome to filter out cases with missing values for machine learning algorithms in training and test of cases involving small or heavily imbalanced datasets (i.e., reducing the minority class). This new approach is a multiple clustering imputation method built for breast cancer screening data with missing values. A variation of this methodology is used in the lung cancer POMDP model described in the previous chapter to impute missing values for the estimation of an individualized initial belief of lung cancer using the Tammemägi model. This chapter addresses the hypothesis that uncertainty is inherent in observational datasets. This is achieved through the use of a multiple clustering imputation approach used with the Gail model [125] by providing confidence intervals around risk prediction.

In breast cancer, models are used to estimate an absolute risk of cancer in women, which

influence decisions related to prescribing a risk-reducing pharmacologic intervention (e.g., selective estrogen receptor modulators) or more aggressive screening strategies (e.g., surveillance using breast magnetic resonance imaging). Tamoxifen is one example of a medication investigated for its effectiveness in the prevention of invasive breast cancer for high-risk women. The Breast Cancer Prevention Trial showed that women with a 5-year absolute risk of 1.67% and greater can reduce their risk of invasive cancer by 49% when undergoing chemoprevention compared to taking a placebo [126]. However, use of tamoxifen is not completely without risks and is associated with adverse events such as uterine cancer and blood clotting in the legs or lungs [127]. The purpose of these risk models is to provide physicians and patients with a reasoning tool to weigh the trade-offs between the effects of the intervention with the absolute risks of various health outcomes [128]. While risk prediction models aid in considering potential benefits and costs, these models also have notable limitations. First, models such as the Gail model provide an average risk for a group of women with similar risk factors, not an individual probability of cancer. As such, the interpretation of the predicted risk is unclear for a given individual. Second, uncertainty is an inherent part of risk assessment, given that not all factors related to cancer risk are known or can be measured to the desired precision. Studies have also shown that patient-reported information such as social history and patient outcomes are unreliable [129–131]. For clinicians who utilize cancer risk models to make decisions about potential interventions, an understanding about the sensitivity and reliability of self-reported risk factors such as the age of first live birth and family history should be known in the situation that such information is unreliable or missing. For example, heredity information is often complex to elicit from a patient, particularly if she is not completely aware of her siblings' and ancestors' health statuses. Additionally, any risk factor reporting age is often rounded up to the nearest year rather than the true age in months or days. Finally, information that is required to execute the risk model may not be available for a variety of reasons. Missing data are unavoidable in the fast pace, real-world clinical environment. Many models such as the Gail model are a form of a logistic regression

model that requires all risk factors to be inputted in order to compute the coefficients for the model or generate an estimated risk. If the patient cannot be subsequently reached to obtain the missing information, data-driven methods such as imputation must be performed to utilize these models. However, the effect of imputation on the validity of risk models has not been thoroughly explored in the medical literature [132,133]. For instance, datasets often suffer from population bias such as when the majority of patients are white. In the case of missing data, imputing instances of minority values from unbalanced datasets introduces bias as well as uncertainty in imputed values. In this work, we present a systematic approach to assess the effect of uncertainty and missing values on risk predictions. Comprehensive assessment of uncertainties in estimated risk metrics requires consideration of uncertainties about the input or parameters of the risk model (parameter uncertainty) as well as uncertainties associated with the form and assumptions of the model (model uncertainty) [134]. The scope of the present work is the treatment of parameter uncertainty. We utilize breast cancer screening as a driving example. Leveraging a large retrospective dataset of women undergoing routine screening, our approach discovers subgroups of similar women from which meaningful value ranges for a given risk factor can be determined. A clustering technique with multiple imputation is used to identify similar patients. Bootstrapping is then used to sample values of similar cases. These values are inputted into the Gail risk model to generate a confidence interval (C.I.) around the absolute risk prediction. We subsequently evaluate the sensitivity of the model to varying inputs. By expressing cancer risk using a C.I., we formalize how uncertainty is expressed, providing additional context to aid physicians in interpreting risk predictions and making management decisions. We analyze the frequency by which uncertainty associated with risk estimates would have potentially changed whether the patient would have been categorized as "high-risk" (e.g., cross the 1.67% risk threshold).

6.2 Background

6.2.1 Predicting absolute risk of breast cancer: The Gail Model

Cancer screening is a large population-based intervention that is at the center of great debate, especially in older patients or for certain cancers such as breast and prostate [125,135]. Breast cancer screening is particularly contentious. A number of models are in use today to stratify patients into different risk groups [136]. The Gail model is among the earliest and most widely used to estimate absolute risk. The model incorporates age, age at menarche, age at first birth, the number of first-degree relatives with breast cancer, the number of previous breast biopsies, and race in its assessment. The Gail model has been validated in specific cohorts of white American women with specific risk factors but has since been adjusted for individuals of different race and ethnicity. The model calculates the absolute risk of breast cancer by breaking the risk estimation into 3 sub-problems: 1) the estimation of the relative risk using a logistic regression; 2) the estimation of the baseline age-specific breast cancer hazard rate; and 3) the estimation of a long-term probability of developing breast cancer from competing risks, relative risk and the baseline hazard [3]. The Gail model was used to compute risk values for our test population. The probability that a woman at age a with a relative risk $r(t)$ will develop cancer by age a + t can be computed following Equation 6.2.1:

$$P(\alpha, \tau, r) = \sum_j \frac{h_{1j}r_j}{h_{1j}r_j + h_{2j}} \frac{s_1(\tau_j - 1)}{s_1(\alpha)} \frac{s_2(\tau_j - 1)}{s_2(\alpha)} (1 - \exp{-\Delta_j(h_{1j}r_j + h_{2j})}) \qquad (6.2.1)$$

where j is a defined age interval, h_2 is the risk of death due to other causes (competing hazards), s_2 is the probability of surviving the competing hazards, s_1 is the probability of surviving the death due to breast cancer, t_j is the time at the j-th age interval, α is the baseline age, and t is the time in years between baseline age and predicted age (typically set to 5 years). More information on the implementation of the Gail model can be found in [137–140].

6.2.2 Handling uncertainty in the data

Simulation-based methods such as Markov Chain Monte Carlo and bootstrapping have certain advantages compared to point estimate imputation methods when dealing with missing or uncertain cases. Even though computationally they are less efficient, they provide a confidence measure in their estimation making them more useful than a point estimate. They simulate possible uncertain values to generate a C.I. that represents the degree of uncertainty. Similarly, this approach can be applied when imputing missing values. Multiple imputation involves the simulation of a user-defined number of complete subsets m which are used to impute missing values. For each missing value, m possible imputed values are generated, reflecting the uncertainty about the true value of the variable. These m imputed values can be used to compute CIs [141, 142]. Another class of methods, model-based imputation, refers to estimating the joint distribution among risk factors from which imputed values are generated. To learn such a model, a training set is required to define the joint distribution. Imputation is then performed on a test set with missing values. Finally, clustering-based imputation approaches are used to identify similar cases, from which an imputed value for the missing variable is assigned [143]. These approaches are typically implemented using a combination of k-means and k-nearest neighbors (kNN) algorithms [144]. The kNN algorithm is frequently used to cluster cases using variables that do not have missing data, from which a set of values of similar cases can be obtained to inform the imputation process.

A significant limitation of these existing methods is the need to utilize complete information for training, which limits the number of cases that can be used. To account for uncertain and imputed values in our breast cancer dataset with the Gail risk model, we propose a multiple clustering imputation methodology that solves the limitations of traditional model-based imputation methods while providing a more informed breast cancer risk representation with CIs. Our proposed methodology imputes missing values from cases with complete information using multiple clusters of similar cases. Unlike methods that require complete data, our approach maximizes the use of available data, even the ones with missing

values. We use bootstrapping to calculate m unique clusters for each case with missing data. Using these m clusters, we generate a range of possible values for missing values, which is used to provide a CI of the imputed value. Subsequently, this range of imputed values can be used with risk models to generate a CI of risk values.

6.3   Methods

6.3.1   Dataset

Data on women who underwent breast screening at UCLA was obtained through an institutional review board (IRB)-approved protocol. The dataset consists of 47,980 cases collected during a five-year period. A detailed description of this dataset is shown in Chapter 4 section 4.2.3. The 5-year absolute risk for each case was calculated using the Gail model. The implementation of the Gail risk model that we used as part of this analysis did not adjust for Native Americans. Individuals who self-reported as part of this race category were excluded from our analysis. Individuals with multiple races were not excluded from our analysis.

6.3.2   Overall Approach

Our approach to investigating the influence of uncertainty is illustrated in Figure 2. We posit that using CIs defined by similar cases for certain input variables can change the interpretation of the absolute risk that is generated by the Gail model. The dataset was randomly split into training (60%) and testing (40%) sets, consisting of 28,788 and 19,192 cases, respectively. Categorical variables such as race were transformed into binary indicator variables, resulting in four variables representing each race and ethnicity categories. Within the testing set, we only considered cases that had complete information, resulting in a total of 1,850 cases. We focused our analysis on this subset of the test set. Missing values were simulated for each case using an unbiased random number generator. The number generation process consists of two random number generators, each producing a value from

0 to 8 (matching the number of input variables). Each random number generator was used to populate a list of 8 elements, corresponding to the number of variables. For elements in each list with the same random integer, the value for the corresponding variable was set as missing. The training set served as a knowledge base of retrospective cases that informs how missing values of the test cases could be imputed. Continuous variables such as current age and age at menarche were varied by ± 1 years. The variability introduced into age variables was constrained for two reasons: 1) we hypothesized that the likelihood of a patient getting her age incorrect was small and that the error was more likely due to rounding to the nearest year; and 2) imputation of age would be extremely difficult from the other variables collected. Subsequently, the multiple clustering imputation (MCI) method was used to identify plausible values for categorical variables and generate a range of possible imputed values for the continuous variables. Given the range of imputed values generated using the MCI method, the risk estimate for each case was calculated using the Gail model, yielding a distribution of 5-year absolute cancer risk predictions as well as a 95% CI around the median.
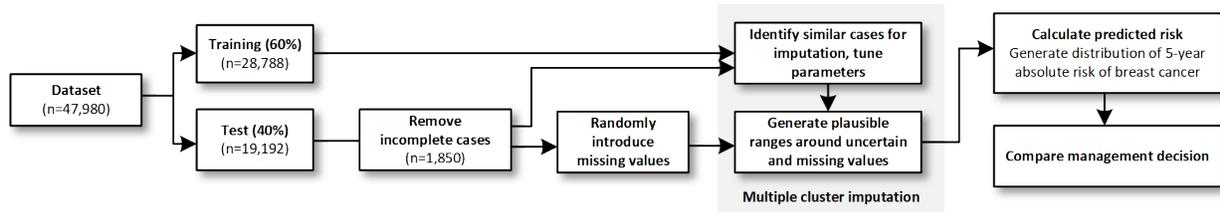
Figure 6.1: Overall approach. Process by which data collected on women undergoing breast screening were split into a train-ing and test set. The training set was used to perform multiple cluster imputation to generate ranges for uncertain or missing values introduced in the test set. These ranges were used to calculate a range absolute risk scores and interpreted using the current approach of identifying women with a 5-year absolute risk of 1.67% or above as can-didates for chemoprevention.

	B	W	A	H	Age	AM	AFLB	Biopsies	FDR_BC
Original data point	0	1	0	0	48	14	39	0	1
Missing Values	0	1	0	0	48	NaN	39	NaN	1
Imputed data point	0	1	0	0	48	(11.7-14.3)	39	0	1

Table 6.1: Data imputation. An example of an imputed case with the range of possible input values considered for the continuous variables and a point estimate for categorical variables. B: Black, W: White, A: Asian, H: Hispanic, AM: Age at menarche, AFLB: Age at first live birth, Biopsies: Number of biopsies, and FDR_BC: 1st degree relatives with breast cancer. NaN: corresponds to a missing value.

6.3.3 Multiple Cluster Imputation (MCI)

Figure 3 illustrates the basic process for generating values using the MCI approach. The algorithm proceeded as follows: for each test case being considered, variables with missing values are identified (Figure 3-A). From within the training set, we identify cases that had observed values for the variables that are not missing in the test case. A strength of our approach is maximizing the number of prior cases that are used in this process through the use of training cases that have missing information. As such, we will make use of cases that have varying levels of completeness for the variables outside of the ones being imputed. An iterative selection process is performed to generate clusters of cases based on how complete the cases are (Figure 3-B). We define a parameter called tolerance value to constrain the level of missing information that may exist in a cluster and used for imputation. For example, if the tolerance value is set to 80%, the algorithm would only select cases that have at least 80% of their variables with an observed value outside of the variables that are being imputed. In our case with nine variables, a tolerance of 80% permits only one additional variable to be missing. The entire training dataset is then examined for all possible combinations of variables where only one additional variable is missing. This process is repeated with increasing tolerance values until the value reaches 100%. This bootstrapping process generates multiple clusters of varying levels of completeness, with replacement, from which similar cases can be selected.

A ball tree algorithm [145] is used to select the most similar cases from each cluster com-
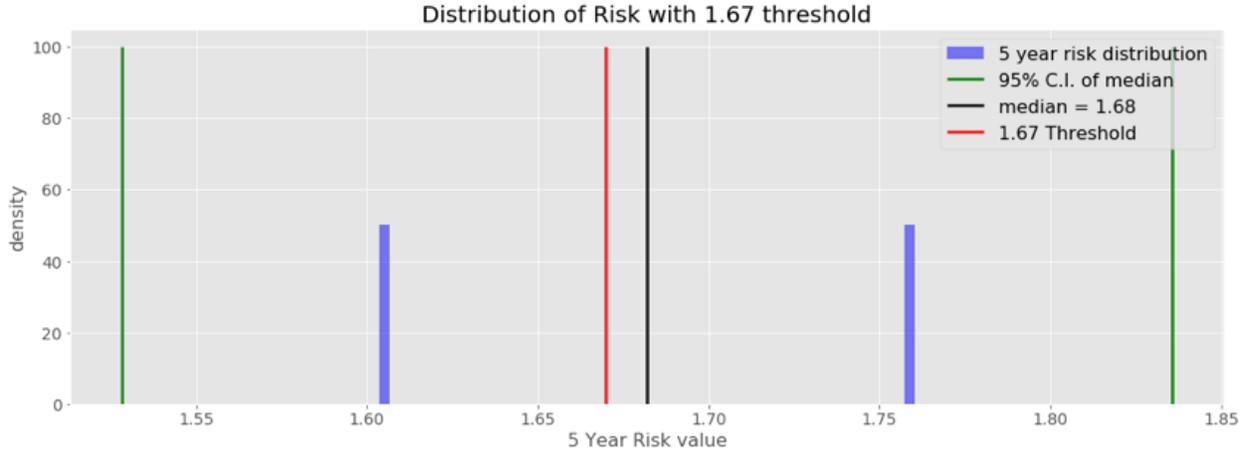
Figure 6.2: Example distribution of a single case. Distribution of risk predictions for a single patient based on 100 different simulated input. The original risk estimate for the individual is 1.60. The median value is 1.68, with 1.53-1.84 confidence interval.

pared to the test case (Figure 3-C). The ball tree is a binary tree where every node consists of a hypersphere that contains a subset of cases to be searched. The ball tree algorithm used is obtained from scikit-learn [146], and our analysis is implemented in Python. The radius of the hypersphere is user-defined and specified as an input parameter to the algorithm. All cases inside a hypersphere are considered as similar cases. As a space partitioning algorithm, the ball tree efficiently projects points/cases in a multi-dimensional space. The ball tree data structure is a hierarchical binary tree in which each node in the binary tree is split into two clusters with data points added in each cluster based on distance from the centroid of each cluster. In this work, the radius of the hypersphere is adaptive and proportional to the tolerance value: smaller tolerance values are associated with smaller hyperspheres. Once a set of cases have been identified from each cluster, then all the cases are combined into one group. These similar cases are used to impute missing values (Figure 3-D). Two types of variables are considered: continuous and ordinal/categorical. Continuous variables are imputed based on the range of values from similar cases in the training set. The minimum and maximum values in these cases define the range of permissible values; we then fit a normal distribution, taking the 50% CI of this distribution from the median. Conversely, ordinal

and categorical variables are imputed based on the most frequent value for a given variable. Table 2 provides an example of how imputation was performed on a case with two missing values. Imputed values and the range of imputed values are estimated for categorical and continuous variables missing values, respectively. Figure 4 depicts an example of a breast cancer risk with continuous variables variability and a risk confidence interval.
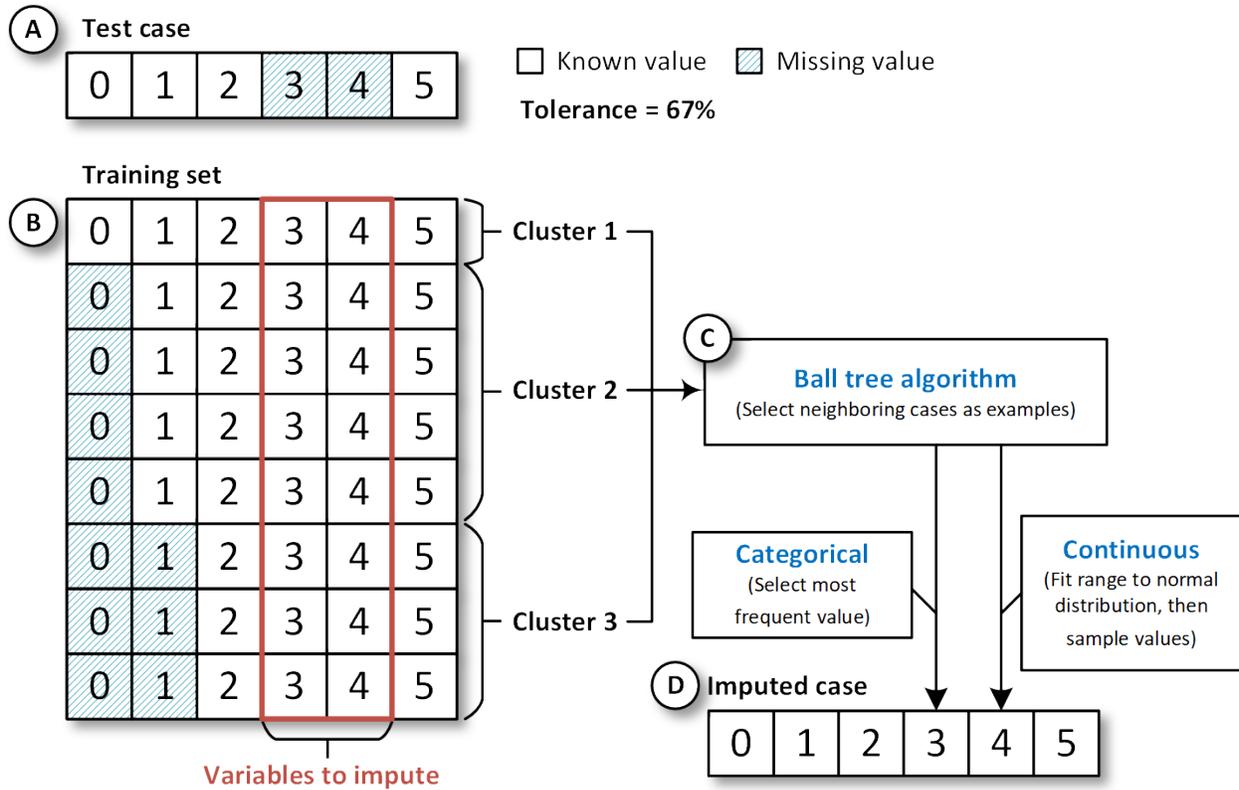
Figure 6.3: Multiple imputation clustering. Given a test case with two missing values (A), we examine the training set for cases that have values for variables 3 & 4 (B). Within that subset, cases are grouped into clusters based on the percentage of observed variables; all clusters will have observed variables above a predefined tolerance value. A ball tree algorithm (C) is used to select the training cases that are most similar to the test case; the range of values defines the permissible values from which the final imputed values are selected (D).

	Complete	No missing values with variability	Missing values with no variability	Missing values with variability
High-risk (HR)	579	462	531	427
Low-risk (LR)	1271	1167	1287	1205
Uncertain (U)	-	221	32	218
HR → HR	-	462	531	427
LR → LR	-	1167	1253	1174
HR → LR	-	0	34	31
LR → HR	-	0	0	0
HR → U	-	117	14	121
LR → U	-	104	18	97

Table 6.2: Summary of interpretation changes by introducing the C.I. associated with the risk prediction. The uncer-tain category highlights the cases where the original risk estimate was either above or below 1.67%, but when a range of possible input values is considered, the decision threshold falls within the C.I. of the risk estimate.

6.4 Evaluation

Using the generated predictions of risk, we performed two types of analyses. In the first analysis, we explored the impact of intentionally varying continuous variables such as current age and age at menarche when complete information was available (i.e., no imputed values) to determine the effect of rounding on predicted risk. Within the testing set, variability in known values of current age and age at menarche (collectively referred to as the continuous age variables) was introduced by calculating the risk based on 0.1 increments between -1 and 1 years from the inputted value, resulting in 20 risk estimates. In the second type of analysis, we examined the effect of imputing missing values using the MCI approach and the effect of intentionally varying continuous age variables. We fitted the continuous variables of all the similar cases on a normal distribution around the median and defined the range as the 50% CI around the median. This range was also split into ten linear steps. Overall, we present four analyses: the original complete dataset without introducing variability on the continuous age variables ("complete"), the original complete dataset with variability introduced on the continuous age variables ("No missing values with variability"), the imputed dataset without variability introduced on the continuous age variables ("Missing values with no variability"), and the imputed dataset with variability introduced on the continuous age variables, if ranges

were not already imputed due to the value being missing ("Missing values with variability"). We evaluated how often the risk model resulted in a predicted absolute cancer risk that would change the management of a patient (e.g., the risk range predicted for each test case crossed the 1.67% threshold). We also evaluated which combinations of feature values would change the categorization of a given patient (e.g., if the patient moves from low-risk to high-risk).

6.5 Results

6.5.1 Implication of risk predictions under uncertainty

The MCI method was used to impute the test set of 1,850 cases. Table 3 summarizes changes in management interpretation when CIs surrounding a risk prediction is provided. In the "complete" column, there are 579 and 1,271 high and low-risk predictions, respectively. When variability was introduced in the continuous age variables, as summarized in the "no missing values with variability" column, the risk category of 221 cases changed from high- or low-risk to "uncertain" given that the CI overlaps with the decision threshold of 1.67%. Out of those 221 cases, 117 were originally high-risk, and 104 were low-risk individuals. No cases changed status form high to low-risk or vice versa. When we randomly introduced missing values into the test data, represented by the "missing values with no variability" column, the total number of uncertain individuals were 32 of which 14 were originally high-risk, and 18 were low-risk. 34 cases changed category from high-risk to low-risk. No cases changed from low- to high-risk. In the "missing values with variability" column, a higher number of high-risk cases changed to uncertain cases compared with other columns. Of the 218 uncertain cases, 121 were previously high-risk, and 97 were low-risk. Additionally, 31 cases that were originally high-risk changed to low-risk. No low-risk cases changed to high-risk. Analysis of Risk Predictions under Uncertainty In the "no missing values with variability" analysis, when age and age at menarche were used independently (varied one at a time), the number of uncertain cases was 96 and 124 for current age and age at menarche, respectively. Age

at menarche had a stronger impact on predicted risk than current age. Table 4 summarizes the average value for each variable, stratified by risk group (HR, LR, U). In addition, several trends that reinforce prior findings were noted: 1) older women were associated with a higher risk of breast cancer; 2) women who started menarche at an older age were associated with a lower cancer risk; 3) the number of biopsies was proportionate with risk; and 4) women with more 1st-degree relatives with a history of breast cancer had higher risk themselves. Uncertain cases had average values for variables in-between average values found in high and low-risk groups. Moreover, when missing values were introduced in the analysis, the number of uncertain cases increased when variability was introduced in the continuous age variables. We estimated the percentage of missing values per variable in the high to low-risk, low to high-risk, and high or low to uncertain risk groups. The variables with the highest percentage of missing values in the group that changed from high- to low-risk were and the number of 1st-degree relatives with breast cancer, the age at first live birth, and the number of prior biopsies, in descending order. The main variables with the highest percent-age of missing values in the groups that changed from high- or low-risk to uncertain were the age at menarche and the number of prior biopsies.

6.5.2   Availability

We have made our analysis available in the form of Jupyter notebooks[1].

6.6   Discussion

In this chapter, we examine the effect of uncertainty on the input values of the Gail model when estimating risk. In addition, we evaluated an approach for imputing a range of missing values for a patient to generate an individualized breast cancer risk CI, using previously observed cases. While many of the variables collected as part of the Gail model are straight-

---

[1]https://github.com/panas89/multipleClusteringImputation

		B	W	A	H	Age	AM	AFLB	Biopsies	FDR_BC
Complete	HR	0.07	0.78	0.09	0.08	60.35	12.63	27.36	0.5	0.56
	LR	0.07	0.65	0.17	0.17	46.07	12.91	21.97	0.07	0.09
No missing values with variability	HR	0.06	0.78	0.1	0.09	60.89	12.61	27.51	0.56	0.65
	LR	0.08	0.65	0.17	0.17	45.24	12.91	21.57	0.06	0.08
	U	0.07	0.75	0.11	0.1	56.85	12.8	26.61	0.22	0.24
	HR → HR	0.05	0.84	0.06	0.06	60.93	12.65	27.01	0.54	0.64
	LR → LR	0.08	0.64	0.17	0.18	44.68	12.88	21.61	0.05	0.08
	HR → LR	0	0.73	0.22	0.05	62.59	13.76	20.51	0.24	0.11
	LR → HR	0.11	0.06	0.53	0.39	57.28	12.31	32.44	0.19	0.42
	HR → U	0.04	0.85	0.09	0.06	59.04	12.79	24.6	0.22	0.22
	LR → U	0.14	0.58	0.16	0.19	53.07	12.81	30.07	0.21	0.26
Missing values with no variability	HR	0.06	0.82	0.1	0.1	60.06	12.79	25.43	0.55	0.58
	LR	0.06	0.67	0.16	0.18	45.46	12.96	19.91	0.07	0.07
	U	0.03	0.88	0.06	0.06	57.91	13	24.69	0.09	0.25
	HR → HR	0.05	0.87	0.05	0.07	60.38	12.82	25.22	0.55	0.57
	LR → LR	0.06	0.66	0.16	0.18	44.6	12.96	19.92	0.06	0.07
	HR → LR	0.02	0.77	0.19	0.05	58.18	13.05	19.74	0.31	0.07
	LR → HR	0.13	0.31	0.57	0.43	55.25	12.6	27.78	0.19	0.5
	HR → U	0	0.95	0.1	0	58.76	13	27.38	0.05	0.29
	LR → U	0.09	0.73	0	0.18	56.27	13	19.55	0.18	0.18
Missing values with variability	HR	0.05	0.79	0.09	0.09	61.36	12.64	24.88	0.54	0.6
	LR	0.07	0.68	0.15	0.16	45.49	12.94	19.2	0.05	0.07
	U	0.08	0.75	0.1	0.1	58.05	12.75	24.37	0.22	0.19
	HR → HR	0.05	0.85	0.05	0.07	61.42	12.68	24.34	0.51	0.6
	LR → LR	0.07	0.67	0.15	0.17	44.69	12.92	19.26	0.04	0.08
	HR → LR	0.03	0.76	0.18	0.08	58.71	13.22	18.17	0.18	0.06
	LR → HR	0.1	0.13	0.55	0.29	57.84	12.36	31.02	0.19	0.36
	HR → U	0.03	0.86	0.07	0.06	60.03	12.8	22.15	0.21	0.19
	LR → U	0.17	0.54	0.17	0.17	54.36	12.67	28.49	0.23	0.18

Table 6.3: Summary of high-risk (HR), low-risk (LR), and uncertain (U) cases average feature values. Left: Binary variables' mean frequency in each risk group. Right: Continuous/ordinal variables' mean value in each risk group. B: Black, W: White, A: Asian, H: Hispanic, AM: Age at menarche, AFLB: Age at first live birth, Biopsies: Number of biopsies, and FDR_BC: 1st degree relatives with breast cancer.

forward to provide, risk models are becoming increasingly complex, and the impact of uncertainty or invalid data should be explored. For example, breast cancer risk models such as Tyrer-Cuzick [147] and BRCAPRO [148] ask for a detailed family history of cancer from first-, second-, and even third-degree relatives, which may be difficult to report precisely. A better understanding of how uncertain or unreliable inputs into these risk models is needed to better inform subsequent management decisions on whether a patient is considered "high-risk" or not. From our analysis, we conclude that uncertainty in input and missing values can potentially change the risk category of an individual when using the Gail model. In-

terestingly, throughout the four analyses shown in Table 3, low-risk cases never changed to high-risk, implying that the Gail model is more robust to low-risk uncertainty than it is to high-risk (high-risk cases being downgraded to low-risk). In Table 4, we demonstrated that a significant number of uncertain cases was classified primarily due to uncertainty in variables such as current age and age at menarche. Additionally, the majority of cases classified as uncertain were primarily missing values such as the age of first live birth and number of biopsies. Cases that were classified as "low-risk" but were actually "high-risk" upon further analysis had missing values for age at menarche, the number of biopsies, and breast cancer history among 1st-degree relatives.

Our study has several limitations. The imputation approach had difficulty providing reasonable estimates for current age; therefore, we chose not to introduce missing values to that variable. We believe the information provided by the other variables was not sufficient to provide meaningful estimates of the age variables. We also assumed that the distribution of continuous variables was normal. For example, the variable age at first live birth had zero values for women without a first birth. A normal distribution was not suitable for this variable; hence it was instead modeled as an ordinal/categorical variable. Future work may consider additional clinical risk factors that could serve as surrogate measures. In addition, while we employed and examined the effect of varying model parameters such as the range of the continuous variables' values and tolerance values, a full search was not performed, hence the performance of the algorithm may not be optimal. We introduced missing values at random into the dataset, but values were frequently missing not at random in real-world scenarios. Bias could be introduced into the missing value generation by adding weights to specific variables that are more frequently missing in practice. We also weighted variables equally when using the MCI method; future work can examine how these weights can be customized for individual variables.

While variables such as current age should be readily accessible, this analysis underscores the need to ensure that all of these variables are accurately recorded, given their impact on

the final risk estimate. Cases that were unchanged generally had low percentages of missing values for all variables and any variability introduced on the continuous age variables had no effect on risk as their values were either very high or low. Our work highlights the utility of conducting sensitivity analyses as part of validating risk prediction models. Furthermore, we believe that reporting of CIs may be more informative than simply interpreting a point estimate of risk. Several studies have shown the utility and potential challenges of representing uncertainty associated with risk predictions to decision makers, including clinicians and patients [149–151]. By conveying the risk as a distribution, clinicians can understand the uncertainty associated with a risk estimate and better determine whether the patient's situation is clearly "high-risk" and should be given risk-reducing interventions or "uncertain" and should undergo further testing. Narrower CIs imply less variability (more confidence) in risk estimate and vice-versa.

6.7   Summary of findings

In this chapter, we present a novel imputation methodology that addresses the inherent uncertainty in data and missing values. This is accomplished using multiple clustering from different random samples of the training dataset with replacement (i.e., bootstraping) and a k-ball tree algorithm that samples similar cases from each cluster to generate a range of plausible values for uncertain or missing values. Subsequently, this range is used to generate an individualized distribution of risk using risk models. The 95% confidence interval of each distribution is used to depict the uncertainty in data. This methodology was built and evaluated on the Athena breast cancer screening dataset with the Gail model, and applied on the NLST lung cancer screening dataset with the Tammemägi risk model.

Chapter 7

Conclusion

7.1 Overview

This chapter summarizes the results and findings of this dissertation. The limitations and potential future research avenues of this work are outlined below.

7.2 Summary of contributions

This dissertation presents methods that can help with more individualized sequential decision making in clinical environments. Specifically, I addressed the need for methods and tools to optimize lung cancer screening by improving early lung cancer detection, while reducing the false positive rate associated with imaging-based LDCT screening. My approach was to develop methods that efficiently learn POMDP models from data. The specific contributions of this dissertation are summarized as follows:

- A dynamic Bayesian network method to individualize the lung cancer screening process. I presented a novel DBN that models the hidden lung cancer disease state space and can be used to provide the probability of a positive biopsy over time. I also described how to add domain knowledge into the design of this DBN.

- A robust multiple clustering imputation approach for the imputation of missing values. I developed a novel approach that overcomes clustering methods limitations by using methods like bootstrapping to impute missing values and to depict the inherent uncertainty in clinical values and imputed missing values.

- An adaptive maximum entropy inverse reinforcement learning algorithm. I described how to use expert's temporal decisions to learn state and action rewards for POMDPs and MDPs and improved the convergence of this algorithm to a unique solution using an adaptive step size. Using a multiplicative model I used the state and action rewards to define POMDP reward functions used in lung and breast cancer screening.

- A lung cancer POMDP model to individualize the lung cancer screening sequential decision making process. I showed how to design an observational model that captures the interactions of LDCT imaging observations and how to define the initial belief of cancer using a patient's demographic and clinical variables.

I performed an extensive evaluation to identify a suitable DBN structure for lung cancer screening. The evaluation involved the sequential prediction of whether a patient should perform biopsy (predicting the probability of a positive biopsy) or not against the cancer status of a patient over time. Five different (three structural, two expert-defined) models were trained on balanced datasets and externally validated on the remaining cases of the LDCT arm of the NLST trial ($\sim$25,000 cases) with performance comparable to experts. Through this work we gained the following insights in modeling dynamic structures. We noticed that the main factor for driving models' performance was the temporal structure presented in Chapter 3, as the structure-learned models did not improve model performance. This finding motivated the selection of a dynamic model that follows the same dynamic structure and takes into account patient long-term benefits and disease outcome for the selection of appropriate screening decisions. As such we used the cancer state space of the aforementioned DBNs to design a POMDP model. To build a POMDP model that suggests sensible

recommendations for screening the definition of an appropriate reward function is imperative. I explored the use of the MaxEnt IRL algorithm for this purpose. I used experts' prior screening recommendations with this algorithm to learn a reward function for POMDPs. The MaxEnt IRL algorithm was used to learn state and action rewards. A reward function is defined by state-action pair rewards. The state and action rewards were used with a multiplicative model to demonstrate the difference in utility in performing each action in each state. This methodology was evaluated in lung and breast cancer screening, with performance comparable to experts. To deal with missing data, I developed a new multiple clustering imputation approach for handling missing values and representing uncertainty in data using confidence intervals. This approach combines the advantages of multiple clustering approaches and bootstraping while imputing missing values without exclusion of cases with missing values in variables other than the ones to be imputed. Collectively, these techniques were used to inform the scalable creation of a lung cancer screening POMDP. In practice, the observation matrix of POMDPs is discrete and tabular (i.e., each column of the matrix represents a unique observation) and thus is limited computationally to a fixed number of observations. I implemented a DBN model that compresses all observations over time as a probability of disease. Even though the observation matrix of this application is discrete I have laid the foundations into extending it in a continuous probability observation matrix. Given such an observation matrix the observational model used to generate probability observations can be replaced with any model that uses imaging or other temporal features to predict cancer. Furthermore, through this work I identified that the transition matrix of this POMDP is action invariant and learning the transition and observation matrix of a POMDP from data can be accomplished through the use of a naïve Bayes DBN model.

7.3 Limitations and future work

The work in this dissertation has limitations as well. Below, a sample of such limitations and improvements are briefly discussed.

In Chapter 3, the developed *Forward-Arrow* and *Reverse-Arrow* DBNs use the physicians' LDCT imaging interpretation as temporal imaging features. A more accurate analysis would use LDCT imaging findings. As such, in my proposed observation model DBN, in Chapter 5, I used imaging fidings to learn the structure and parameters of the DBN. In addition, the POMDP reward function is computed using state visitation frequencies with the MaxEnt IRL algorithm – but using the value iteration algorithm, more than one combination of rewards can potentially define the same problem. To overcome this, a policy iteration algorithm can be used rather than value iteration to learn optimal policies, as the policy space is finite in comparison to the rewards space (hence the policy iteration algorithm is guaranteed to optimally converge). Future work can also involve the exploration of transfer learning to reuse the weights learned using the MaxEnt IRL algorithm between other datasets and domains.

The POMDP state and observation spaces are discrete, but a continuous space would model states and observations more accurately. This approach may be achieved using Gaussian conditional probability tables. Moreover, in this dissertation I described a way to learn a representation of the observation space although our 3-state cancer state space is expert-defined. Future work should expand this observational model to consider multiple IPNs, as well as incorporating a richer set of imaging features derived from deep learning, which have demonstrated high classification performance in detecting malignant pulmonary nodules. In addition, POMDPs in healthcare would benefit from methodologies that can learn a representation and structure of the state space from data. Nevertheless, the more complex and larger a POMDP model in terms of states, actions, and observations, the more computationally intractable it becomes. Using approximation algorithms such as QMDP [66] to solve such models efficiently disregards the state uncertainty in diseases something that is

preserved through the use of exact POMDP solvers.

The transition model of the POMDP is discrete in time, while lung and breast cancer screening tests occurred roughly at 1-year intervals. Yet assuming that screening occurs annually (i.e., at fixed frequency) does not model the data generation process accurately. Building on existing work in efficient learning of continuous-time hidden Markov models for disease progression [152] the transition matrix of POMDPs can be modeled as continuous. A comparison of discrete models at higher time resolution (i.e., screening in months or days) against continuous time POMDPs is warranted.

Regarding evaluating the POMDP model, the low number of malignant nodules did not allow us to have a hold-out external validation test set. Instead, we performed a five-fold cross validation to obtain an unbiased estimate of performance. Future work includes conducting an external validation study of this NLST-based POMDP. Moreover, I trained and evaluated the POMDP model on single nodule trajectories, while in lung cancer patients may have multiple nodules. This limitation should be handled in future work by combining imaging observations from multiple nodules simultaneously and making action recommendation on all nodules.

This work can benefit from semi-supervised learning. Using existing unlabeled data, the use of the expectation-maximization algorithm can be leveraged to learn from unlabeled data. For instance, consider the expectation (E) and maximization (M) step of this algorithm. The E step: given that observations are probabilities of cancer, using imaging data and a pre-trained model to predict the malignancy of nodules from raw images we can generate observations for an unlabeled dataset. Subsequently, using a pre-trained POMDP model we can generate a belief and recommended actions for each patient over time. The belief will be used to predict the state of the patient (i.e., state with the highest probability). For the M-step: using the observations, states, and actions from the E-step, the POMDP can be retrained with additional data. The EM-algorithm will converge to a solution once the performance of the POMDP does not change.

Lastly, the lung cancer screening POMDP model performs binary decision making. Moving forward a treatment-based POMDP should be formulated to fully grasp the utility of temporal credit assignment with multiple treatment decisions over time. Assigning optimal treatment decisions can demonstrate the utility of learned rewards function from data to improve lung cancer survival. A potential extension of this work is thus to design a treatment-based POMDP so that when a patient leaves the screening process, he/she will enter the treatment process with the screening agent's insight. A pipeline that combines these two models may improve early lung cancer prediction and lung cancer survival.

Going beyond lung cancer screening, many patients suffer from numerous chronic diseases, simultaneously, such as diabetes or kidney disease, and are administered multiple treatments. Currently, in medical practise, there is no real assessment between disease experts to account for multiple treatments complications for future treatments regiments when administered without long-term benefits in mind. Designing multi-agent systems that can recommend optimal individualized treatment decisions given a holistic knowledge about a patient's health status and long-term benefits is the next frontier. The competing agents will suggest treatment or screening policies given an individual's diseases. At each epoch the agent with the highest short- and long- term individualized benefit (e.g., disease treatment, minimum complications, early disease detection benfit) will be used to suggest a policy.

7.4   Concluding remarks

Screening with LDCT has the potential to detect lung cancer at an early stage; and when detected earlier, more choices for treatment are available, along with improved chances of survival. However, LDCT screening is associated with a high false positive rate. This dissertation demonstrates a lung cancer screening decision making methodology that helps to address this challenge by reducing false positives while maintaining a satisfactory true positive rate. Additionally, a statistical analysis of false positives and early true positives

underscored which covariates are indicative of lung cancer. This methodology learns an agent that can be used for efficient decision recommendations.

Appendix A

Appendix

A.1 Eligibility criteria

The eligibility criteria used to obtain the complete set of $25,846$ cases from the CT arm of the NLST dataset were: 1) the participant to be eligible to participate in the NLST trial in terms of the NLST eligibility criteria (e.g., age between 55-74 years old); 2) the participant's last contact status to be either active or deceased; and 3) the participant's case to be neither withdrawn or lost.

A.2 Variables

Variables used from the NLST data and the associated categories/discretizations in the dynamic Bayesian network are as follows:

Variable Name	Description	Discretization
Age	Age of the individual	Under 60 years old; Between 60 and 70 years old; and More than 70 years old
Gender	Gender of the study subject	Male, female
Smoking status	The smoking status of the individual at the outset of the NLST.	Yes, no
Body mass index (BMI)	Height/weight ratio of the individual at the start of the NLST	Underweight, normal, overweight, obese
Cancer history	Specifies if the individual had a prior history of bladder, breast, cervical, colorectal, esophageal, larynx, lung, nasal, oral, pancreatic, pharynx, stomach, thyroid, or transitional cell cancer.	Yes, no
Disease history	Boolean variable representing the individual's history of diagnosis of asthma (adult or childhood), COPD, emphysema, fibrosis of the lung, sarcodosis, or tuberculosis.	Yes, no
Work history	Represents work-based exposures related to the development of lung cancer, including asbestos, coal, and other chemicals.	Yes, no
Family history of lung cancer	Boolean variable indicating if an immediate family member (parent, sibling, child) was previously diagnosed with lung cancer.	Yes, no
Cancer	This variable represents the state of the individual to have a suspected lung cancer, based on Figure 3.1.	
LDCT	The outcome of the imaging study for the individual, based on radiologist interpretation.	Screening with abnormalities detected and growth since prior study; Screening with abnormalities detected but no growth or change since prior study; no abnormalities
Biopsy	The results of a diagnostic biopsy.	Positive, negative
Death	Boolean variable giving the probability of death.	Yes, no

A.3 Prediction of future cancer cases

	DBN A Predictions		DBN B Predictions	
First Screening	150 (tp)	47 (fn)	121 (tp)	76 (fn)
	124 (fp)	77 (tn)	64 (fp)	121 (tn)
Second Screening	58 (tp)	76 (fn)	58 (tp)	76 (fn)
	20 (fp)	172 (tn)	19 (fp)	172 (tn)
Third Screening	45 (tp)	58 (fn)	45 (tp)	58 (fn)
	13 (fp)	175 (tn)	13 (fp)	175 (tn)

Table A.1: Top: Contingency table that represents an evaluation of the DBN predictions from the first screen with all cancer cases in the trial in the 10 random balanced test sets, including the cancer cases of the first screening. Middle: a Contingency table that represents an evaluation of the DBN predictions from the second screen with the remaining cancer cases in the trial, including the cancer cases of the second screening. Bottom: A Contingency table that represents an evaluation of the DBN predictions from the third screen with all the remaining cancer cases in the trial, including the cancer cases of the third screening. The 150 true positive cases shown above on the first screening of DBN A, consist of the 51 true positives predicted by the model in the first screening evaluation without taking into consideration the remaining cancer cases of the trial. By including the additional future cancer cases the DBN is able to predict an additional 99 cancer cases which in the initial evaluation were considered as false positives. This means that the majority of false positives predicted in the first screening in future screenings are true cancer cases.

A.4   Calibration curves
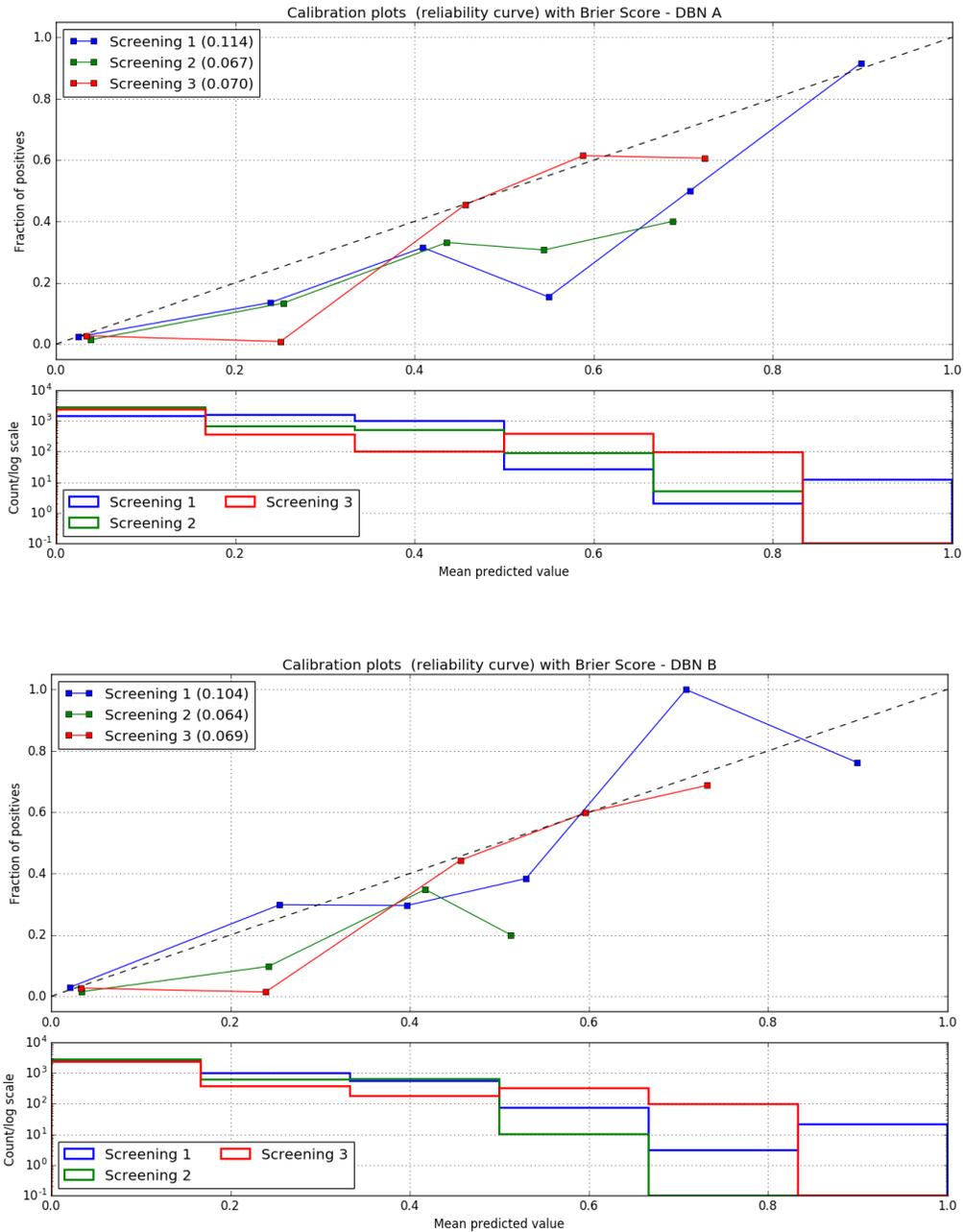
Figure A.1: The calibration curves of the DBN models for each screening as well as the Brier Score. The Brier score decreases with time between screenings. Bottom: Histogram of the positive cases over the probability of a positive Biopsy for each screening.

A.5   The DBN networks

A.6   Statistics

We present the results of the performance of each DBN structure over the same random balanced test sets of 400 cases (200 cancer and 200 non-cancer cases). All DBNs were trained on balanced training sets of 400 cases (200 cancer and 200 non-cancer cases). The thresholds used in these evaluations are 0.04, 0.21 and 0.25 for each screening, respectively.
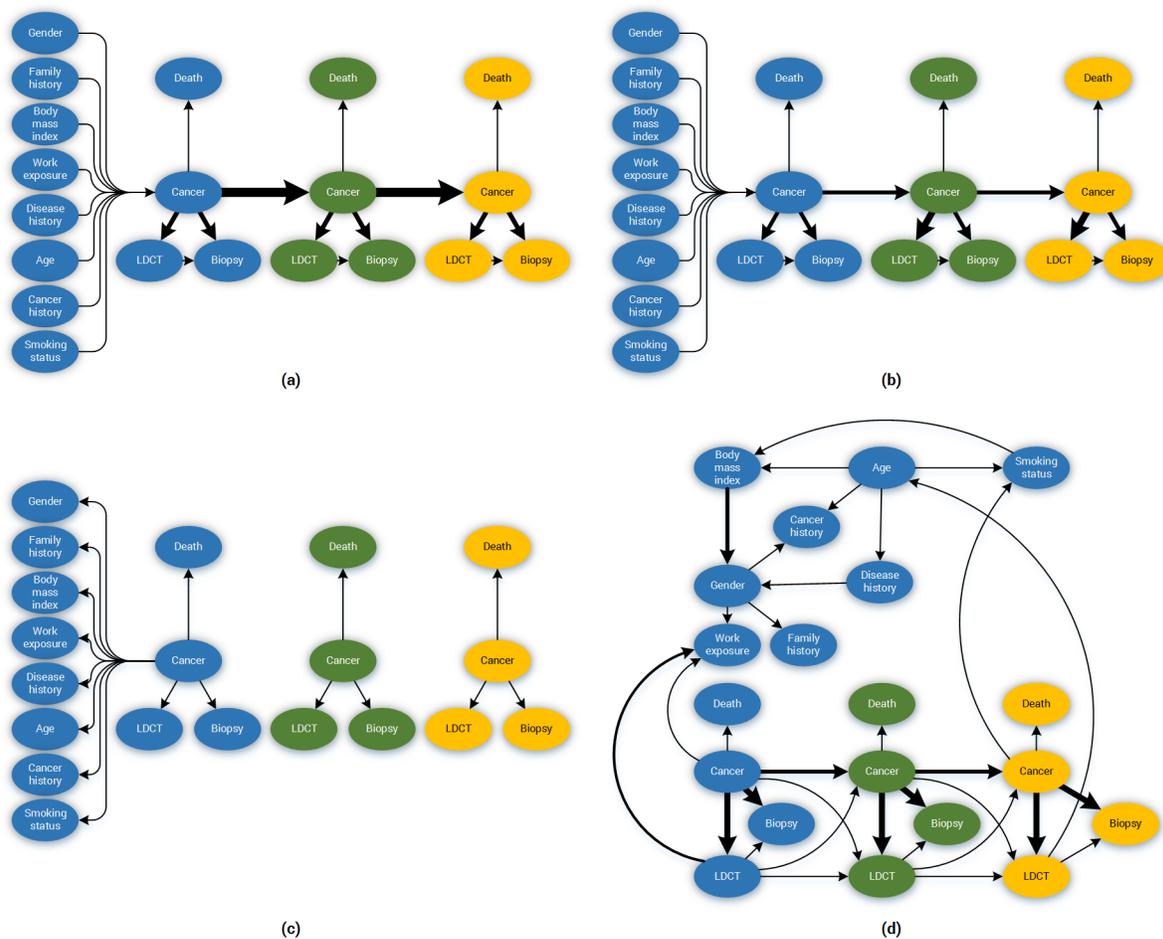
Figure A.2: The network Structure and the strength of influence depicted by the arrow thickness connecting the two variables. (a) The *Forward-Arrow* DBN without the NoisyMax gate; (b) The *Forward-Arrow* DBN with a NoisyMax gate as a cancer node at $t = 0$; (c) The *Reversed-Arrow* DBN; (d) The Learned Network with compositional nodes. The Learned DBN without compositional variables is not depicted due to the high complexity in structure.

A.6.1 The *Forward-Arrow* DBN without a NoisyMax gate

The *Forward-Arrow* DBN without a NoisyMax Gate was not parameterized using the EM algorithm. All nodes CPT tables' probabilities were empirically estimated from the dataset observations except from the Biopsy (abstracted from literature) and Death (death rate at baseline) nodes which were fixed nodes and the Cancer node at baseline. The Cancer variable would be impossible to parameterize without imposing some domain assumptions about an individual's cancer state as this node consists of 2304 parameters and 3 states (Non-cancer, In Situ, Invasive Cancer). The data do not contain sufficient observations to represent every single parameter (i.e., combination of parent state to effect node state). We dealt with this parameterization problem by using the following two assumptions. First, we assumed that every state combination with no instances in the In Situ or Invasive Cancer state in our data would imply that the majority of instances are in the Non-cancer state. Second, when we had data instances for either the Situ or Invasive-cancer state, we computed the probabilities of those states and assumed that the remaining cases were in the Non-cancer states (i.e., probability complement). The reason we pursued this parameterization approach is that most existing training algorithms do not support the use of missing data (e.g., dead patients with no observations in subsequent screenings). For example, EM would be a more appropriate algorithm in the case of missing values (i.e., missing value of age or BMI). In such a case an EM algorithm would instead estimate a statistical estimate of that value. We believe it would be undesirable to estimate the disease status of a deceased individual

Structure Learning	
Dataset number of cases	25046
Learning Algorithm	Bayesian Search
Algorithm Parameters	
Max parent count	8
Iterations	20
Sample size	50
Seed	0
Link Probability	0.1
Prior Link Probability	0.001
Background Knowledge	
Forced Arcs	5
Nodes assigned to tiers	6

Table A.2: Structure learning algorithm parameters.

The *Forward-Arrow* DBN without a NoisyMax gate						
	Screen 1		Screen 2		Screen 3	
Rates	0.927 (tp)	0.073 (fn)	0.903 (tp)	0.097 (fn)	0.854 (tp)	0.146 (fn)
	0.347 (fp)	0.653 (tn)	0.228 (fp)	0.772 (tn)	0.139 (fp)	0.861 (tn)
Counts	51 (tp)	4 (fn)	28 (tp)	3 (fn)	35 (tp)	6 (fn)
	119 (fp)	224 (tn)	67 (fp)	227 (tn)	35 (fp)	216 (tn)

Table A.3: The tp, fp, tn, fn rates and the counts of tp, fp, tn, fn of the DBN for each screen respectively. The thresholds used for each screening were 0.04, 0.21 and 0.25 for screen 1,2 and 3 respectively.

in subsequent screenings as deceased/diagnosed with cancer individuals were removed from

the screening process of the trial.

	AUCs	AUCs C.I.	Interval
First Screening	0.789	0.774 - 0.804	0.0304
Second Screening	0.844	0.819 - 0.869	0.0496
Third Screening	0.884	0.863 - 0.906	0.0435

Table A.4: The reported AUCs of the ROC and the C.I. of the AUCs for each screening.

A.6.2   The *Forward-Arrow* DBN with a NoisyMax gate

The *Forward-Arrow* DBN with a NoisyMax gate						
	Screen 1		Screen 2		Screen 3	
Rates	0.96 (tp)	0.04 (fn)	0.87 (tp)	0.13 (fn)	0.83 (tp)	0.17 (fn)
	0.65 (fp)	0.35 (tn)	0.17 (fp)	0.83 (tn)	0.10 (fp)	0.90 (tn)
Counts	53 (tp)	2 (fn)	27 (tp)	4 (fn)	35 (tp)	7 (fn)
	221 (fp)	121 (tn)	50 (fp)	244 (tn)	24 (fp)	227 (tn)

Table A.5: The tp, fp, tn, fn rates and the counts of tp, fp, tn, fn of the DBN for each screen respectively. The thresholds used for each screening were 0.04, 0.21 and 0.25 for screen 1,2 and 3 respectively.

	AUCs	AUCs C.I.	Interval
First Screening	0.778	0.757 - 0.800	0.043
Second Screening	0.857	0.834 - 0.880	0.046
Third Screening	0.887	0.869 - 0.905	0.035

Table A.6: The reported AUCs of the ROC and the C.I. of the AUCs for each screening.

A.6.3 *Reversed-Arrow* DBN

Reversed-Arrow DBN						
	Screen 1		**Screen 2**		**Screen 3**	
Rates	0.93 (tp)	0.07 (fn)	0.87 (tp)	0.13 (fn)	0.83 (tp)	0.17 (fn)
	0.39 (fp)	0.61 (tn)	0.17 (fp)	0.83 (tn)	0.10 (fp)	0.90 (tn)
Counts	51 (tp)	4 (fn)	27 (tp)	4 (fn)	34 (tp)	7 (fn)
	134 (fp)	208 (tn)	50 (fp)	244 (tn)	24 (fp)	227 (tn)

Table A.7: The tp, fp, tn, fn rates and the counts of tp, fp, tn, fn of the DBN for each screen respectively. The thresholds used for each screening were 0.04, 0.21 and 0.25 for screen 1,2 and 3 respectively.

	AUCs	AUCs C.I.	Interval
First Screening	0.798	0.776 - 0.821	0.045
Second Screening	0.858	0.832 - 0.884	0.052
Third Screening	0.887	0.866 - 0.907	0.041

Table A.8: The reported AUCs of the ROC and the C.I. of the AUCs for each screening.

A.6.4 Learned DBN with compositional variables (structure learning)

Learned DBN with compositional variables					
	Screen 1		Screen 2		Screen 3
Rates	0.93 (tp)	0.07 (fn)	0.87 (tp)	0.13 (fn)	0.81 (tp) 0.19 (fn)
	0.36 (fp)	0.64 (tn)	0.18 (fp)	0.82 (tn)	0.10 (fp) 0.90 (tn)
Counts	51 (tp)	4 (fn)	27 (tp)	4 (fn)	34 (tp) 8 (fn)
	122 (fp)	220 (tn)	53 (fp)	241 (tn)	26 (fp) 225 (tn)

Table A.9: The tp, fp, tn, fn rates and the counts of tp, fp, tn, fn of the DBN for each screen respectively. The thresholds used for each screening were 0.04, 0.21 and 0.25 for screen 1,2 and 3 respectively. Bottom:

	AUCs	AUCs C.I.	Interval
First Screening	0.790	0.769 - 0.810	0.040
Second Screening	0.862	0.839 - 0.886	0.047
Third Screening	0.877	0.858 - 0.896	0.038

Table A.10: The reported AUCs of the ROC and the C.I. of the AUCs for each screening.

A.6.5 Learned DBN without compositional variables

Learned DBN without compositional variables						
	Screen 1		Screen 2		Screen 3	
Rates	0.95 (tp)	0.05 (fn)	0.81 (tp)	0.19 (fn)	0.83 (tp)	0.17 (fn)
	0.42 (fp)	0.58 (tn)	0.17 (fp)	0.83 (tn)	0.11 (fp)	0.89 (tn)
Counts	52 (tp)	3 (fn)	26 (tp)	6 (fn)	34 (tp)	7 (fn)
	145 (fp)	198 (tn)	51 (fp)	244 (tn)	28 (fp)	222 (tn)

Table A.11: The tp, fp, tn, fn rates and the counts of tp, fp, tn, fn of the DBN for each screen respectively. The thresholds used for each screening were 0.04, 0.21 and 0.25 for screen 1,2 and 3 respectively.

	AUCs	AUCs C.I.	Interval
First Screening	0.751	0.654 - 0.849	0.195
Second Screening	0.853	0.832 - 0.875	0.043
Third Screening	0.878	0.859 - 0.897	0.038

Table A.12: The reported AUCs of the ROC and the C.I. of the AUCs for each screening.

A.6.6 Naïve Bayes (NB)

Naïve Bayes						
	Screen 1		Screen 2		Screen 3	
Rates	0.927 (tp)	0.073 (fn)	0.871 (tp)	0.129 (fn)	0.833 (tp)	0.167 (fn)
	0.392 (fp)	0.608 (tn)	0.170 (fp)	0.830 (tn)	0.096 (fp)	0.904 (tn)
Counts	51 (tp)	4 (fn)	27 (tp)	4 (fn)	35 (tp)	7 (fn)
	134 (fp)	208 (tn)	50 (fp)	244 (tn)	24 (fp)	227 (tn)

Table A.13: Top: The tp, fp, tn, fn rates and the counts of tp, fp, tn, fn of the DBN for each screen respectively. The thresholds used for each screening were 0.04, 0.21 and 0.25 for screen 1,2 and 3 respectively. Bottom:

	AUCs	AUCs C.I.	Interval
First Screening	0.799	0.777 - 0.821	0.044
Second Screening	0.865	0.844 - 0.885	0.041
Third Screening	0.886	0.866 - 0.907	0.041

Table A.14: The reported AUCs of the ROC and the C.I. of the AUCs for each screening.

A.7 The Probability Distributions over each screen of confirmed cancer and Non-cancer cases

A.7.1 The *Forward-Arrow* DBN without a NoisyMax gate
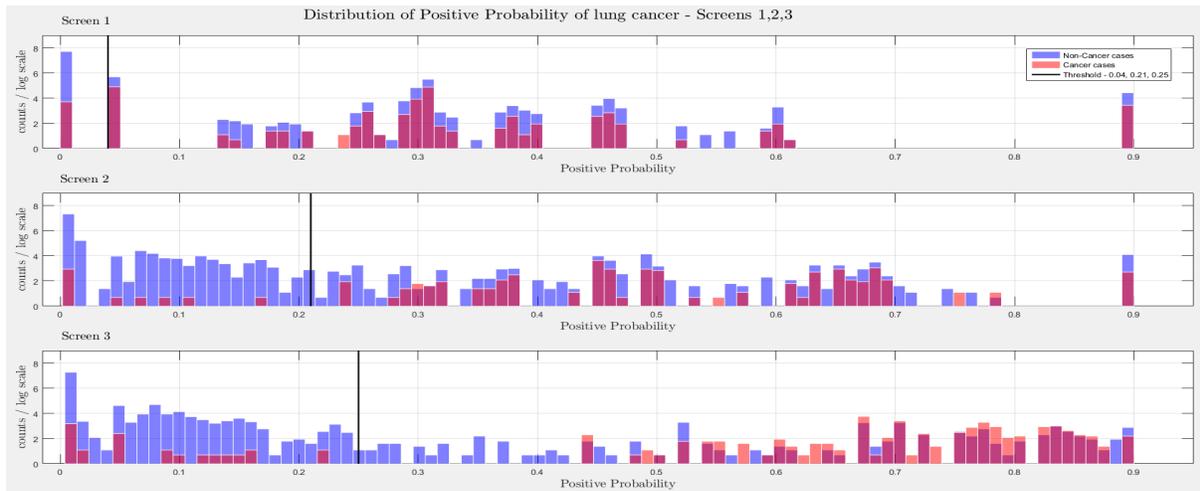
Figure A.3: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.7.2 The *Forward-Arrow* DBN with a NoisyMax gate
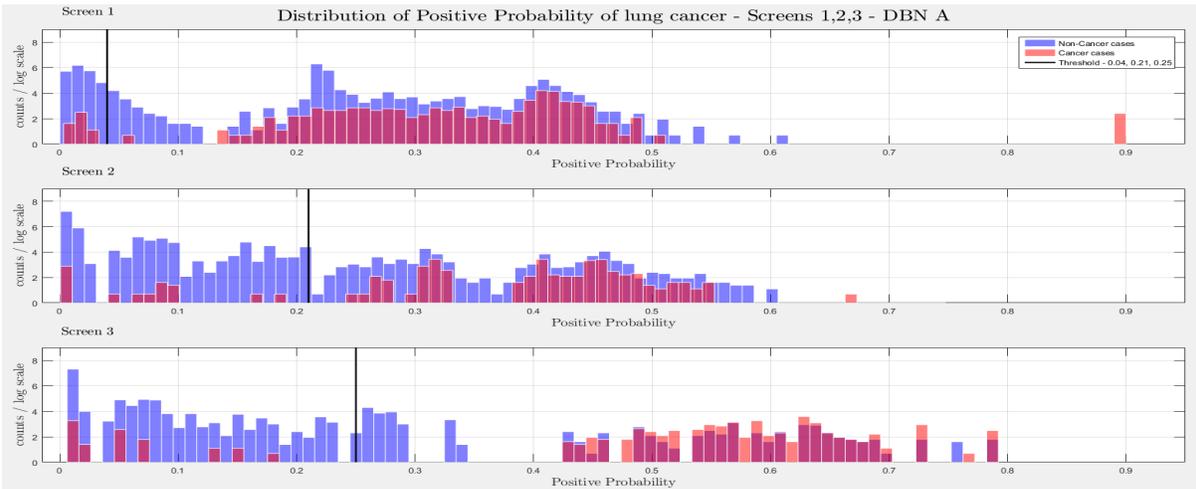
Figure A.4: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.7.3 *Reversed-Arrow* DBN
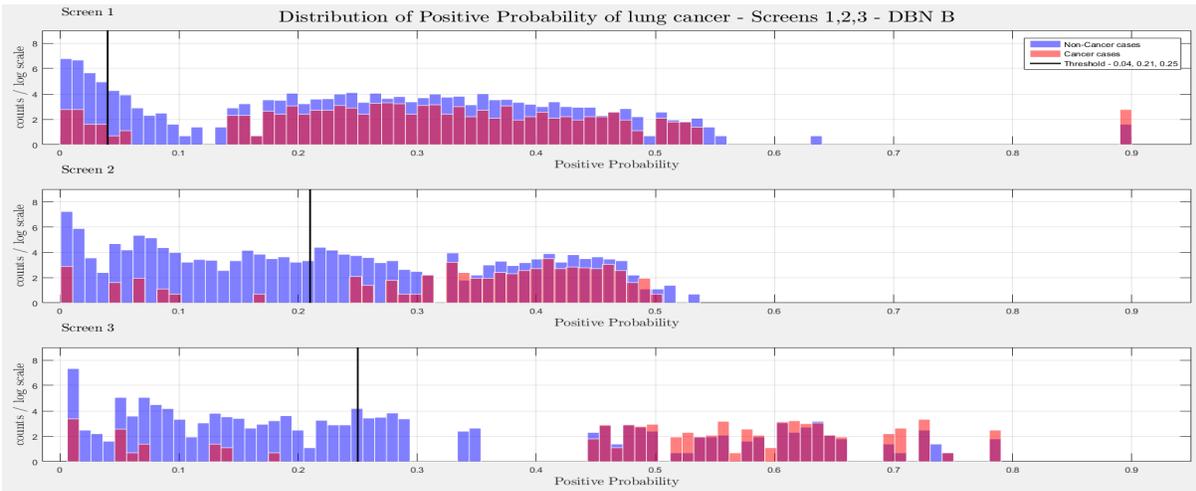
Figure A.5: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.7.4 Learned DBN with compositional variables
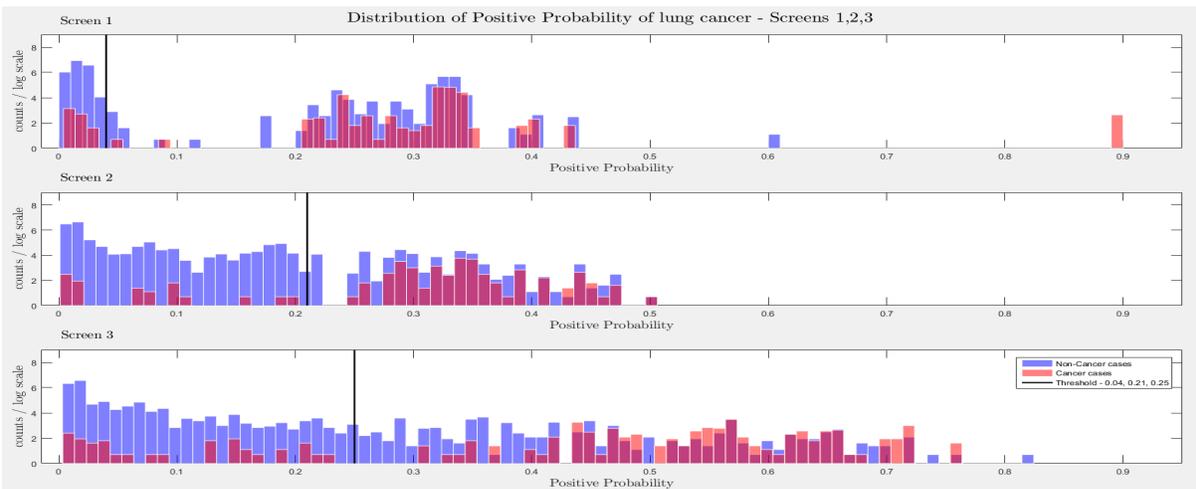
Figure A.6: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.7.5 Learned DBN without compositional variables
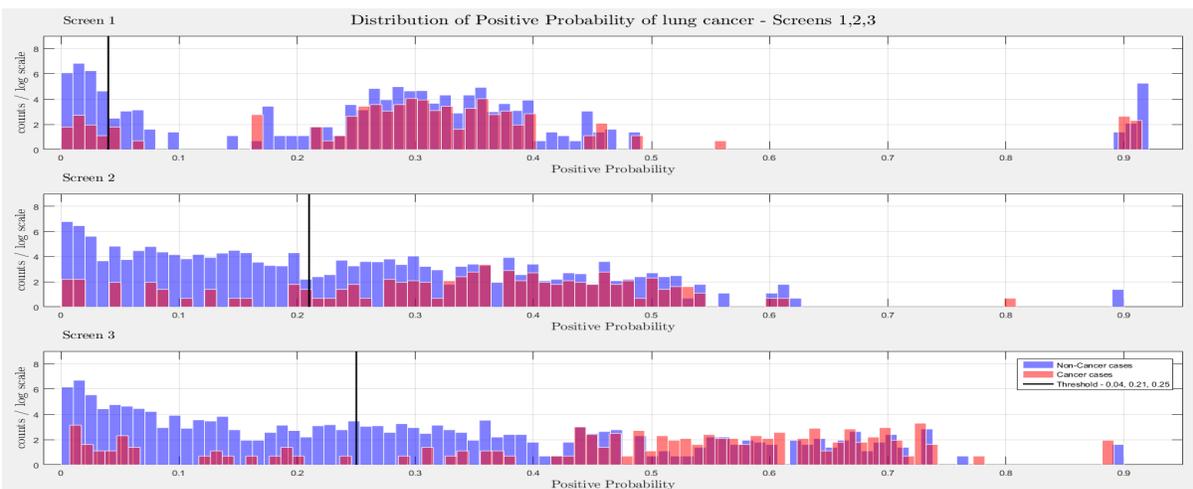
Figure A.7: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.7.6 10-fold cross validation of the *Forward-Arrow* DBN with a NoisyMax gate
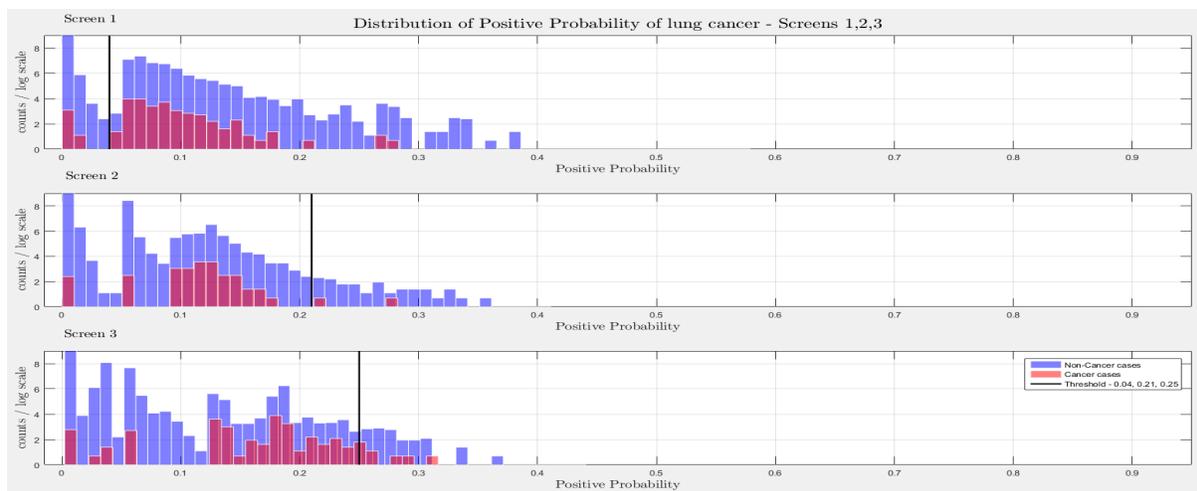
Figure A.8: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.7.7 Naïve Bayes
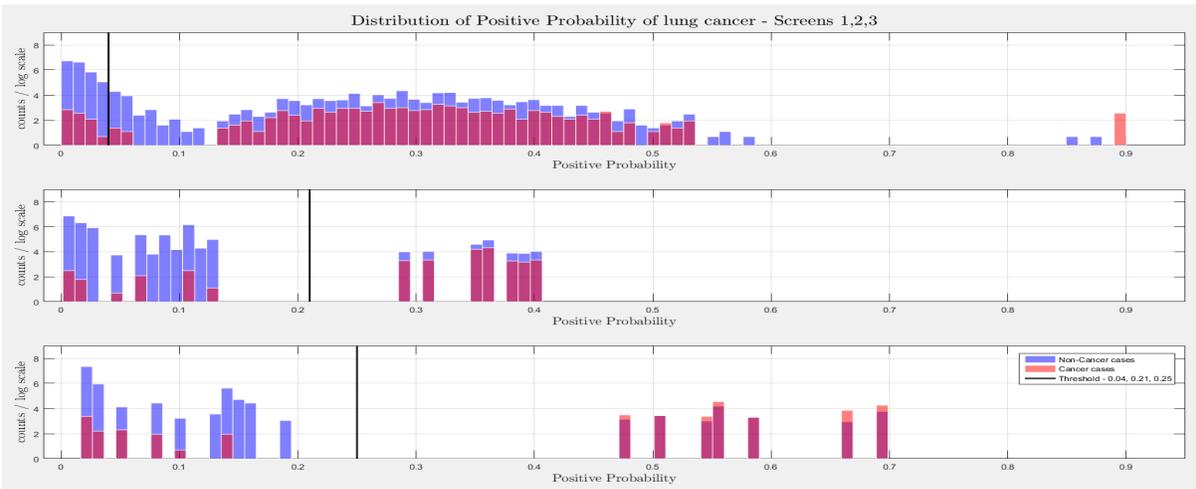
Figure A.9: The combined probability distributions for a positive biopsy for all cases across the 10 random test sets, for each screen. Red indicates all confirmed cancer cases in the trial, irrespective of time. Blue indicates the confirmed non-cancer cases. The 3 subplots depict the probability of a positive biopsy in each of the three screening points of the trial. With successive screenings we can see that the probability of a positive biopsy for non-cancer (blue) and cancer (red) cases tends to move towards the left and right side of each subplot, respectively. The solid black lines represent the thresholds chosen to discriminate cancer cases from non-cancer cases in the DBN predictions.

A.8 F-Scrore curves

A.8.1 The *Forward-Arrow* DBN without a NoisyMax gate
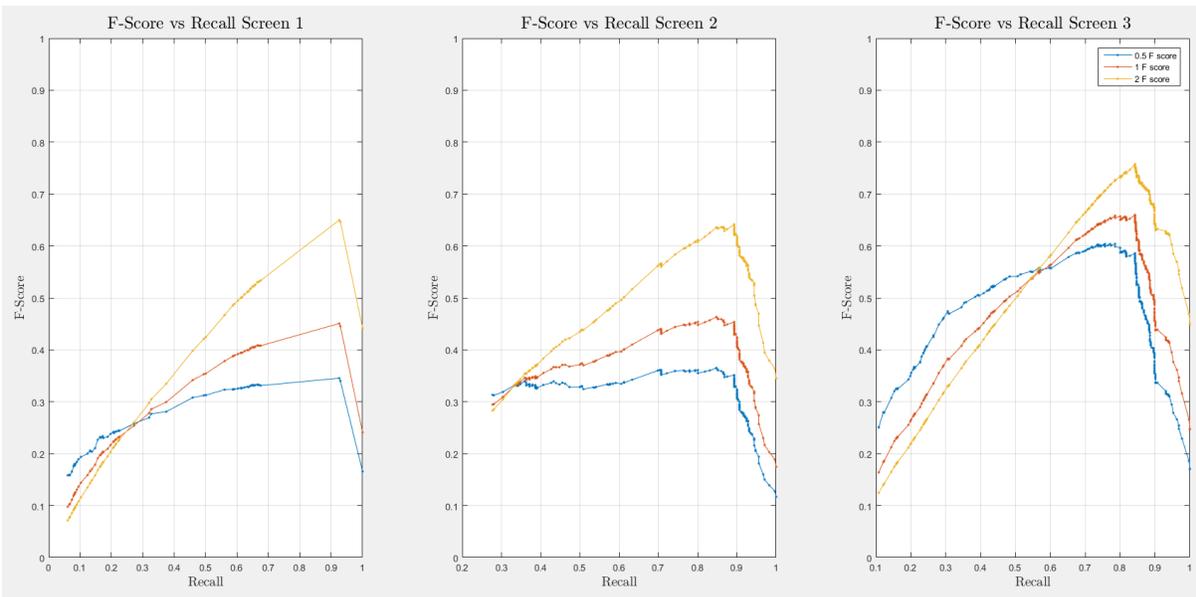
Figure A.10: F-score over recall curve.

A.8.2 The *Forward-Arrow* DBN with a NoisyMax gate
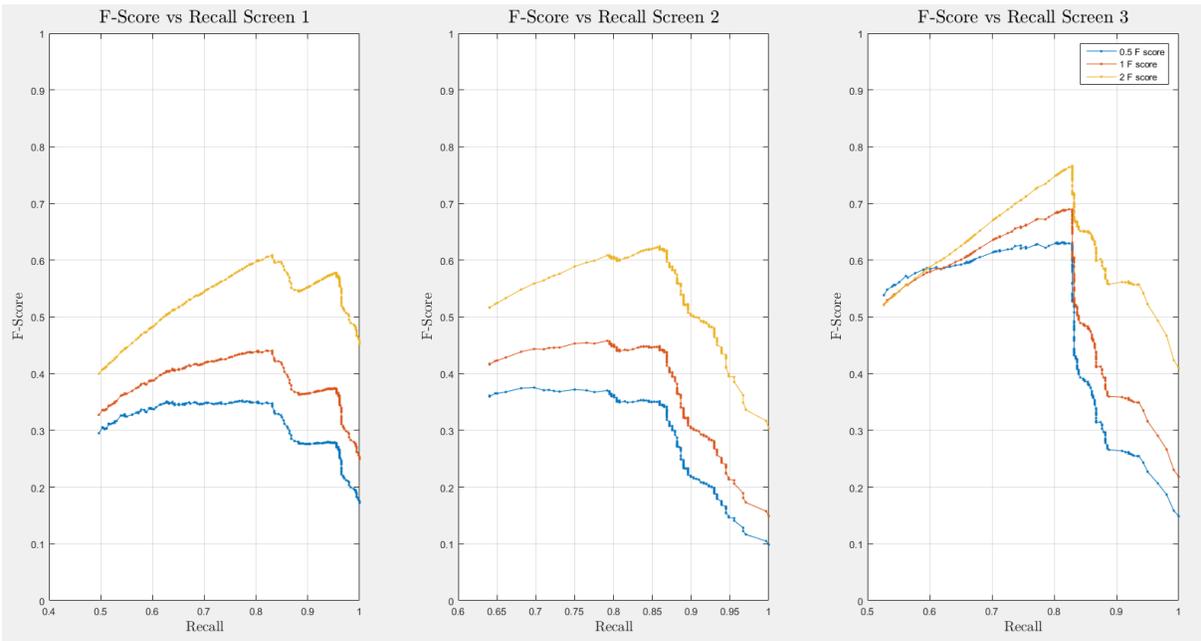
Figure A.11: F-score over recall curve.

A.8.3 *Reversed-Arrow* DBN
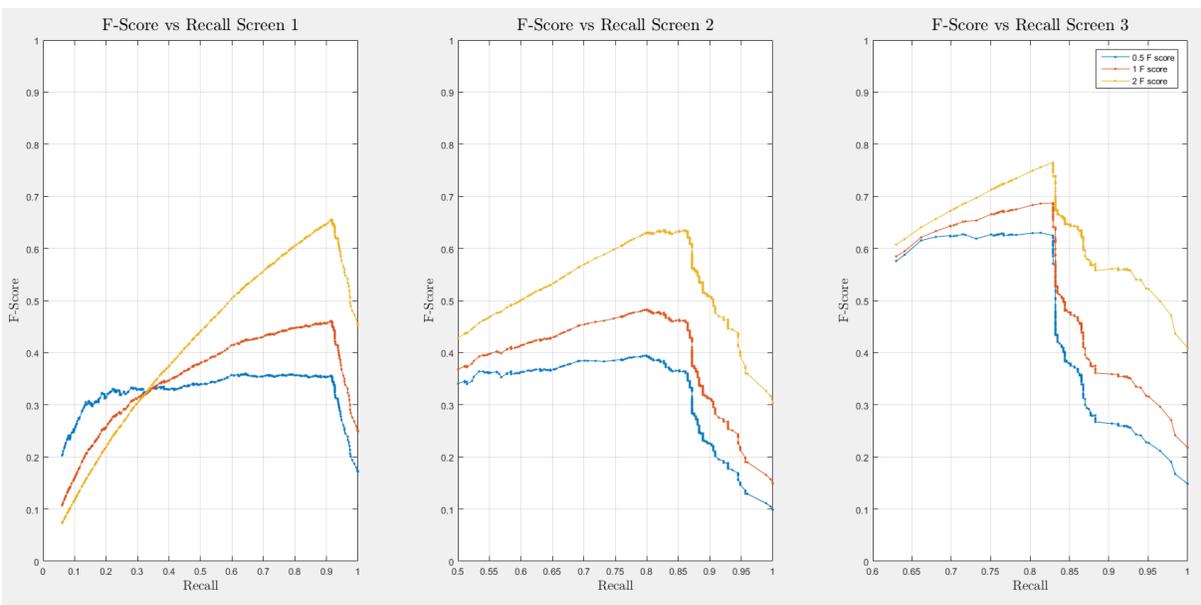
Figure A.12: F-score over recall curve.

A.8.4   Learned DBN with compositional variables
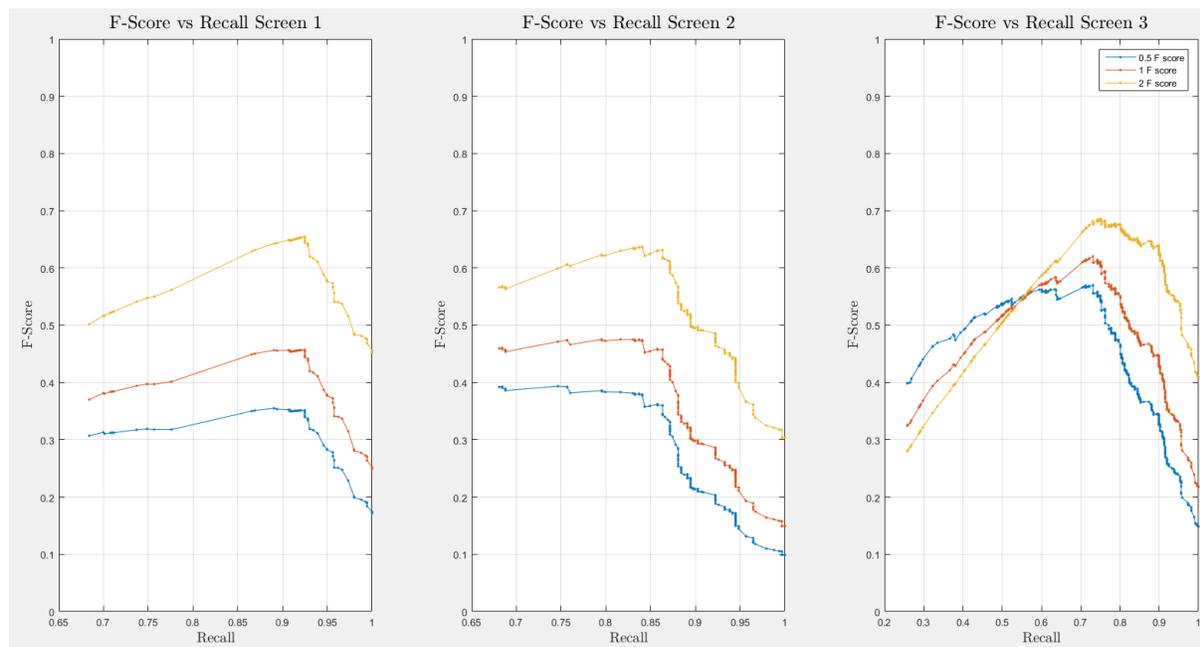
Figure A.13:   F-score over recall curve.

A.8.5   Learned DBN without compositional variables
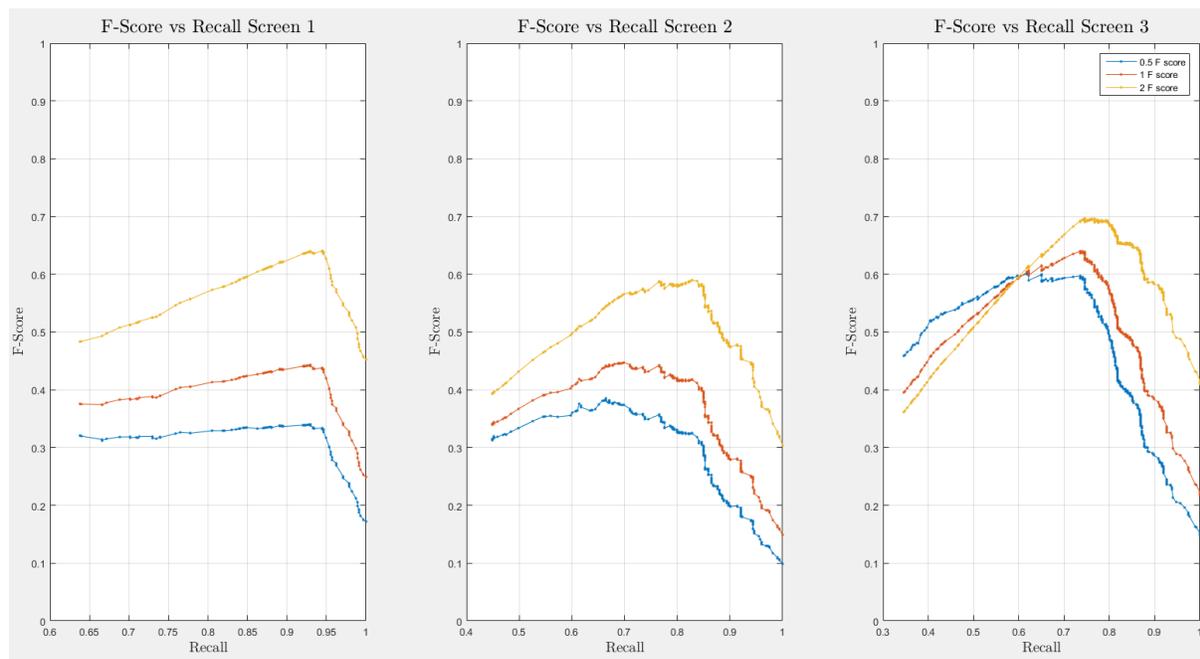
Figure A.14:   F-score over recall curve.

A.8.6 10-fold cross validation of the *Forward-Arrow* DBN with a NoisyMax gate
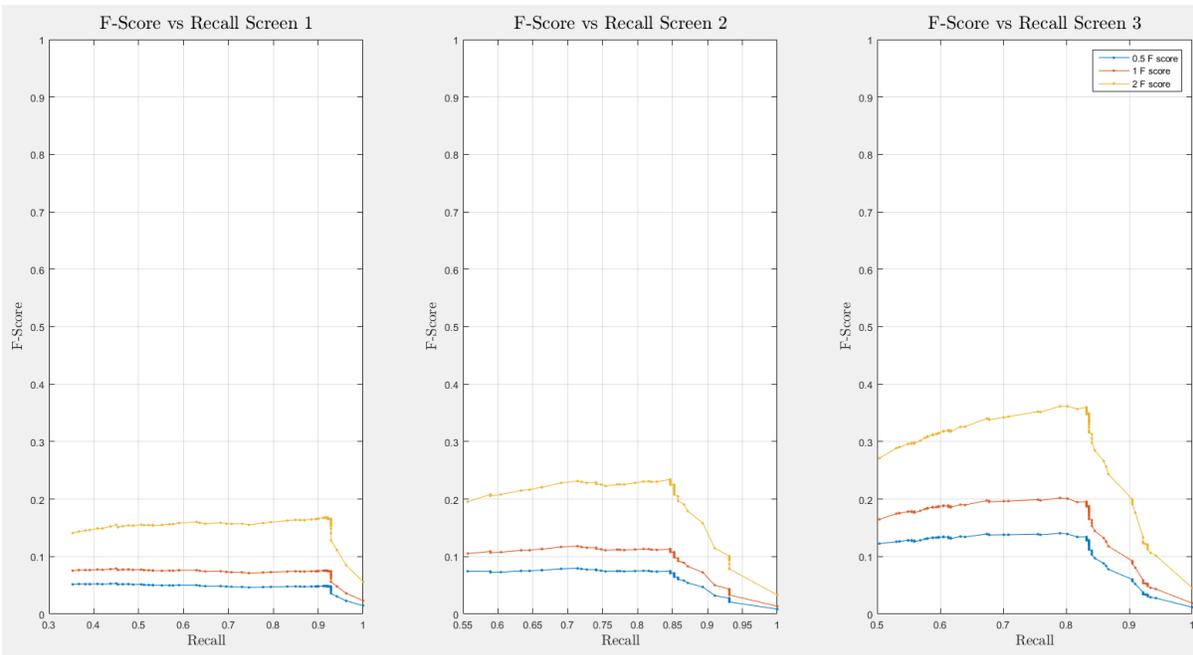
Figure A.15: F-score over recall curve.

A.8.7 Naïve Bayes (NB)
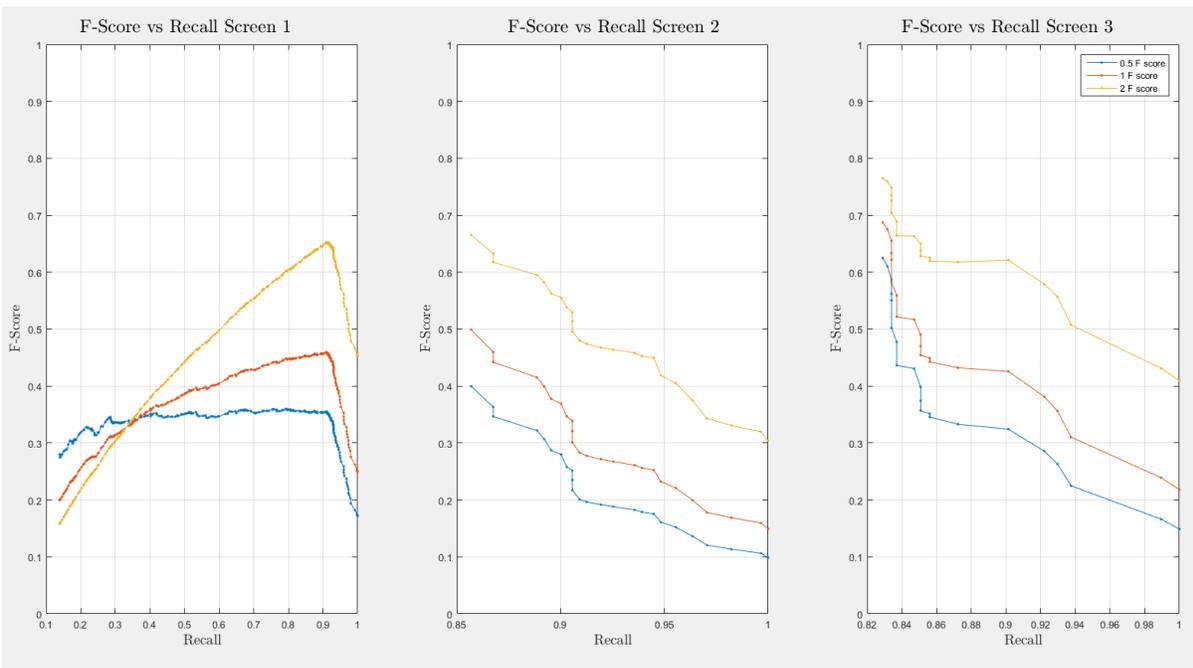
Figure A.16: F-score over recall curve.

A.9 PR Curves of the original model

A.9.1 The *Forward-Arrow* DBN without a NoisyMax gate
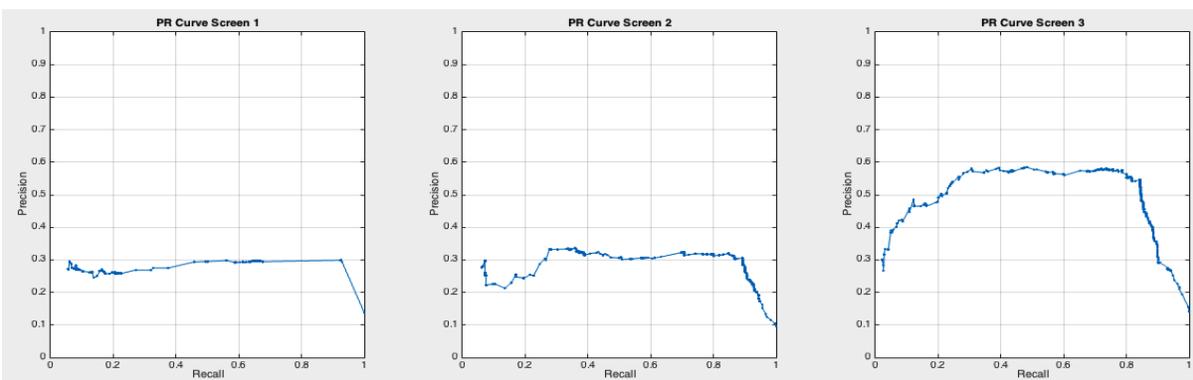
Figure A.17: The precision and recall curve.

A.9.2   The *Forward-Arrow* DBN with a NoisyMax gate
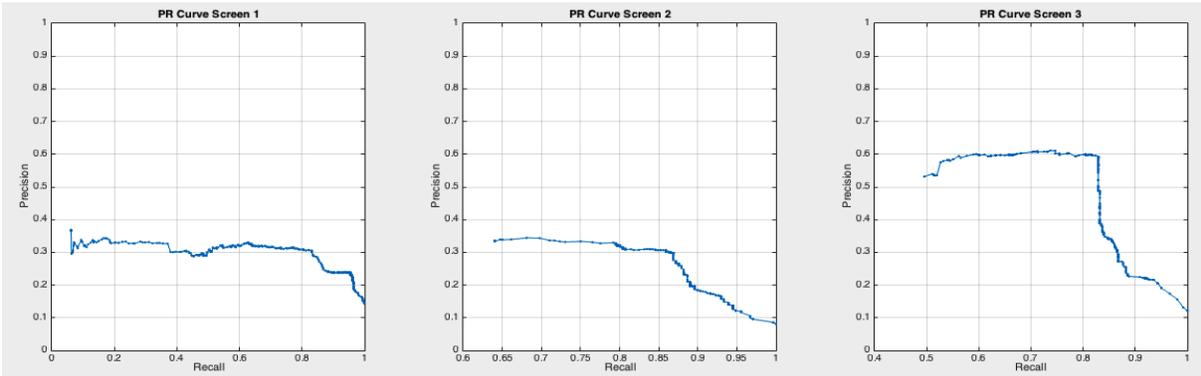
Figure A.18:   The precision and recall curve.

A.9.3   *Reversed-Arrow* DBN
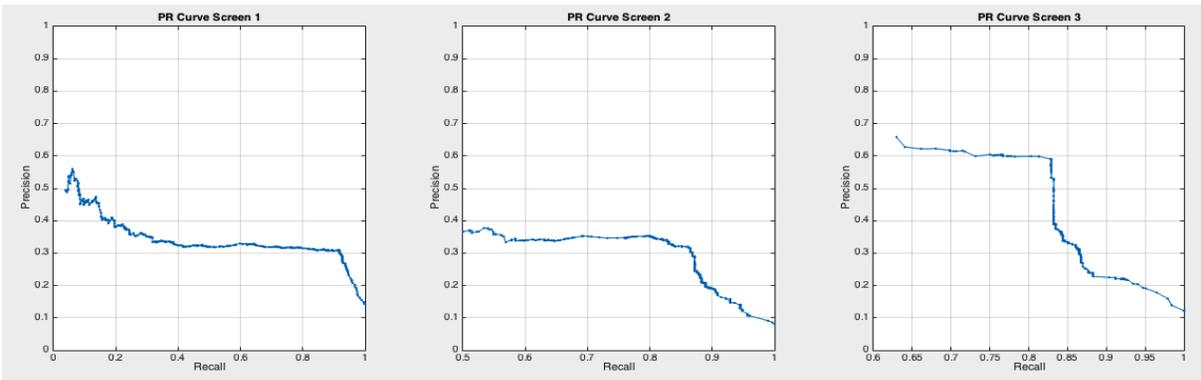
Figure A.19:   The precision and recall curve.

A.9.4 Learned DBN with compositional variables
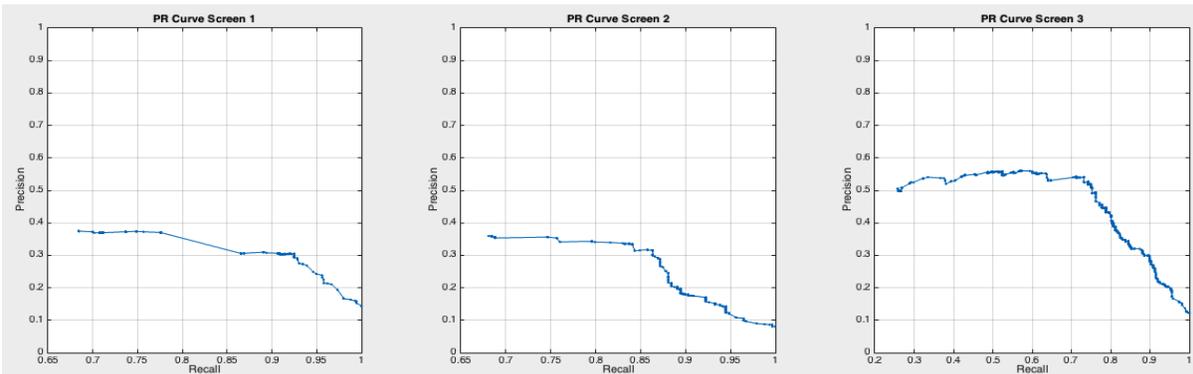
Figure A.20: The precision and recall curve.

A.9.5 Learned DBN without compositional variables
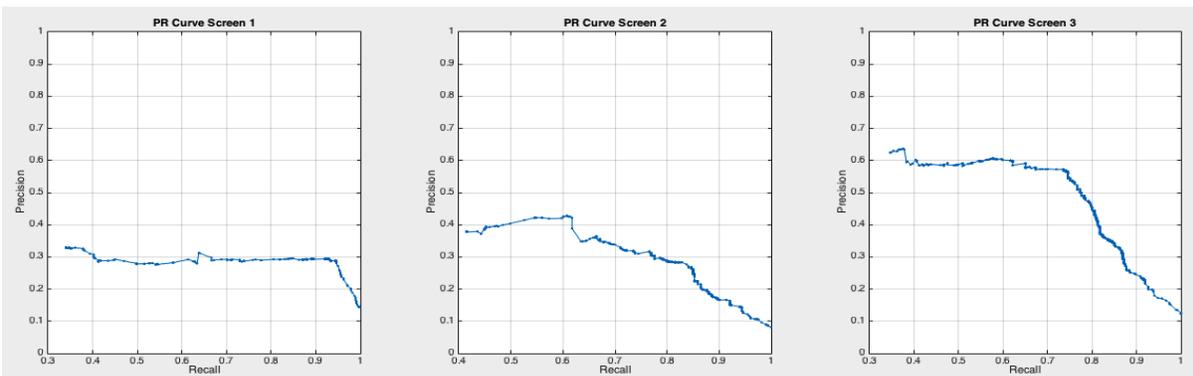
Figure A.21: The precision and recall curve.

A.9.6 10-fold cross validation of the *Forward-Arrow* DBN with a NoisyMax gate
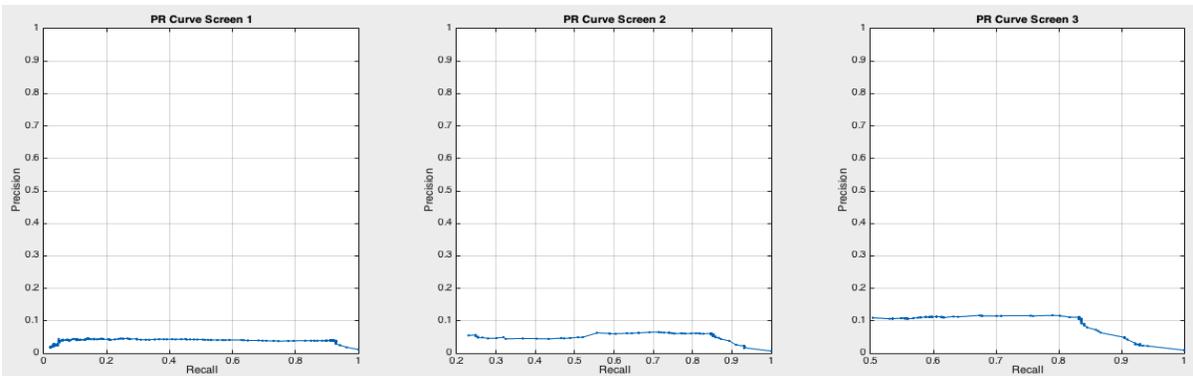
Figure A.22: The precision and recall curve.

A.9.7 Naïve Bayes (NB)
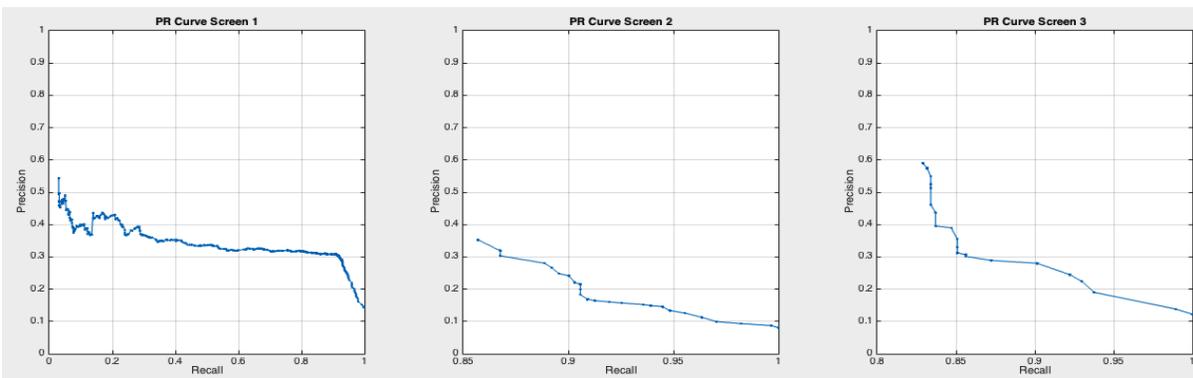
Figure A.23: The precision and recall curve.

A.10 Missing values statistics

		Age	BMI	Family History	Disease History	Cancer History	Smoking Status	Work Exposure	Gender
Count	**Present values**	25846	25573	25846	25846	25846	25846	25846	25846
	Missing values	0	93	0	0	0	0	0	0
Fraction	**Present values**	1	0.9964	1	1	1	1	1	1
	Missing values	0	0.0036	0	0	0	0	0	

Table A.15: Parent Nodes Missing value counts.

		LDCT Screen 1 Outcome	LDCT Screen 2 Outcome	LDCT Screen 3 Outcome
Count	**Present values**	25827	24335	23696
	Missing values	19	1511	2150
Fraction	**Present values**	0.9993	0.942	0.917
	Missing values	0.0007	0.058	0.083

Table A.16: LDCT nodes outcomes missing values. The missing values of these nodes consist of individuals that died, were diagnosed with cancer and are administered treatment and individuals that missed a screening exam.

A.11   Effect size range

Table A.17: Magnitude of effect size, Cohen et al [115].

Magnitude of effect size	Cramer's V or ϕ	Cohen's d	r^2 or η^2
Small	0.1	0.2	0.01
Medium	0.3	0.5	0.059
Large	0.5	0.8	0.14

Cramer's V was used for χ^2 and Fisher tests.

r^2 or η^2 was used for Student test.

Cohen's r^2 was used for Wilcoxon-Mann-Whitney test.

A.12   False Positives Analysis

Testing data.

	Physician	POMDP	p	effect-size
cigsmok			< **0.001**	0.432
0	527/735 (71.7 %)	237/839 (28.25 %)		
1	208/735 (28.3 %)	602/839 (71.75 %)		
diagcopd			< **0.001**	0.187
0	728/733 (99.32 %)	758/834 (90.89 %)		
1	5/733 (0.68 %)	76/834 (9.11 %)		
famHist			< **0.001**	0.327
0	670/735 (91.16 %)	530/839 (63.17 %)		
1	65/735 (8.84 %)	309/839 (36.83 %)		
pCancHist			< **0.001**	0.152
0	728/735 (99.05 %)	778/839 (92.73 %)		

1	7/735 (0.95 %)	61/839 (7.27 %)		
gender			0.185	0.033
1	435/735 (59.18 %)	525/839 (62.57 %)		
2	300/735 (40.82 %)	314/839 (37.43 %)		
race			< **0.001**	0.09
1	701/735 (95.37 %)	746/839 (88.92 %)		
2	11/735 (1.5 %)	75/839 (8.94 %)		
4	20/735 (2.72 %)	10/839 (1.19 %)		
5	0/735 (0 %)	5/839 (0.6 %)		
6	3/735 (0.41 %)	3/839 (0.36 %)		
educat			< **0.001**	0.129
1	1/577 (0.17 %)	19/617 (3.08 %)		
2	9/577 (1.56 %)	79/617 (12.8 %)		
4	84/577 (14.56 %)	120/617 (19.45 %)		
5	180/577 (31.2 %)	211/617 (34.2 %)		
6	140/577 (24.26 %)	104/617 (16.86 %)		
7	152/577 (26.34 %)	67/617 (10.86 %)		
8	11/577 (1.91 %)	17/617 (2.76 %)		
sctpreatt0			< **0.001**	0.078
1	566/664 (85.24 %)	213/280 (76.07 %)		
2	57/664 (8.58 %)	50/280 (17.86 %)		
3	35/664 (5.27 %)	15/280 (5.36 %)		
4	6/664 (0.9 %)	2/280 (0.71 %)		
sctmargins0			< **0.001**	0.239
1	1/684 (0.15 %)	38/276 (13.77 %)		
2	603/684 (88.16 %)	184/276 (66.67 %)		
3	80/684 (11.7 %)	54/276 (19.57 %)		

BMI	$29.06(4.94), 2$	$26.4(4.26), 3$	$< \mathbf{0.001}$	0.072
smokeIntensity	$29.16(11.57), 0$	$28.58(11.35), 0$	0.24	< 0.001
smokeyr	$34.65(5.07), 14$	$45.77(6.38), 1$	$< \mathbf{0.001}$	**0.514**
smokeQuitTime	$5.72(5.36), 27$	$1.76(3.79), 9$	$< \mathbf{0.001}$	**0.186**
age	$58.39(3.27), 57$	$64.57(5.29), 49$	$< \mathbf{0.001}$	**0.321**
LargestDiam0	$5.39(1.88), 0$	$2.71(4.85), 0$	$< \mathbf{0.001}$	**0.215**

Table A.18: Comparison between physicians and POMDP (baseline screen). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage)". Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			$< \mathbf{0.001}$	0.343
0	605/875 (69.14 %)	279/802 (34.79 %)		
1	270/875 (30.86 %)	523/802 (65.21 %)		
diagcopd			$< \mathbf{0.001}$	0.137
0	862/872 (98.85 %)	745/796 (93.59 %)		
1	10/872 (1.15 %)	51/796 (6.41 %)		
famHist			$< \mathbf{0.001}$	0.259
0	774/875 (88.46 %)	537/802 (66.96 %)		
1	101/875 (11.54 %)	265/802 (33.04 %)		
pCancHist			$< \mathbf{0.001}$	0.107
0	862/875 (98.51 %)	758/802 (94.51 %)		
1	13/875 (1.49 %)	44/802 (5.49 %)		

gender			**0.037**	0.051
1	504/875 (57.6 %)	503/802 (62.72 %)		
2	371/875 (42.4 %)	299/802 (37.28 %)		
race			**< 0.001**	0.074
1	826/875 (94.4 %)	719/802 (89.65 %)		
2	21/875 (2.4 %)	68/802 (8.48 %)		
4	25/875 (2.86 %)	10/802 (1.25 %)		
5	0/875 (0 %)	1/802 (0.12 %)		
6	3/875 (0.34 %)	4/802 (0.5 %)		
educat			**< 0.001**	0.099
1	2/685 (0.29 %)	14/599 (2.34 %)		
2	11/685 (1.61 %)	61/599 (10.18 %)		
4	101/685 (14.74 %)	105/599 (17.53 %)		
5	217/685 (31.68 %)	199/599 (33.22 %)		
6	166/685 (24.23 %)	123/599 (20.53 %)		
7	173/685 (25.26 %)	85/599 (14.19 %)		
8	15/685 (2.19 %)	12/599 (2 %)		
sctpreatt0			**< 0.001**	0.087
1	557/651 (85.56 %)	181/244 (74.18 %)		
2	58/651 (8.91 %)	48/244 (19.67 %)		
3	32/651 (4.92 %)	13/244 (5.33 %)		
4	4/651 (0.61 %)	2/244 (0.82 %)		
sctpreatt1			**0.023**	0.056
1	528/595 (88.74 %)	339/406 (83.5 %)		
2	47/595 (7.9 %)	56/406 (13.79 %)		
3	16/595 (2.69 %)	9/406 (2.22 %)		
4	4/595 (0.67 %)	2/406 (0.49 %)		

			p	effect-size
sctmargins0			**< 0.001**	0.233
1	0/675 (0 %)	30/240 (12.5 %)		
2	595/675 (88.15 %)	164/240 (68.33 %)		
3	80/675 (11.85 %)	46/240 (19.17 %)		
sctmargins1			**< 0.001**	0.145
1	7/620 (1.13 %)	38/416 (9.13 %)		
2	525/620 (84.68 %)	303/416 (72.84 %)		
3	88/620 (14.19 %)	75/416 (18.03 %)		
BMI	28.9(5.06), 4	26.9(4.63), 3	**< 0.001**	0.041
smokeIntensity	28.77(11.26), 0	28.63(11.47), 0	0.545	< 0.001
smokeyr	35.24(5.29), 17	43.51(7.13), 2	**< 0.001**	**0.323**
smokeQuitTime	5.4(5.36), 32	2.29(4.28), 11	**< 0.001**	0.112
age	58.62(3.49), 69	63.05(5.53), 56	**< 0.001**	**0.166**
LargestDiam0	4.36(2.55), 0	2.45(4.7), 0	**< 0.001**	0.133
LargestDiam1	4(3.59), 31	4.1(5.56), 38	0.297	< 0.001

Table A.19: Comparison between physicians and POMDP (2nd screen). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage)". Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			**< 0.001**	0.278
0	643/948 (67.83 %)	306/766 (39.95 %)		
1	305/948 (32.17 %)	460/766 (60.05 %)		

diagcopd			< **0.001**	0.074
0	928/946 (98.1 %)	728/763 (95.41 %)		
1	18/946 (1.9 %)	35/763 (4.59 %)		
famHist			< **0.001**	0.191
0	830/948 (87.55 %)	554/766 (72.32 %)		
1	118/948 (12.45 %)	212/766 (27.68 %)		
pCancHist			< **0.001**	0.08
0	931/948 (98.21 %)	730/766 (95.3 %)		
1	17/948 (1.79 %)	36/766 (4.7 %)		
gender			0.079	0.042
1	547/948 (57.7 %)	475/766 (62.01 %)		
2	401/948 (42.3 %)	291/766 (37.99 %)		
race			< **0.001**	0.071
1	894/948 (94.3 %)	690/766 (90.08 %)		
2	26/948 (2.74 %)	61/766 (7.96 %)		
4	25/948 (2.64 %)	12/766 (1.57 %)		
6	3/948 (0.32 %)	3/766 (0.39 %)		
educat			< **0.001**	0.085
1	1/746 (0.13 %)	13/567 (2.29 %)		
2	11/746 (1.47 %)	43/567 (7.58 %)		
4	116/746 (15.55 %)	99/567 (17.46 %)		
5	236/746 (31.64 %)	184/567 (32.45 %)		
6	184/746 (24.66 %)	127/567 (22.4 %)		
7	182/746 (24.4 %)	89/567 (15.7 %)		
8	16/746 (2.14 %)	12/567 (2.12 %)		
sctpreatt0			< **0.001**	0.084
1	543/627 (86.6 %)	177/235 (75.32 %)		

2	54/627 (8.61 %)	44/235 (18.72 %)		
3	26/627 (4.15 %)	12/235 (5.11 %)		
4	4/627 (0.64 %)	2/235 (0.85 %)		
sctpreatt1			< **0.001**	0.066
1	528/588 (89.8 %)	301/361 (83.38 %)		
2	43/588 (7.31 %)	51/361 (14.13 %)		
3	13/588 (2.21 %)	8/361 (2.22 %)		
4	4/588 (0.68 %)	1/361 (0.28 %)		
sctpreatt2			**0.02**	0.055
1	503/572 (87.94 %)	403/490 (82.24 %)		
2	52/572 (9.09 %)	74/490 (15.1 %)		
3	11/572 (1.92 %)	10/490 (2.04 %)		
4	6/572 (1.05 %)	3/490 (0.61 %)		
sctmargins0			< **0.001**	0.234
1	0/650 (0 %)	29/232 (12.5 %)		
2	578/650 (88.92 %)	160/232 (68.97 %)		
3	72/650 (11.08 %)	43/232 (18.53 %)		
sctmargins1			< **0.001**	0.176
1	0/611 (0 %)	32/370 (8.65 %)		
2	529/611 (86.58 %)	272/370 (73.51 %)		
3	82/611 (13.42 %)	66/370 (17.84 %)		
sctmargins2			< **0.001**	0.099
1	8/582 (1.37 %)	30/500 (6 %)		
2	501/582 (86.08 %)	389/500 (77.8 %)		
3	73/582 (12.54 %)	81/500 (16.2 %)		
BMI	28.83(5.05), 6	27.3(4.82), 2	< **0.001**	0.025
smokeIntensity	28.81(11.19), 0	29.15(11.94), 0	0.889	< 0.001

smokeyr	$35.51(5.46), 18$	$41.62(7.25), 4$	**< 0.001**	**0.199**
smokeQuitTime	$5.24(5.39), 32$	$2.82(4.74), 16$	**< 0.001**	0.068
age	$58.69(3.58), 77$	$61.81(5.33), 57$	**< 0.001**	0.088
LargestDiam0	$3.85(2.75), 0$	$2.46(4.72), 0$	**< 0.001**	0.083
LargestDiam1	$3.61(3.61), 30$	$3.79(5.43), 38$	0.442	< 0.001
LargestDiam2	$3.34(2.58), 54$	$4.74(4.17), 59$	**< 0.001**	0.044

Table A.20: Comparison between physicians and POMDP (3rd screen). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage)". Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			**< 0.001**	0.276
0	639/947 (67.48 %)	299/753 (39.71 %)		
1	308/947 (32.52 %)	454/753 (60.29 %)		
diagcopd			**< 0.001**	0.072
0	927/945 (98.1 %)	716/750 (95.47 %)		
1	18/945 (1.9 %)	34/750 (4.53 %)		
famHist			**< 0.001**	0.193
0	830/947 (87.65 %)	544/753 (72.24 %)		
1	117/947 (12.35 %)	209/753 (27.76 %)		
pCancHist			**< 0.001**	0.079
0	930/947 (98.2 %)	718/753 (95.35 %)		
1	17/947 (1.8 %)	35/753 (4.65 %)		

gender			0.079	0.043
1	545/947 (57.55 %)	466/753 (61.89 %)		
2	402/947 (42.45 %)	287/753 (38.11 %)		
race			**< 0.001**	0.069
1	892/947 (94.19 %)	678/753 (90.04 %)		
2	27/947 (2.85 %)	60/753 (7.97 %)		
4	25/947 (2.64 %)	12/753 (1.59 %)		
6	3/947 (0.32 %)	3/753 (0.4 %)		
educat			**< 0.001**	0.082
1	2/745 (0.27 %)	12/556 (2.16 %)		
2	11/745 (1.48 %)	42/556 (7.55 %)		
4	116/745 (15.57 %)	96/556 (17.27 %)		
5	233/745 (31.28 %)	183/556 (32.91 %)		
6	184/745 (24.7 %)	125/556 (22.48 %)		
7	183/745 (24.56 %)	88/556 (15.83 %)		
8	16/745 (2.15 %)	10/556 (1.8 %)		
sctpreatt0			**< 0.001**	0.086
1	544/628 (86.62 %)	174/232 (75 %)		
2	54/628 (8.6 %)	44/232 (18.97 %)		
3	26/628 (4.14 %)	12/232 (5.17 %)		
4	4/628 (0.64 %)	2/232 (0.86 %)		
sctpreatt1			**< 0.001**	0.065
1	528/588 (89.8 %)	298/356 (83.71 %)		
2	43/588 (7.31 %)	50/356 (14.04 %)		
3	13/588 (2.21 %)	7/356 (1.97 %)		
4	4/588 (0.68 %)	1/356 (0.28 %)		
sctpreatt2			**0.019**	0.056

1	503/571 (88.09 %)	396/481 (82.33 %)		
2	51/571 (8.93 %)	72/481 (14.97 %)		
3	11/571 (1.93 %)	10/481 (2.08 %)		
4	6/571 (1.05 %)	3/481 (0.62 %)		
sctmargins0			< **0.001**	0.229
1	0/651 (0 %)	27/229 (11.79 %)		
2	579/651 (88.94 %)	159/229 (69.43 %)		
3	72/651 (11.06 %)	43/229 (18.78 %)		
sctmargins1			< **0.001**	0.174
1	0/611 (0 %)	31/365 (8.49 %)		
2	529/611 (86.58 %)	270/365 (73.97 %)		
3	82/611 (13.42 %)	64/365 (17.53 %)		
sctmargins2			< **0.001**	0.1
1	8/581 (1.38 %)	30/491 (6.11 %)		
2	501/581 (86.23 %)	383/491 (78 %)		
3	72/581 (12.39 %)	78/491 (15.89 %)		
BMI	28.82(5.06), 6	27.32(4.82), 2	< **0.001**	0.024
smokeIntensity	28.77(11.19), 0	29.02(11.83), 0	0.961	< 0.001
smokeyr	35.52(5.47), 18	41.64(7.25), 4	< **0.001**	**0.199**
smokeQuitTime	5.21(5.38), 32	2.79(4.72), 16	< **0.001**	0.068
age	58.68(3.56), 76	61.79(5.31), 56	< **0.001**	0.088
LargestDiam0	3.86(2.74), 0	2.47(4.74), 0	< **0.001**	0.083
LargestDiam1	3.61(3.61), 30	3.8(5.44), 37	0.442	< 0.001
LargestDiam2	3.35(2.58), 55	4.73(4.17), 58	< **0.001**	0.043

Table A.21: Comparison between physicians and POMDP (post screening). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage)". Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

A.13   Early TPs analysis

Testing data.

	Physician	POMDP	p	effect-size
cigsmok			0.084	0.229
0	13/30 (43.33 %)	5/27 (18.52 %)		
1	17/30 (56.67 %)	22/27 (81.48 %)		
diagcopd			0.238	0.127
0	28/30 (93.33 %)	22/27 (81.48 %)		
1	2/30 (6.67 %)	5/27 (18.52 %)		
famHist			0.855	0.024
0	16/30 (53.33 %)	16/27 (59.26 %)		
1	14/30 (46.67 %)	11/27 (40.74 %)		
pCancHist			0.66	0.016
0	28/30 (93.33 %)	24/27 (88.89 %)		
1	2/30 (6.67 %)	3/27 (11.11 %)		
gender			0.098	0.219
1	16/30 (53.33 %)	21/27 (77.78 %)		

2	14/30 (46.67 %)	6/27 (22.22 %)		
race			0.238	0.127
1	28/30 (93.33 %)	22/27 (81.48 %)		
2	2/30 (6.67 %)	5/27 (18.52 %)		
educat			0.838	0.106
1	1/21 (4.76 %)	1/19 (5.26 %)		
2	1/21 (4.76 %)	3/19 (15.79 %)		
4	4/21 (19.05 %)	4/19 (21.05 %)		
5	7/21 (33.33 %)	7/19 (36.84 %)		
6	4/21 (19.05 %)	2/19 (10.53 %)		
7	4/21 (19.05 %)	2/19 (10.53 %)		
sctpreatt0			0.324	0.218
1	18/29 (62.07 %)	6/6 (100 %)		
2	6/29 (20.69 %)	0/6 (0 %)		
3	5/29 (17.24 %)	0/6 (0 %)		
sctmargins0			0.87	0.101
1	11/27 (40.74 %)	4/7 (57.14 %)		
2	7/27 (25.93 %)	1/7 (14.29 %)		
3	9/27 (33.33 %)	2/7 (28.57 %)		
BMI	$27.15(7.99), 0$	$24.54(3.61), 0$	0.198	0.029
smokeIntensity	$30.17(11.56), 0$	$29.93(11.08), 0$	0.98	< 0.001
smokeyr	$43.83(5.79), 0$	$45.7(6.01), 0$	0.238	0.025
smokeQuitTime	$1.86(3.1), 1$	$0.73(2.6), 1$	0.059	0.065
age	$63.83(4.73), 6$	$63.92(5.36), 2$	0.952	< 0.001
LargestDiam0	$13.83(20.18), 0$	$3.52(6.67), 0$	$< \mathbf{0.001}$	**0.319**

Table A.22: Comparison between physicians and POMDP (early prediction of 2nd screen with a_0). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage). Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			**0.016**	0.31
0	14/26 (53.85 %)	7/34 (20.59 %)		
1	12/26 (46.15 %)	27/34 (79.41 %)		
diagcopd			1	< 0.001
0	23/26 (88.46 %)	30/34 (88.24 %)		
1	3/26 (11.54 %)	4/34 (11.76 %)		
famHist			0.101	0.212
0	20/26 (76.92 %)	18/34 (52.94 %)		
1	6/26 (23.08 %)	16/34 (47.06 %)		
pCancHist			1	< 0.001
0	25/26 (96.15 %)	32/34 (94.12 %)		
1	1/26 (3.85 %)	2/34 (5.88 %)		
gender			0.725	0.045
1	14/26 (53.85 %)	21/34 (61.76 %)		
2	12/26 (46.15 %)	13/34 (38.24 %)		
race			1	0.094
1	25/26 (96.15 %)	31/34 (91.18 %)		
2	1/26 (3.85 %)	1/34 (2.94 %)		

4	0/26 (0 %)	1/34 (2.94 %)		
6	0/26 (0 %)	1/34 (2.94 %)		
educat			0.053	0.208
2	2/21 (9.52 %)	4/27 (14.81 %)		
4	6/21 (28.57 %)	7/27 (25.93 %)		
5	4/21 (19.05 %)	13/27 (48.15 %)		
6	4/21 (19.05 %)	1/27 (3.7 %)		
7	5/21 (23.81 %)	1/27 (3.7 %)		
8	0/21 (0 %)	1/27 (3.7 %)		
sctpreatt0			1	0.065
1	16/26 (61.54 %)	2/3 (66.67 %)		
2	8/26 (30.77 %)	1/3 (33.33 %)		
4	2/26 (7.69 %)	0/3 (0 %)		
sctmargins0			1	0.11
1	5/22 (22.73 %)	0/2 (0 %)		
2	9/22 (40.91 %)	1/2 (50 %)		
3	8/22 (36.36 %)	1/2 (50 %)		
BMI	$25.84(3.83), 0$	$25.4(5.28), 0$	0.438	0.01
smokeIntensity	$33.65(16.34), 0$	$29.12(12.34), 0$	0.346	0.015
smokeyr	$40.96(6.86), 0$	$47.59(7.57), 0$	**< 0.001**	**0.171**
smokeQuitTime	$4.12(5.32), 1$	$1.45(3.96), 1$	**0.01**	0.116
age	$61.32(4.5), 1$	$65.91(5.34), 2$	**< 0.001**	**0.167**
LargestDiam0	$10.5(7.63), 0$	$0.85(3.23), 0$	**< 0.001**	**0.71**

Table A.23: Comparison between physicians and POMDP (early prediction of 3rd screen with a_0). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage). Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			0.051	0.226
0	20/43 (46.51 %)	7/32 (21.88 %)		
1	23/43 (53.49 %)	25/32 (78.12 %)		
diagcopd			0.451	0.048
0	40/43 (93.02 %)	28/32 (87.5 %)		
1	3/43 (6.98 %)	4/32 (12.5 %)		
famHist			0.258	0.131
0	32/43 (74.42 %)	19/32 (59.38 %)		
1	11/43 (25.58 %)	13/32 (40.62 %)		
pCancHist			0.572	0.03
0	42/43 (97.67 %)	30/32 (93.75 %)		
1	1/43 (2.33 %)	2/32 (6.25 %)		
gender			1	< 0.001
1	24/43 (55.81 %)	17/32 (53.12 %)		
2	19/43 (44.19 %)	15/32 (46.88 %)		
race			**0.038**	0.176
1	39/43 (90.7 %)	29/32 (90.62 %)		
2	4/43 (9.3 %)	0/32 (0 %)		

4	0/43 (0 %)	1/32 (3.12 %)		
6	0/43 (0 %)	2/32 (6.25 %)		
educat			0.195	0.156
2	3/33 (9.09 %)	4/25 (16 %)		
4	8/33 (24.24 %)	6/25 (24 %)		
5	9/33 (27.27 %)	12/25 (48 %)		
6	7/33 (21.21 %)	1/25 (4 %)		
7	5/33 (15.15 %)	1/25 (4 %)		
8	1/33 (3.03 %)	1/25 (4 %)		
sctpreatt0			1	0.065
1	16/26 (61.54 %)	2/3 (66.67 %)		
2	8/26 (30.77 %)	1/3 (33.33 %)		
4	2/26 (7.69 %)	0/3 (0 %)		
sctpreatt1			0.754	0.107
1	22/34 (64.71 %)	4/6 (66.67 %)		
2	7/34 (20.59 %)	1/6 (16.67 %)		
3	2/34 (5.88 %)	1/6 (16.67 %)		
4	3/34 (8.82 %)	0/6 (0 %)		
sctmargins0			1	0.11
1	5/22 (22.73 %)	0/2 (0 %)		
2	9/22 (40.91 %)	1/2 (50 %)		
3	8/22 (36.36 %)	1/2 (50 %)		
sctmargins1			0.645	0.126
1	9/31 (29.03 %)	2/6 (33.33 %)		
2	12/31 (38.71 %)	1/6 (16.67 %)		
3	10/31 (32.26 %)	3/6 (50 %)		
BMI	26.34(3.32), 0	25.29(5.44), 0	0.097	0.037

			p	effect-size
smokeIntensity	32.44(15.13), 0	26.5(10), 0	0.079	0.041
smokeyr	42.4(6.47), 0	46.25(7.99), 0	**0.029**	0.064
smokeQuitTime	3.26(4.79), 1	1.55(4.07), 1	**0.032**	0.063
age	61.69(4.79), 1	64.83(5.86), 2	**0.019**	0.076
LargestDiam0	6.35(7.85), 0	0.91(3.32), 0	**< 0.001**	**0.25**
LargestDiam1	7.86(8.04), 1	2.98(7.15), 0	**< 0.001**	**0.225**

Table A.24: Comparison between physicians and POMDP (early prediction of 3rd screen with a_1). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage). Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			0.789	0.032
0	14/45 (31.11 %)	9/24 (37.5 %)		
1	31/45 (68.89 %)	15/24 (62.5 %)		
diagcopd			0.687	< 0.01
0	41/45 (91.11 %)	21/24 (87.5 %)		
1	4/45 (8.89 %)	3/24 (12.5 %)		
famHist			0.39	0.104
0	34/45 (75.56 %)	15/24 (62.5 %)		
1	11/45 (24.44 %)	9/24 (37.5 %)		
pCancHist			0.333	0.089
0	43/45 (95.56 %)	21/24 (87.5 %)		
1	2/45 (4.44 %)	3/24 (12.5 %)		

gender			1	< 0.001
1	30/45 (66.67 %)	16/24 (66.67 %)		
2	15/45 (33.33 %)	8/24 (33.33 %)		
race			**0.012**	0.253
1	43/45 (95.56 %)	20/24 (83.33 %)		
2	0/45 (0 %)	4/24 (16.67 %)		
4	2/45 (4.44 %)	0/24 (0 %)		
educat			0.072	0.216
1	0/26 (0 %)	1/15 (6.67 %)		
2	0/26 (0 %)	2/15 (13.33 %)		
4	10/26 (38.46 %)	3/15 (20 %)		
5	9/26 (34.62 %)	5/15 (33.33 %)		
6	4/26 (15.38 %)	2/15 (13.33 %)		
7	3/26 (11.54 %)	0/15 (0 %)		
8	0/26 (0 %)	2/15 (13.33 %)		
sctpreatt0			0.731	0.155
1	28/41 (68.29 %)	5/5 (100 %)		
2	9/41 (21.95 %)	0/5 (0 %)		
3	4/41 (9.76 %)	0/5 (0 %)		
sctmargins0			0.287	0.175
1	11/43 (25.58 %)	3/5 (60 %)		
2	24/43 (55.81 %)	2/5 (40 %)		
3	8/43 (18.6 %)	0/5 (0 %)		
BMI	26.52(4.71), 0	27.08(4.83), 0	0.668	< 0.001
smokeIntensity	28.22(10.07), 0	31.12(13.3), 0	0.442	< 0.001
smokeyr	46.33(5.51), 0	45.92(7.26), 0	0.807	< 0.001
smokeQuitTime	2(4.13), 1	3.33(5.11), 0	0.387	0.011

age	65.36(5.32), 3	66.79(4.45), 0	0.246	0.021
LargestDiam0	8.02(4.91), 0	1.46(2.93), 0	**< 0.001**	**0.397**

Table A.25: Comparison between physicians and POMDP (early prediction of post-screening with a_0). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage). Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			0.267	0.124
0	22/65 (33.85 %)	8/15 (53.33 %)		
1	43/65 (66.15 %)	7/15 (46.67 %)		
diagcopd			1	0.019
0	57/65 (87.69 %)	14/15 (93.33 %)		
1	8/65 (12.31 %)	1/15 (6.67 %)		
famHist			0.166	0.155
0	49/65 (75.38 %)	8/15 (53.33 %)		
1	16/65 (24.62 %)	7/15 (46.67 %)		
pCancHist			0.234	0.074
0	62/65 (95.38 %)	13/15 (86.67 %)		
1	3/65 (4.62 %)	2/15 (13.33 %)		
gender			1	< 0.001
1	43/65 (66.15 %)	10/15 (66.67 %)		
2	22/65 (33.85 %)	5/15 (33.33 %)		
race			**0.035**	0.241

1	61/65 (93.85 %)	12/15 (80 %)		
2	1/65 (1.54 %)	3/15 (20 %)		
4	3/65 (4.62 %)	0/15 (0 %)		
educat			0.361	0.148
1	0/35 (0 %)	1/12 (8.33 %)		
2	2/35 (5.71 %)	2/12 (16.67 %)		
4	11/35 (31.43 %)	4/12 (33.33 %)		
5	12/35 (34.29 %)	2/12 (16.67 %)		
6	6/35 (17.14 %)	1/12 (8.33 %)		
7	3/35 (8.57 %)	1/12 (8.33 %)		
8	1/35 (2.86 %)	1/12 (8.33 %)		
sctpreatt0			1	0.117
1	30/43 (69.77 %)	3/3 (100 %)		
2	9/43 (20.93 %)	0/3 (0 %)		
3	4/43 (9.3 %)	0/3 (0 %)		
sctpreatt1			0.28	0.108
1	37/48 (77.08 %)	4/7 (57.14 %)		
2	7/48 (14.58 %)	2/7 (28.57 %)		
3	4/48 (8.33 %)	1/7 (14.29 %)		
sctmargins0			0.268	0.156
1	12/45 (26.67 %)	2/3 (66.67 %)		
2	25/45 (55.56 %)	1/3 (33.33 %)		
3	8/45 (17.78 %)	0/3 (0 %)		
sctmargins1			0.745	0.092
1	8/47 (17.02 %)	1/7 (14.29 %)		
2	27/47 (57.45 %)	3/7 (42.86 %)		
3	12/47 (25.53 %)	3/7 (42.86 %)		

BMI	$26.26(4.62), 0$	$27.13(5.05), 0$	0.613	< 0.001
smokeIntensity	$28.57(10.36), 0$	$32.33(12.66), 0$	0.302	0.013
smokeyr	$46.12(5.94), 0$	$42.6(6.29), 0$	0.062	0.048
smokeQuitTime	$2.31(4.29), 1$	$4.2(5.2), 0$	0.116	0.031
age	$65.53(5.41), 6$	$65(5.59), 1$	0.763	< 0.001
LargestDiam0	$5.78(5.44), 0$	$1.33(2.79), 0$	$\mathbf{< 0.001}$	0.128
LargestDiam1	$6.53(4.74), 4$	$4.59(6.07), 1$	0.111	0.034

Table A.26: Comparison between physicians and POMDP (early prediction of post-screening with a_1). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage). Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

	Physician	POMDP	p	effect-size
cigsmok			0.309	0.107
0	27/78 (34.62 %)	7/13 (53.85 %)		
1	51/78 (65.38 %)	6/13 (46.15 %)		
diagcopd			1	< 0.001
0	70/78 (89.74 %)	12/13 (92.31 %)		
1	8/78 (10.26 %)	1/13 (7.69 %)		
famHist			1	< 0.001
0	58/78 (74.36 %)	10/13 (76.92 %)		
1	20/78 (25.64 %)	3/13 (23.08 %)		
pCancHist			1	< 0.001
0	73/78 (93.59 %)	12/13 (92.31 %)		

1	5/78 (6.41 %)	1/13 (7.69 %)		
gender			1	< 0.01
1	50/78 (64.1 %)	9/13 (69.23 %)		
2	28/78 (35.9 %)	4/13 (30.77 %)		
race			0.724	0.059
1	71/78 (91.03 %)	12/13 (92.31 %)		
2	4/78 (5.13 %)	1/13 (7.69 %)		
4	3/78 (3.85 %)	0/13 (0 %)		
educat			0.091	0.177
1	0/44 (0 %)	1/11 (9.09 %)		
2	3/44 (6.82 %)	1/11 (9.09 %)		
4	13/44 (29.55 %)	3/11 (27.27 %)		
5	16/44 (36.36 %)	1/11 (9.09 %)		
6	7/44 (15.91 %)	2/11 (18.18 %)		
7	4/44 (9.09 %)	1/11 (9.09 %)		
8	1/44 (2.27 %)	2/11 (18.18 %)		
sctpreatt0			1	0.117
1	30/43 (69.77 %)	3/3 (100 %)		
2	9/43 (20.93 %)	0/3 (0 %)		
3	4/43 (9.3 %)	0/3 (0 %)		
sctpreatt1			0.367	0.093
1	39/51 (76.47 %)	3/5 (60 %)		
2	8/51 (15.69 %)	1/5 (20 %)		
3	4/51 (7.84 %)	1/5 (20 %)		
sctpreatt2			1	0.055
1	32/42 (76.19 %)	7/9 (77.78 %)		
2	8/42 (19.05 %)	2/9 (22.22 %)		

3	1/42 (2.38 %)	0/9 (0 %)		
4	1/42 (2.38 %)	0/9 (0 %)		
sctmargins0			0.268	0.156
1	12/45 (26.67 %)	2/3 (66.67 %)		
2	25/45 (55.56 %)	1/3 (33.33 %)		
3	8/45 (17.78 %)	0/3 (0 %)		
sctmargins1			0.554	0.076
1	8/50 (16 %)	1/5 (20 %)		
2	29/50 (58 %)	2/5 (40 %)		
3	13/50 (26 %)	2/5 (40 %)		
sctmargins2			0.746	0.109
1	5/43 (11.63 %)	0/9 (0 %)		
2	27/43 (62.79 %)	6/9 (66.67 %)		
3	11/43 (25.58 %)	3/9 (33.33 %)		
BMI	$26.73(4.78), 0$	$26.52(5.19), 0$	0.755	< 0.001
smokeIntensity	$28.49(10.29), 0$	$36.54(15.99), 0$	0.075	0.035
smokeyr	$45.33(5.93), 0$	$40.31(7.78), 0$	**0.043**	0.053
smokeQuitTime	$2.29(4.24), 1$	$4.77(5.78), 0$	0.086	0.033
age	$64.89(5.46), 6$	$62.92(6.49), 1$	0.3	0.013
LargestDiam0	$4.82(5.41), 0$	$1.54(2.96), 0$	**0.027**	0.053
LargestDiam1	$5.79(5.14), 4$	$3.17(4.34), 1$	0.096	0.032
LargestDiam2	$5.24(5.33), 11$	$5.17(3.81), 1$	0.774	< 0.001

Table A.27: Comparison between physicians and POMDP (early prediction of post-screening with a_2). For quantitative covariates: "mean (sd), missing data", and for categorical covariates: "effective/ total effective (percentage). Student test or Wilcoxon-Mann-Whitney test, χ^2 test or Fisher test used when appropriate.

A.13.1 Variables' categories

Variable	Variable Explanation	Categories
cigsmok	Smoking status at T0 Participant	0="Former" 1="Current"
diagcopd	COPD: Ever diagnosed prior to trial?	0="No" 1="Yes"
famHist	Family History of lung cancer, 1st degree relative	0="No" 1="Yes"
pCancHist	Personal cancer history, all types of cancer	0="No" 1="Yes"
gender		1="Male" 2="Female"
race		1="White" 2="Black" 3="Hispanic" 4="Asian" 5="American Indian or Alaskan Native" 6="Native Hawaiian or Other Pacific Islander"
BMI	Body mass Index	continuous
smokeIntensity	Average number of cigarettes per day	continuous
smokeYr	Total years of smoking	continuous
smokeQuitTime	Time of Quitting Smoking	continuous
age	Age at T0	continuous

sct_pre_att0_2	Predominant attenuation T0-2	1="Soft Tissue" 2="Ground Glass" 3="Mixed" 4="Other"
sct_margins0_2	Margins T0-2	1="Spiculated" 2="Smooth" 3="Poorly defined"
LargestDiam0_2	Largest nodule diameter (mm) T0-2	continuous

Table A.28: The categories and description of each variable.

A.14 Comparison of POMDP and DBN:

|	POMDP			DBN 2016		
	TN rate	TP rate/Recall	Precision	TN rate	TP rate/Recall	Precision
	Screening T0 (Cancers = 32, Non-Cancers = 1,047)					
a_0	0.47	0.97	0.05	0.43	1.00	0.06
	Screening T1 (Cancers = 17, Non-Cancers = 1,030)					
a_0	0.47	0.67	0.02	0.43	0.47	0.01
$a_{0.5}$	0.47	0.67	0.02			
a_1	0.34	0.98	0.02	0.30	0.99	0.02
	Screening T2 (Cancers = 21, Non-Cancers = 1,009)					
a_0	0.47	0.55	0.02	0.42	0.30	0.01
$a_{0.5}$	0.47	0.55	0.02			
a_1	0.34	0.69	0.02	0.30	0.48	0.01
$a_{1.5}$	0.34	0.70	0.02			
a_2	0.25	0.96	0.03	0.18	1.00	0.03
	Post Screening (Cancers = 19, Non-Cancers = 990)					
a_0	0.48	0.68	0.02	0.43	0.49	0.02
$a_{0.5}$	0.47	0.68	0.02			
a_1	0.35	0.81	0.02	0.30	0.71	0.02
$a_{1.5}$	0.34	0.81	0.02			
a_2	0.25	0.93	0.02	0.18	0.88	0.02

Table A.29: POMDP Vs DBN 2016 model. Testing data.

A.15 POMDP performance - comparison of observation models

	POMDP – DBN			POMDP – Exhaustive search			POMDP – LR		
	TN rate	TP rate/Recall	Precision	TN rate	TP rate/Recall	Precision	TN rate	TP rate/Recall	Precision
Screening T0 (Cancers = 32, Non-Cancers = 1,047)									
a_0	0.84	0.83	0.14	0.76	0.87	0.1	0.9	0.79	0.2
Screening T1 (Cancers = 17, Non-Cancers = 1,030)									
a_0	0.84	0.31	0.03	0.76	0.36	0.03	0.9	0.21	0.03
$a_{0.5}$	0.84	0.31	0.03	0.7	0.74	0.04	0.88	0.45	0.06
a_1	0.74	0.69	0.04	0.7	0.8	0.04	0.85	0.63	0.07
Screening T2 (Cancers = 21, Non-Cancers = 1,009)									
a_0	0.84	0.19	0.02	0.77	0.23	0.02	0.9	0.12	0.02
$a_{0.5}$	0.84	0.19	0.02	0.7	0.38	0.03	0.88	0.15	0.03
a_1	0.75	0.34	0.03	0.69	0.4	0.03	0.85	0.22	0.03
$a_{1.5}$	0.74	0.36	0.03	0.64	0.74	0.04	0.84	0.37	0.05
a_2	0.69	0.71	0.05	0.63	0.8	0.04	0.84	0.37	0.05
Post Screening (Cancers = 19, Non-Cancers = 990)									
a_0	0.84	0.24	0.03	0.77	0.32	0.03	0.9	0.15	0.03
$a_{0.5}$	0.84	0.24	0.03	0.7	0.4	0.03	0.89	0.19	0.03
a_1	0.75	0.39	0.03	0.63	0.42	0.03	0.86	0.25	0.03
$a_{1.5}$	0.74	0.39	0.03	0.64	0.46	0.03	0.85	0.26	0.03
a_2	0.69	0.47	0.03	0.63	0.5	0.03	0.85	0.26	0.03

Table A.30: POMDP model performance using the DBN, an Exhaustive search model (all combinations of observations), and a logistic regression model, respectively. Testing data.

A.16 Box plots of all cases:
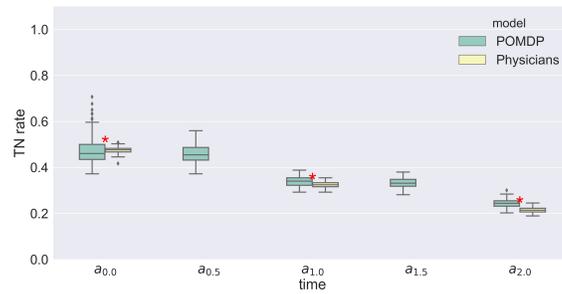
(a) TN rate.
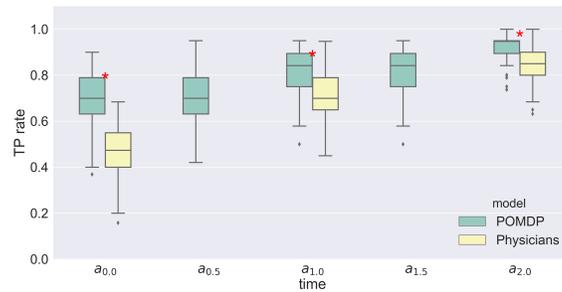
(b) Recall/TP rate.
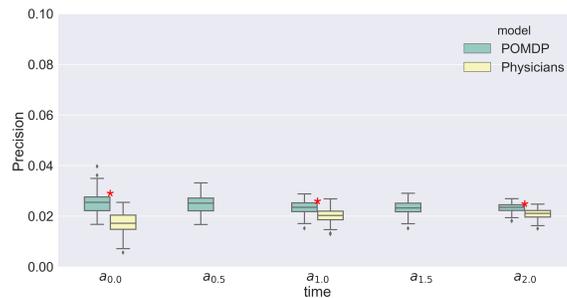
(c) Precision.

Figure A.24: Box plots of the performance (TN, TP, precision) of the POMDP and physicians on cases from the NLST testing set, from the start of the trial through to last screening. Blue and yellow represent the POMDP and experts, respectively. Red stars depict instances where the performance measure between the physicians and model are significantly different.

Bibliography

[1] L. Pack Kaelbling, M. L. Littman, A. R. Cassandra, L. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and Acting in Partially Observable Stochastic Domains," *Artificial Intelligence* **101** (1998), no. 1-2 99–134.

[2] R. S. Sutton and A. G. Barto, "Introduction to Reinforcement Learning," *Learning* **4** (1998), no. 1996 1–5.

[3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: A Cancer Journal for Clinicians* **68** (1, 2018) 7–30.

[4] Surveillance Epidemiology and End Results., "Cancer Statistics Review, 1975-2011 - Previous Version - SEER Cancer Statistics Review," 2014.

[5] National Lung Screening Trial Research Team, D. R. Aberle, A. M. Adams, C. D. Berg, W. C. Black, J. D. Clapp, R. M. Fagerstrom, I. F. Gareen, C. Gatsonis, P. M. Marcus, J. D. Sicks, and P. P.-t.-p. Transmission, "Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening," *The New England journal of medicine* **365** (8, 2011) 333–340.

[6] B. Bunn, "IASLC Issues Statement on Lung Cancer Screening with Low-Dose Computed Tomography."

[7] Z. Saghir, A. Dirksen, H. Ashraf, K. S. Bach, J. Brodersen, P. F. Clementsen, M. Døssing, H. Hansen, K. F. Kofoed, K. R. Larsen, J. Mortensen, J. F. Rasmussen,

N. Seersholm, B. G. Skov, H. Thorsen, P. Tønnesen, and J. H. Pedersen, "CT screening for lung cancer brings forward early disease. The randomised Danish lung cancer screening trial: Status after five annual screening rounds with low-dose CT," *Thorax* **67** (4, 2012) 296–301.

[8] D. R. Aberle, S. DeMello, C. D. Berg, W. C. Black, B. Brewer, T. R. Church, K. L. Clingan, F. Duan, R. M. Fagerstrom, I. F. Gareen, C. A. Gatsonis, D. S. Gierada, A. Jain, G. C. Jones, I. Mahon, P. M. Marcus, J. M. Rathmell, J. Sicks, D. Sarah, C. D. Berg, W. C. Black, B. Brewer, T. R. Church, K. L. Clingan, F. Duan, R. M. Fagerstrom, I. F. Gareen, C. A. Gatsonis, D. S. Gierada, A. Jain, G. C. Jones, I. Mahon, P. M. Marcus, J. M. Rathmell, J. Sicks, and N. Team, "Results of the two incidence screenings in the National Lung Screening Trial," *N. Engl. J. Med.* **369** (2013), no. 10 920–931.

[9] E. F. Patz, P. Pinsky, C. Gatsonis, J. D. Sicks, B. S. Kramer, M. C. Tammemägi, C. Chiles, W. C. Black, and D. R. Aberle, "Overdiagnosis in Low-Dose Computed Tomography Screening for Lung Cancer," *JAMA Internal Medicine* **174** (2014), no. 2 269–274.

[10] M. Wiering and M. van Otterlo, *Reinforcement Learning*, vol. 12. 2012.

[11] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts, "Modeling medical treatment using Markov decision processes," *Operations Research and Health Care* (2005) 597–616.

[12] G. Tusch, "Optimal sequential decisions in liver transplantation based on a POMDP model," in *ECAI*, pp. 186–190, 2000.

[13] M. Leshno, Z. Halpern, and N. Arber, "Cost-effectiveness of colorectal cancer screening in the average risk population," *Health care management science* **6** (2003), no. 3 165–174.

[14] T. Ayer, O. Alagoz, and N. K. Stout, "OR Forum—A POMDP Approach to Personalize Mammography Screening Decisions," *Operations Research* **60** (2012), no. 5 1019–1034.

[15] F. S. Erenay, O. Alagoz, and A. Said, "Optimizing Colonoscopy Screening for Colorectal Cancer Prevention and Surveillance," *Manufacturing & Service Operations Management* **16** (2014), no. 3 381–400.

[16] J. Zhang, B. T. Denton, H. Balasubramanian, N. D. Shah, and B. A. Inman, "Optimization of Prostate Biopsy Referral Decisions," *Manufacturing & Service Operations Management* **14** (2012), no. 4 529–547.

[17] S. Elson, R. Hiatt, and C. Anton, *The Athena breast health network: Developing a rapid learning system in breast cancer prevention, screening, treatment, and care*, vol. 140. Springer US, 7, 2013.

[18] P. B. Bach, M. W. Kattan, M. D. Thornquist, M. G. Kris, R. C. Tate, M. J. Barnett, L. J. Hsieh, and C. B. Begg, "Variations in Lung cancer risk among smokers," *Journal of the National Cancer Institute* **95** (2003), no. 6 470–478.

[19] C. J. Etzel and P. B. Bach, "Estimating individual risk for lung cancer," *Seminars in Respiratory and Critical Care Medicine* **32** (2011), no. 1 3–9.

[20] K. A. Cronin, M. H. Gail, Z. Zou, P. B. Bach, J. Virtamo, and D. Albanes, "Validation of a model of lung cancer risk prediction among smokers," *Journal of the National Cancer Institute* **98** (2006), no. 9 637–640.

[21] K. Jayasurya, G. Fung, S. Yu, and D.-O. C, "Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy,".

[22] O. Y. Raji, S. W. Duffy, O. F. Agbaje, S. G. Baker, D. C. Christiani, A. Cassidy, and J. K. Field, "Predictive accuracy of the Liverpool Lung Project risk model for stratifying patients for computed tomography screening for lung cancer: a case-control and cohort validation study," *Ann. Intern. Med.* **157** no. 4 242–250.

[23] M. R. Spitz, W. K. Hong, C. I. Amos, X. Wu, M. B. Schabath, Q. Dong, S. Shete, and C. J. Etzel, "A risk model for prediction of lung cancer," *Journal of the National Cancer Institute* **99** (2007), no. 9 715–726.

[24] P. Maisonneuve, V. Bagnardi, M. Bellomi, L. Spaggiari, G. Pelosi, C. Rampinelli, R. Bertolotti, N. Rotmensz, J. K. Field, A. DeCensi, and G. Veronesi, "Lung cancer risk prediction to select smokers for screening CT - A model based on the Italian COSMOS trial," *Cancer Prevention Research* **4** (2011), no. 11 1778–1789.

[25] C. M. Tammemagi, P. F. Pinsky, N. E. Caporaso, P. A. Kvale, W. G. Hocking, T. R. Church, T. L. Riley, J. Commins, M. M. Oken, C. D. Berg, and P. C. Prorok, "Lung cancer risk prediction: Prostate, lung, colorectal and ovarian cancer screening trial models and validation," *Journal of the National Cancer Institute* **103** (2011), no. 13 1058–1068.

[26] L. Yu and H. Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution," *International Conference on Machine Learning (ICML)* (2003) 1–8.

[27] K. Polat and S. Güneş, "Computer aided medical diagnosis system based on principal component analysis and artificial immune recognition system classifier algorithm," *Expert Systems with Applications* **34** (2008), no. 1 773–779.

[28] H. Boström, "Maximizing the area under the ROC curve using incremental reduced error pruning," *Proceedings of the International Conference on Machine Learning 2005 workshop on ROC analysis in machine learning* (2005).

[29] I. Hendrickx and A. D. Van Bosch, "Hybrid algorithms with instance-based classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3720 LNAI, pp. 158–169, 2005.

[30] P. Thamilselvan and J. G. Sathiaseelan, "An enhanced k nearest neighbor method to detecting and classifying MRI lung cancer images for large amount data," *International Journal of Applied Engineering Research* **11** (2016), no. 6 4223–4229.

[31] M. Annette, M. C. Tammemagi, J. R. Mayo, H. Roberts, G. Liu, K. Soghrati, K. Yasufuku, S. Martel, F. Laberge, M. Gingras, A.-K. Sukhinder, C. D. Berg, K. Evans, R. Finley, J. Yee, J. English, P. Nasute, J. Goffin, S. Puksa, L. Stewart, S. Tsai, M. R. Johnston, D. Manos, G. Nicholas, G. D. Goss, J. M. Seely, K. Amjadi, A. Tremblay, P. Burrowes, M. Paul, R. Bhatia, M.-S. Tsao, S. Lam, A. McWilliams, M. C. Tammemagi, J. R. Mayo, H. Roberts, G. Liu, K. Soghrati, K. Yasufuku, S. Martel, F. Laberge, M. Gingras, S. Atkar-Khattra, C. D. Berg, K. Evans, R. Finley, J. Yee, J. English, P. Nasute, J. Goffin, S. Puksa, L. Stewart, S. Tsai, M. R. Johnston, D. Manos, G. Nicholas, G. D. Goss, J. M. Seely, K. Amjadi, A. Tremblay, P. Burrowes, P. MacEachern, R. Bhatia, M.-S. Tsao, and S. Lam, "Probability of cancer in pulmonary nodules detected on first screening CT.," *The New England journal of medicine* **369** (2013), no. 10 910–9.

[32] B.-C. Kim, Y. S. Sung, and H.-I. Suk, "Deep feature learning for pulmonary nodule classification in a lung CT," in *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*, pp. 1–3, IEEE, 2, 2016.

[33] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, C.-S. Huang, D. Shen, and C.-M. Chen, "Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans.," *Scientific reports* **6** (2016) 24454.

[34] S. K, Jayanthiladevi A, and Kavitha C, "Automatic detection of lung cancer nodules by employing intelligent fuzzy c- means and support vector machine,".

[35] A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, "Computer-aided diagnosis system for lung nodules based on computed tomography using shape analysis, a genetic algorithm, and SVM," *Medical and Biological Engineering and Computing* **55** (10, 2017) 1129–1146.

[36] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences: temporal and static modeling," *Computer Vision and Image Understanding* **91** (2003), no. 1 160–187.

[37] K. Bhaskaran, A. Gasparrini, S. Hajat, L. Smeeth, and B. Armstrong, "Time series regression studies in environmental epidemiology," *International Journal of Epidemiology* **42** (8, 2013) 1187–1195.

[38] A. M., K. D.R., Y. M., and E. M., "Time series analysis of cholera in matlab, bangladesh, during 1988-2001," *Journal of Health, Population and Nutrition* **31** (2013), no. 1 11–19.

[39] W. Hu, S. Tong, K. Mengersen, and D. Connell, "Weather Variability and the Incidence of Cryptosporidiosis: Comparison of Time Series Poisson Regression and SARIMA Models," *Annals of Epidemiology* **17** (2007), no. 9 679–688.

[40] M. A. J. van Gerven, B. G. Taal, P. J. F. Lucas, M. A. J. V. Gerven, B. G. Taal, and P. J. F. Lucas, "Dynamic Bayesian networks as prognostic models for clinical patient management," *Journal of biomedical informatics* **41** (2008), no. 4 515–529.

[41] A. E. Nicholson and M. J. Flores, "Combining state and transition models with dynamic Bayesian networks," *Ecological Modelling* **222** (2011), no. 3 555–566.

[42] T. Kohda and W. Cui, "Risk-based reconfiguration of safety monitoring system using dynamic Bayesian network," *Reliability Engineering and System Safety* **92** (2007), no. 12 1716–1723.

[43] Y. Li, S. Swift, and A. Tucker, "Modelling and analysing the dynamics of disease progression from cross-sectional studies," *Journal of Biomedical Informatics* **46** (2013), no. 2 266–274.

[44] K. Daphne and N. Friedman, *Probabilistic graphical models: principles and techniques.* MIT press, 2009.

[45] N. Friedman, K. Murphy, and S. Russell, "Learning the Structure of Dynamic Probabilistic Networks," *15th Annual Conference on Uncertainty in Artificial Intelligence* (1999) 139–147.

[46] H. Ltifi, G. Trabelsi, M. B. Ayed, and A. M. Alimi, "Dynamic Decision Support System Based on Bayesian Networks Application to fight against the Nosocomial Infections," *arXiv preprint arXiv:1211.2126* **1** (2012), no. 1 22–29.

[47] T. Charitos, L. C. van der Gaag, S. Visscher, K. A. M. Schurink, P. J. F. Lucas, L. C. V. D. Gaag, S. Visscher, K. A. M. Schurink, P. J. F. Lucas, L. C. van der Gaag, S. Visscher, K. A. M. Schurink, and P. J. F. Lucas, "A dynamic Bayesian network for diagnosing ventilator-associated pneumonia in {ICU} patients," *Expert Systems with Applications* **36** (2009), no. 2 1249–1258.

[48] M. Verduijn, P. M. J. Rosseel, N. Peek, E. de Jonge, B. A. J. M. de Mol, E. D. Jonge, and B. A. J. M. D. Mol, "Prognostic Bayesian networks. II: An application in the domain of cardiac surgery," *Journal of Biomedical Informatics* **40** (2007), no. 6 619–630.

[49] A. Costa, M. Freestone, W. Marsh, J. Coid, A. C. Constantinou, M. Freestone, W. Marsh, and J. Coid, "Causal inference for violence risk management and decision support in forensic psychiatry," *Decision Support Systems* **80** (2015) 42–55.

[50] G. Cuaya, A. Muñoz-Meléndez, L. N. N. Carrera, E. F. Morales, I. Quiñones, A. I. Pérez, A. Alessi, M.-M. Angélica, L. N. N. Carrera, E. F. Morales, I. Quiñones, A. I. Pérez, A. Alessi, A. Mu??oz-Mel??ndez, L. N. N. Carrera, E. F. Morales, I. Qui??ones, A. I. P??rez, and A. Alessi, "A dynamic Bayesian network for estimating the risk of falls from real gait data," *Medical and Biological Engineering and Computing* **51** (2, 2013) 29–37.

[51] E. W. Watt and A. A. T. Bui, "Evaluation of a dynamic bayesian belief network to predict osteoarthritic knee pain using data from the osteoarthritis initiative.," in *AMIA 2008 Symposium*, pp. 788–92, 2008.

[52] P. M. Pardalos, P. Xanthopoulos, M. Zervakis, K. P. Exarchos, G. Rigas, Y. Goletsis, and D. I. Fotiadis, *Modelling of Oral Cancer Progression Using Dynamic Bayesian Networks*, vol. 25. Springer US, 2012.

[53] A. Stojadinovic, A. Bilchik, D. Smith, J. S. Eberhardt, E. B. Ward, A. Nissan, E. K. Johnson, M. Protic, G. E. Peoples, I. Avital, and S. R. Steele, "Clinical decision support and individualized prediction of survival in colon cancer: Bayesian belief network model," *Annals of Surgical Oncology* **20** (2013), no. 1 161–174.

[54] R. M. Austin and A. Onisko, "Increased cervical cancer risk associated with extended screening intervals after negative human papillomavirus test results : Bayesian risk estimates using the Pittsburgh Cervical Cancer Screening Model," *Journal of the American Society of Cytopathy* **1** (2015), no. 1 9–14.

[55] Nicandro, Cruz-Ramírez, H. GG, Acosta-Mesa, C.-C. Humberto, L. A, Nava-Fernández, R. EE, Barrientos-Martínez, N. Cruz-Ramírez, H. G. Acosta-mesa,

H. Carrillo-calvet, L. A. Nava-fernández, R. E. Barrientos-martínez, L. Alonso Nava-Fernández, and R. E. Barrientos-martínez, "Diagnosis of breast cancer using Bayesian networks: a case study.," *Computers in Biology and Medicine* **37** (2007), no. 11 1553–1564.

[56] M. Velikova, P. J. F. P. J. F. Lucas, N. Ferreira, M. Samulski, N. Karssemeijer, C. N. de Ferreira, N. Ferreira, M. Samulski, and N. Karssemeijer, "A decision support system for breast cancer detection in screening programs.," *ECAI* **178** (2008) 658–662.

[57] O. Gevaert, F. D. Smet, D. Timmerman, Y. Moreau, and B. D. Moor, "Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks.,".

[58] J. H. Oh, J. Craft, R. Lozi, M. Vaidya, Y. Meng, J. O. Deasy, J. D. Bradley, and I. Naqa, "A Bayesian network approach for modeling local failure in lung cancer.,".

[59] P. J. F. Lucas, N. C. D. Bruijn, K. Schurink, N. C. de Bruijn, and K. Schurink, "A probabilistic and decision-theoretic approach to the management of infectious disease at the {ICU},".

[60] S. M. Maskery, H. Hu, J. Hooke, C. D. Shriver, and M. N. Liebman, "A Bayesian derived network of breast pathology co-occurrence.,".

[61] A. R. Cassandra, M. L. Littman, and N. L. Zhang, "Incremental Pruning: A Simple, Fast, Exact Method for Partially Observable Markov Decision Processes," *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence* (2, 1997) 54–61.

[62] W. S. Lovejoy, "A survey of algorithmic methods for partially observed Markov decision processes," *Annals of Operations Research* **28** (1991), no. 1 47–65.

[63] E. J. Sondik and N. M. Apr, "The Optimal Control of Partially Observable Markov Processes Over the Infinite Horizon : Discounted Costs The Optimal Control of Partially Observable Markov," *Source: Operations Research Operations Research Society of America* **26** (1978), no. 2 282–304.

[64] C. C. White, "A survey of solution techniques for the partially observed Markov decision process," *Annals of Operations Research* **32** (12, 1991) 215–230.

[65] S. Russell and P. Norvig, *Artificial Intelligence_ A Modern Approach (3rd Edition)*. 2010.

[66] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. 2006.

[67] L. Chrisman, "Reinforcement Learning with Perceptual Aliasing: The Perceptual Distinctions Approach," *Proceedings of the Tenth National Conference on Artificial Intelligence* (1992) 183–188.

[68] D. Mescheder, K. Tuyls, and M. Kaisers, "Opponent Modeling with pomdps," in *Belgian/Netherlands Artificial Intelligence Conference*, 2011.

[69] Mccallum and A. Kachites, *RL with Selective Perception and Hidden State*. PhD thesis, 1996.

[70] A. Y. Ng and S. Russell, "Algorithms for inverse reinforcement learning," *Proceedings of the Seventeenth International Conference on Machine Learning* (2000) 663–670.

[71] J. Choi and K. E. Kim, "Inverse reinforcement learning in partially observable environments," in *IJCAI International Joint Conference on Artificial Intelligence*, vol. 12, pp. 1028–1033, JMLR.org, 2009.

[72] C. C. Bennett and K. Hauser, "Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach," *Artificial Intelligence in Medicine* **57** (2013), no. 1 919.

[73] M. Hauskrecht and H. Fraser, "Planning treatment of ischemic heart disease with partially observable Markov decision processes," *Artificial Intelligence in Medicine* **18** (2000), no. 3 221–244.

[74] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Twenty-first international conference on Machine learning - ICML '04*, p. 1, 2004.

[75] M. Babeş-Vroman, V. Marivate, K. Subramanian, and M. Littman, "Apprenticeship learning about multiple intentions," *Proceedings of the 28th International Conference on Machine Learning, ICML 2011* (2011) 897–904.

[76] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey, "Maximum Entropy Inverse Reinforcement Learning.," in *AAAI Conference on Artificial Intelligence*, pp. 1433–1438, 2008.

[77] M. C. Vroman, *Maximum Likelihood Inverse Reinforcement Learning.* PhD thesis, 2014.

[78] B. Ziebart, *Modeling Purposeful Adaptive Behavior with the Principle of Maximum Causal Entropy.* PhD thesis, 2010.

[79] A. Tsoukalas, T. Albertson, and I. Tagkopoulos, "From Data to Optimal Decision Making: A Data-Driven, Probabilistic Machine Learning Approach to Decision Support for Patients With Sepsis," *JMIR Medical Informatics* **3** (2015), no. 1 e11.

[80] J. E. Goulionis, A. Vozikis, V. K. Benos, and D. Nikolakis, "On the decision rules of cost-effective treatment for patients with diabetic foot syndrome," *ClinicoEconomics and Outcomes Research* **2** (8, 2010) 121–126.

[81] A. Vozikis, J. E. Goulionis, and V. K. Benos, "The partially observable Markov decision processes in healthcare: an application to patients with ischemic heart disease {(IHD)}," *Operational Research* (2010).

[82] J. Goulionis and A. Vozikis, "Medical decision making for patients with Parkinson disease under Average Cost Criterion TL - 6," *Aust. New Zealand Health Policy* **6 VN - re**.

[83] O. Alagoz, J. Chhatwal, and E. S. Burnside, "Optimal policies for reducing unnecessary follow-up mammography exams in breast cancer diagnosis," *Decision Analysis* **10** (2013), no. 3 200–224.

[84] L. M. Maillart, J. S. Ivy, S. Ransom, and K. Diehl, "Assessing Dynamic Breast Cancer Screening Policies," *Operations Research* **56** (2008), no. 6 1411–1427.

[85] V. Schlecht, J. Zhang, B. Denton, and B. Inman, "Optimization of PSA-Based Screening Decisions for Prostate Cancer Detection," *Working paper* **5** (2009), no. January 7–24.

[86] J. Chhatwal, O. Alagoz, and E. S. Burnside, "Optimal Breast Biopsy Decision-Making Based on Mammographic Features and Demographic Factors," *Operations Research* **58** (2010), no. 6 1577–1591.

[87] O. Alagoz, T. Ayer, and F. Safa Erenay, "Operations Research Models for Cancer Screening," *Wiley Encyclopedia of Operations Research and Management Science* (2011).

[88] M. A. Greenwood-Hickman, L. Jones, K. Marsh, M. P. Rivera, P. Molina, D. Reuland, J. M. Bowling, K. Birchard, S. Maygarden, and L. Henderson, "Do perceived barriers to lung cancer screening differ between attending physicians and residents?," *J. Clin. Oncol.* **34** (2016), no. 7_suppl 192.

[89] American Joint Committee on Cancer, "Lung Cancer , Staging," 2009.

[90] M. J. Druzdzel, "SMILE : Structural Modeling, Inference, and Learning Engine and GeNIe: A Development Environment for Graphical Decision-Theoretic Models,"

Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99)
99 (1999) 342–343.

[91] L. Uusitalo, "Advantages and challenges of Bayesian networks in environmental modelling," *Ecological Modelling* **203** (2007), no. 3-4 312–318.

[92] P. Kraaijeveld and M. J. Druzdzel, "GeNIeRate: An interactive generator of diagnostic Bayesian network models," in *Proc. 16th Int. Workshop on Principles of Diagnosis*, pp. 175–180, 2005.

[93] R. H. Poe, R. E. Tobin, P. R.H., and T. R.E., "Sensitivity and specificity of needle biopsy in lung malignancy," *American Review of Respiratory Disease* **122** (11, 1980) 725–729.

[94] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering* **21** (2009), no. 9 1263–1284.

[95] K. Orphanou, A. Stassopoulou, and E. Keravnou, "DBN-extended: A Dynamic Bayesian network model extended with temporal abstractions for coronary heart disease prognosis.," *IEEE journal of biomedical and health informatics* **20** (2015), no. 3 944–952.

[96] S. C. van't Westeinde, H. J. de Koning, D. M. Xu, H. C. Hoogsteden, and R. J. van Klaveren, "How to deal with incidentally detected pulmonary nodules less than 10 mm in size on CT in a healthy person," *Lung Cancer* **60** (2008), no. 2 151–159.

[97] S. Shen, A. A. T. Bui, J. Cong, and W. Hsu, "An automated lung segmentation approach using bidirectional chain codes to improve nodule detection accuracy," *Computers in Biology and Medicine* **57** (2, 2015) 139–149.

[98] O. Alagoz, T. Ayer, and F. Safa Erenay, "Operations Research Models for Cancer Screening," *Wiley Encyclopedia of Operations Research and Management Science* (2011).

[99] M. C. Tammemägi, H. A. Katki, W. G. Hocking, T. R. Church, N. Caporaso, P. A. Kvale, A. K. Chaturvedi, G. A. Silvestri, T. L. Riley, J. Commins, and C. D. Berg, "Selection Criteria for Lung-Cancer Screening," *New England Journal of Medicine* **368** (2013), no. 8 728–736.

[100] P. Petousis, A. Naeim, A. Mosleh, and W. Hsu, "Evaluating the Impact of Uncertainty on Risk Prediction: Towards More Robust Prediction Models," in *AMIA - Annual Symposium proceedings*, 2018.

[101] C. J. D'Orsi, *ACR BI-RADS atlas: breast imaging reporting and data system*. American College of Radiology, 2013.

[102] M. Alger, "Deep Inverse Reinforcement Learning," tech. rep., 2016.

[103] Chelsea Finn, "Deep RL Bootcamp Lecture 10B Inverse Reinforcement Learning - YouTube," 2017.

[104] S. Klein, J. P. Pluim, M. Staring, and M. A. Viergever, "Adaptive stochastic gradient descent optimisation for image registration," *International Journal of Computer Vision* **81** (2009), no. 3 227–239.

[105] M. Hauskrecht and H. Milos, "Dynamic decision making in stochastic partially observable medical domains: Ischemic heart disease example," in *Lecture Notes in Computer Science*, pp. 296–299. springerlink.

[106] P. Petousis, S. X. Han, D. Aberle, and A. A. Bui, "Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network," *Artificial Intelligence in Medicine* **72** (2016) 42–55.

[107] A. M. Alaa, K. H. Moon, W. Hsu, and M. Van Der Schaar, "ConfidentCare: A Clinical Decision Support System for Personalized Breast Cancer Screening," *IEEE Transactions on Multimedia* **18** (2016), no. 10 1942–1955.

[108] E. S. Burnside, J. Davis, J. Chhatwal, O. Alagoz, M. J. Lindstrom, B. M. Geller, B. Littenberg, K. a. Shaffer, C. E. Kahn, and C. D. Page, "Probabilistic computer model developed from clinical data in national mammography database format to classify mammographic findings.," *Radiology* **251** (2009), no. 3 663–672.

[109] M. Van der Heijden, M. Velikova, and P. J. F. Lucas, "Learning Bayesian networks for clinical time series analysis," *Journal of Biomedical Informatics* **48** (2014) 94–105.

[110] P. Petousis, S. X. Han, W. Hsu, and A. A. Bui, "Generating Reward Functions using IRL Towards Individualized Cancer Screening," in *CEUR Workshop Proceedings*, vol. 2142, pp. 109–120, 2018.

[111] P. P. Massion and R. C. Walker, "Indeterminate pulmonary nodules: Risk for having or for developing lung cancer?," 12, 2014.

[112] K. P. Murphy, "The Bayes Net Toolbox for Matlab," *Computing Science and Statistics* **33** (2001), no. 2 1024–1034.

[113] D. Koller and N. Friedman, *Probabilistic Graphical Models- Principles and Techniques*, vol. 53. 1989.

[114] M. C. Tammemagi, S. C. Lam, A. M. Mcwilliams, and D. D. Sin, "Incremental Value of Pulmonary Function and Sputum DNA Image Cytometry in Lung Cancer Risk Prediction,".

[115] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, vol. 136. LAWRENCE ERLBAUM ASSOCIATES, second edi ed., 1988.

[116] M. Tomczak and E. Tomczak, "The need to report effect size estimates revisited. An overview of some recommended measures of effect size," *Trends in Sport Sciences* **1** (2014), no. 21 19–25.

[117] C. O. Fritz, P. E. Morris, and J. J. Richler, "Effect size estimates: Current use, calculations, and interpretation," *Journal of Experimental Psychology: General* **141** (2, 2012) 2–18.

[118] C. S. White, E. Dharaiya, E. Campbell, and L. Boroczky, "The Vancouver Lung Cancer Risk Prediction Model: Assessment by Using a Subset of the National Lung Screening Trial Cohort," *Radiology* **283** (2016) 152627.

[119] B. J. McKee, S. M. Regis, A. B. McKee, S. Flacke, and C. Wald, "Performance of ACR Lung-RADS in a Clinical CT Lung Screening Program," *Journal of the American College of Radiology* **13** (3, 2016) R25–R29.

[120] P. F. Pinsky, D. S. Gierada, W. Black, R. Munden, H. Nath, D. Aberle, and E. Kazerooni, "Performance of lung-RADS in the national lung screening trial: A retrospective assessment," *Annals of Internal Medicine* **162** (2015), no. 7 485–491.

[121] C. I. Henschke, R. Yip, D. F. Yankelevitz, and J. P. Smith, "Definition of a positive test result in computed tomography screening for lung cancer," *Annals of Internal Medicine* **158** (2013), no. 4 246–252.

[122] American Lung Association, "Lung Cancer Fact Sheet," 2018.

[123] S. Shen, S. X. Han, D. R. Aberle, A. A. T. Bui, and W. Hsu, "An Interpretable Deep Hierarchical Semantic Convolutional Neural Network for Lung Nodule Malignancy Classification," *arXiv preprint arXiv:1806.00712* (6, 2018).

[124] A. Murphy, M. Skalski, and F. Gaillard, "The utilisation of convolutional neural networks in detecting pulmonary nodules: a review," *The British Journal of Radiology* **91** (10, 2018) 20180028.

[125] L. C. Walter and K. E. Covinsky, "Cancer Screening in Elderly Patients," *Jama* **285** (6, 2001) 2750.

[126] B. Fisher, J. P. Costantino, D. L. Wickerham, C. K. Redmond, M. Kavanah, W. M. Cronin, V. Vogel, A. Robidoux, N. Dimitrov, J. Atkins, M. Daly, S. Wieand, E. Tan-Chiu, L. Ford, and N. Wolmark, "Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study.," *Journal of the National Cancer Institute* **90** (9, 1998) 1371–88.

[127] "Patient education: Medications for the prevention of breast cancer (Beyond the Basics) - UpToDate."

[128] M. H. Gail, "Personalized estimates of breast cancer risk in clinical practice and public health," *Statistics in Medicine* **30** (5, 2011) 1090–1104.

[129] E. Peters, J. Hibbard, P. Slovic, and N. Dieckmann, "Numeracy skill and the communication, comprehension, and use of risk-benefit information," 5, 2007.

[130] R. Al-Abri and A. Al-Balushi, "Patient satisfaction survey as a tool towards quality improvement," 1, 2014.

[131] N. Timmins, "NHS goes to the PROMS," *BMJ* **336** (6, 2008) 1464–1465.

[132] J. A. Sterne, I. R. White, J. B. Carlin, M. Spratt, P. Royston, M. G. Kenward, A. M. Wood, and J. R. Carpenter, "Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.," *BMJ (Clinical research ed.)* **338** (6, 2009) b2393.

[133] J. H. Flory, J. Roy, J. J. Gagne, K. Haynes, L. Herrinton, C. Lu, E. Patorno, A. Shoaibi, and M. A. Raebel, "Missing laboratory results data in electronic health databases: implications for monitoring diabetes risk," *Journal of Comparative Effectiveness Research* **6** (1, 2017) 25–32.

[134] C. . S. N. L. C. Mosleh, A.; Smidts, "Model uncertainty: Its characterization and quantification," in *Proceedings of workshop I in advanced topics in risk and reliability analysis. Model uncertainty: Its characterization and quantification*, 1994.

[135] L. Esserman, Y. Shieh, and I. Thompson, "Rethinking Screening for Breast Cancer and Prostate Cancer," *JAMA* **302** (10, 2009) 1685.

[136] A. N. Freedman, D. Seminara, M. H. Gail, P. Hartge, G. A. Colditz, R. Ballard-Barbash, and R. M. Pfeiffer, "Cancer risk prediction models: A workshop on development, evaluation, and application," *Journal of the National Cancer Institute* **97** (5, 2005) 715–723.

[137] M. H. Gail, L. A. Brinton, D. P. Byar, D. K. Corle, S. B. Green, C. Schairer, and J. J. Mulvihill, "Projecting individualized probabilities of developing breast cancer for white females who are being examined annually," *Journal of the National Cancer Institute* **81** (12, 1989) 1879–1886.

[138] Y. Li, L. Chen, X. Wan, and A. Chiang, "Implementation of Breast Cancer Risk Assessment Tool using SAS ®," *PharmaSUG 2013* (2013).

[139] J. P. Costantino, M. H. Gail, D. Pee, S. Anderson, C. K. Redmond, J. Benichou, and H. S. Wieand, "Validation Studies for Models Projecting the Risk of Invasive and Total Breast Cancer Incidence," *JNCI Journal of the National Cancer Institute* **91** (1999), no. 18 1541–1548.

[140] A. P. A. A. T. B. A. N. W. H. Nova F. Smedley, N Y Elizabeth Chau, "A Platform for Generating and Validating Breast Risk Models from Clinical Data: Towards Patient-Centered Risk Stratified Screening - Semantic Scholar," in *AMIA*, 2015.

[141] P. Loukopoulos, S. Sampath, P. Pilidis, G. Zolkiewski, I. Bennett, F. Duan, and D. Mba, "Dealing with missing data for prognostic purposes," in *2016 Prognostics and System Health Management Conference (PHM-Chengdu)*, pp. 1–5, IEEE, 10, 2016.

[142] J. Schafer, *Analysis of Incomplete Multivariate Data*, vol. 72 of *C&H/CRC Monographs on Statistics & Applied Probability*. Chapman & Hall, 8, 1997.

[143] S. Zhang, J. Zhang, X. Zhu, Y. Qin, and C. Zhang, "Missing Value Imputation Based on Data Clustering," in *Transactions on Computational Science I*, pp. 128–138. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[144] D. T. Satish Gajawada, "Missing Value Imputation Method Based on Clustering and Nearest Neighbours," *International Journal of Future Computer and Communication* **1** (2012).

[145] T. Liu, A. W. Moore, and A. Gray, "New Algorithms for Efficient High-Dimensional Nonparametric Classification," *Journal of Machine Learning Research* **7** (2006), no. Jun 1135–1158.

[146] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and A. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research* **12** (2011), no. Oct 2825–2830.

[147] J. Tyrer, S. W. Duffy, and J. Cuzick, "A breast cancer prediction model incorporating familial and personal risk factors," *Statistics in Medicine* **23** (4, 2004) 1111–1130.

[148] D. M. Euhus, K. C. Smith, L. Robinson, A. Stucky, O. I. Olopade, S. Cummings, J. E. Garber, A. Chittenden, G. B. Mills, P. Rieger, L. Esserman, B. Crawford, K. S. Hughes, C. A. Roche, P. A. Ganz, J. Seldon, C. J. Fabian, J. Klemp, and G. Tomlinson, "Pretest prediction of BRCA1 or BRCA2 mutation by risk counselors and the computer model BRCAPRO," *Journal of the National Cancer Institute* **94** (6, 2002) 844–851.

[149] W. Jorritsma, F. Cnossen, and P. van Ooijen, "Improving the radiologist–CAD interaction: designing for appropriate trust," *Clinical Radiology* **70** (2, 2015) 115–122.

[150] J. M. McGuirl and N. B. Sarter, "Supporting Trust Calibration and the Effective Use of Decision Aids by Presenting Dynamic System Confidence Information," *Human Factors: The Journal of the Human Factors and Ergonomics Society* **48** (12, 2006) 656–665.

[151] P. K. J. Han, W. M. P. Klein, T. Lehman, B. Killam, H. Massett, and A. N. Freedman, "Communication of Uncertainty Regarding Individualized Cancer Risk Estimates," *Medical Decision Making* **31** (3, 2011) 354–366.

[152] Y.-Y. Liu, S. Li, F. Li, L. Song, and J. M. Rehg, "Efficient Learning of Continuous-Time Hidden Markov Models for Disease Progression.," *Advances in neural information processing systems* **28** (1973), no. 12 3599–3607.