Deep Reinforcement Learning for Early Diagnosis of Lung Cancer

Yifan Wang[1,2], **Qining Zhang**[1], **Lei Ying**[1], **Chuan Zhou**[2]

[1] Department of Electrical Engineering and Computer Science, University of Michigan, USA
[2] Department of Radiology, University of Michigan, USA
{wangyfan,qiningz,leiying}@umich.edu, chuan@med.umich.edu

Abstract

Lung cancer remains the leading cause of cancer-related death worldwide, and early diagnosis of lung cancer is critical for improving the survival rate of patients. Performing annual low-dose computed tomography (LDCT) screening among high-risk populations is the primary approach for early diagnosis. However, after each screening, whether to continue monitoring (with follow-up screenings) or to order a biopsy for diagnosis remains a challenging decision to make. Continuing with follow-up screenings may lead to delayed diagnosis but ordering a biopsy without sufficient evidence incurs unnecessary risk and cost. In this paper, we tackle the problem by an optimal stopping approach. Our proposed algorithm, called EarlyStop-RL, utilizes the structure of the Snell envelope for optimal stopping, and model-free deep reinforcement learning for making diagnosis decisions. Through evaluating our algorithm on a commonly used clinical trial dataset (the National Lung Screening Trial), we demonstrate that EarlyStop-RL has the potential to greatly enhance risk assessment and early diagnosis of lung cancer, surpassing the performance of two widely adopted clinical models, namely the Lung-RADS and the Brock model.

Introduction

Resulting in estimated 130,180 deaths in 2022, lung cancer has become the leading cause of cancer-related deaths in the United States (Siegel et al. 2022). The prognosis of lung cancer patients at different clinical stages is significantly different. The 5-year survival rate of stage IA groups (early stage) can exceed 90%, while the survival rate of patients with stage IV (the latest stage) is less than 10% (Ning et al. 2021). Therefore, early diagnosis holds immense significance for individuals with lung cancer.

The primary approach employed in clinical practices to improve early diagnosis of lung cancer is conducting lung cancer screenings among high-risk populations using low-dose computed tomography (LDCT) (Team 2011b; Ardila, Kiraly et al. 2019). After detecting a lung nodule from an LDCT scan, the primary objective of radiologists is to identify the nodule's risk and then establish a definitive diagnosis or ascertain the necessity for subsequent follow-up LDCT examinations. To date, most lung cancer screening studies and programs worldwide have offered sequential annual screening to participants for up to 4 years (Robbins et al. 2022) according to existing international clinical guidelines (Larici et al. 2017). A significant limitation associated with this management framework is the high incidence of false positives (FP) and an excessive number of follow-up LDCTs (Mehta, Mohammed, and Jantz 2017). The high rate of FP will result in considerable clinical and financial costs associated with over-diagnosis such as unnecessary downstream invasive procedures, while an excessive number of follow-up exams will lead to missed or delayed diagnosis.

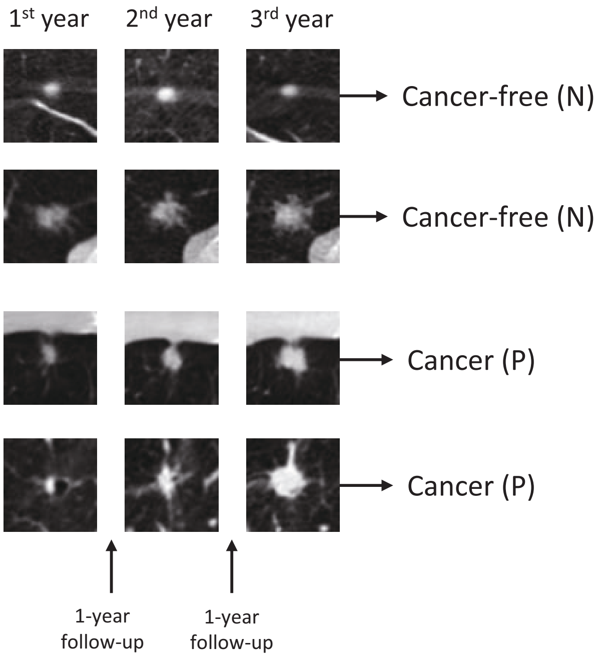

Figure 1: LDCT screening and diagnosis examples from the NLST dataset.

Figure 1 presents examples of four lung nodules on LDCT scans and their diagnosis results from the National Lung Screening Trial (NLST) dataset (https://cdas.cancer.gov/nlst/) (Team 2011a). In the first year, each of the four cases exhibited considerable suspicion of lung cancer. In the subsequent years, distinct radiological attributes became evident. However, by the third screening cycle year, patients afflicted with cancer have already progressed to an advanced pathological stage, and their prospects for survival have substantially decreased. Therefore, it is crucial to diagnose the cancer timely, i.e., in the first or second year before it is too late.

Optimal stopping problem (Poor and Hadjiliadis 2009) roots in the fields of stochastic processes and dynamic programming and gradually becomes a centerpiece for many real-world applications that deal with streaming data. We refer the readers to (Xie et al. 2021) for a recent survey on this topic. In the context of healthcare, this fundamental problem is perfectly akin to the challenges encountered in the field of epidemiology. Various algorithms, such as Cumulative Sum (CUSUM (Ritov 1990)) and Likelihood Ratio Test (LRT (Willsky and Jones 1976)), have been extensively employed for many diseases such as COVID-19 (Braca et al. 2021) or for the early termination of Phase II clinical trials (Nasrollahzadeh and Khademi 2020). The line of research most related to our topic is the Optimal Stopping in Radiation Therapy (OSRT), where several optimal stopping methods have been explored to effectively manage the radiotherapy process for patients with lung cancer (Ajdari et al. 2019).

When facing complex underlying dynamic processes, classical approaches like the Snell envelope approach or Approximate Dynamic Programming (ADP) (Bertsekas and Tsitsiklis 1995) may not be adequate to address optimal stopping problems effectively. On the other hand, Reinforcement Learning (RL), especially model-free reinforcement learning, has been identified as a viable solution to overcome these challenges, particularly through the utilization of deep reinforcement learning algorithms (Ery and Michel 2021; Fathan and Delage 2021).

Our Contributions

In this paper, we focus on the challenge of finding an effective strategy that can provide early diagnoses of lung cancer while maintaining relatively low rates of false positives and false negatives. By considering the biologically stochastic and dynamic nature of lung cancer, we formulate the natural history of lung cancer as a discrete-time Partially Observable Markov Decision Process (POMDP) and view the early diagnosis of lung cancer as an optimal stopping time problem. To solve this problem, we utilize model-free deep reinforcement learning (RL) machinery and the structure of the Snell envelope. We choose RL due to the sequential nature of the data and the unknown progression of lung cancer, which aligns with RL's inherent characteristics. The design of the Snell envelope will yield a stopping rule that is interpretable, leveraging the convexity of the stopping region. This stopping rule can be readily assessed and implemented in clinical settings.

Our work is strongly associated with the Bayesian regime (Shiryaev 1963) and differs fundamentally from the standard optimal stopping problem. Unlike classical stochastic problems that have a directly measurable quantity to indicate the object's condition, such as the number of individuals affected by disease or a machine's work efficiency, the true state of a patient is multifaceted and intricate, which makes it notably difficult to observe directly through several factors and only partial information can be accessed at each stage. Thus, we model the natural history of lung cancer as a POMDP which helps to relate the partial information to the true (unknown) state of the patient. However, it is well-known that solving general POMDPs is extremely difficult. Inspired by the idea of imperfect state information theory (ISI), instead of assuming a deterministic (and unknown) state for the malignancy of a nodule at each stage, the algorithm maintains a probabilistic belief about the state (Wei et al. 2019; Zhang et al. 2022) and solves a Belief MDP problem. Moreover, the nature of our problem is heightened in complexity also due to a more general action space setting where we have multiple stopping actions and may also have multiple continuous actions. At the same time, it is imperative for the policy to possess a level of interpretability that aligns with the distinct stipulations pertinent to its clinical application.

Following the aforementioned novel formulation of the problem concerning the early diagnosis of lung cancer, we develop an algorithm called EarlyStop-RL, which includes a belief update phase and an optimal stopping phase:

Belief Update: After each medical exam, our EarlyStop-RL algorithm will update its assessment of lung cancer beliefs based on historic observations, using the Hidden Markov Model filter method that adheres to a probabilistic model we adapt from a clinical lung nodule dynamic model (Sarapata and De Pillis 2014; Vaghi et al. 2020).

Optimal Stopping: Utilizing the up-to-date beliefs and observations, the agent will either provide a definite diagnosis indicating whether the patient is positive/negative for lung cancer or schedule another follow-up exam to collect more evidence about the patient. As more evidence is gathered, the rates of false negatives and false positives decrease, but we are at risk of potential late diagnosis. In order to achieve a balance between these two types of risks, we propose a cost model that takes into account the clinical significance of both misdiagnosis and delayed diagnosis costs. With this formulation, the goal of our optimal stopping model is to discover an optimal stopping policy that will minimize the overall cost. Therefore, the ideal policy will give a diagnosis at the earliest possible time and at the same time also maintain relatively low rates of false negatives and false positives.

In summary, the main contributions of this paper include:

- **Problem Formulation:** based on the dynamic nature of lung cancer history and the screening process, considering the fundamental trade-off between immediate diagnosis and more evidence, we formulate the natural history of lung cancer as a POMDP and view the early diagnosis of lung cancer as an optimal stopping problem.
- **Deep Reinforcement Learning for Optimal Stopping:** based on the problem formulation, we make a connection between the optimal stopping problem and the model-free reinforcement learning framework, leveraging the power of representation and learning from deep rein-

forcement learning algorithms and utilizing the structure of optimal stopping framework at the same time.

- **Structural Results with Theory Analysis:** following the well-crafted framework and conducting theoretical evaluations, we establish structural results related to the convexity of value functions and the stopping region, resulting in an interpretable stopping and diagnosis policy that outperforms current clinic models, such as Lung-RADS and Brock, by a considerable margin.

The codes and appendixes of this project are published at https://github.com/Yifan-wang-maybe/EarlyStop-RL

Related Works

In order to assist radiologists in making a diagnosis decision during the lung cancer screening process, several clinical guidelines and predictive models have been proposed to estimate the probability of malignancy of nodules and guide management. The Lung CT Screening Reporting and Data System (Lung-RADS) (McKee et al. 2016; Ardila, Kiraly et al. 2019) and the Brock model (McWilliams et al. 2013) are two commonly used clinical models. Lung-RADS operates as a category-based model, while the Brock model functions as a linear regression model with radiological risk factors, such as nodule diameter and attenuation, serving as inputs. These models are widely recognized and used as baselines for developing new models.

Nowadays, the implementation of various Artificial Intelligence (AI) models has yielded significant advancements in the field of computer-aided diagnosis. Notably, deep convolutional neural networks (DCNNs) (see (Mridha et al. 2022) for a comprehensive review) i.e. a study from Google AI (Ardila, Kiraly et al. 2019) demonstrated comparable performance to real clinical practices for the diagnosis of lung cancer. However, most of these methods focus on analyzing image features on an individual CT exam under a supervised approach, which can not accurately reflect the overall risk of lung cancer. While some single-time characteristics such as size or density generally correlate with the probability of malignancy, a definitive assessment of a nodule's biological behavior is unknown clinically until the nodule demonstrates more suspicious features such as growth or stability. Therefore, using a temporal analysis based on the biological progression of cancer over time to estimate lung cancer risk is crucial. There exist some studies using value-iteration reinforcement learning algorithm (Wang et al. 2021) or deep Q-learning algorithm (Liu et al. 2019) to explore serial exams during lung cancer screening. To our knowledge, no existing AI model considers the optimal balance between early diagnosis and follow-up exams for risk management in lung cancer screening, which is our major contribution.

Problem Formulation

This section introduces the formal problem statement and its corresponding mathematical models. As stated in the introduction section, we model the dynamic nature of lung cancer history and the screening process as a partially observable Markov decision process (POMDP). Formally, a POMDP is represented as a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{C}, \Omega, \mathcal{O}, \gamma)$, where $\mathcal{S}$, $\mathcal{A}$, and Ω are the state, action, and observation sets, respectively. $\mathcal{T}$ is the stochastic state transition model and $\mathcal{O}$ is the probabilistic observation model. $\mathcal{C}$ is a bounded cost function and $0 \leq \gamma \leq 1$ is a discount factor.

Notations: We use $\mathbb{1}_{\{\cdot\}}$ to denote the indicator function, Pr to represent the probability measure, and $\mathbb{E}_X$ to denote expectation with respect to a random variable X. Subscript t such as a_t typically represents the time index unless otherwise specified.

State, Action, and State Transition Model

In our early diagnosis of the lung cancer problem, we consider a state space consisting of three distinct states based on the true cancer state of each patient which is not observable: Negative for lung cancer (N), Positive for lung cancer (P), and Evolving (Ev). Therefore, we define state space as:

$$\mathcal{S} = \Big\{\mathrm{Negative(N)}, \mathrm{Evolving(Ev)}, \mathrm{Positive(P)}\Big\}, \quad (1)$$

and we use s_t to represent the state of a patient at time t.

We assume that patients begin in the evolving state (state Ev). Once a patient's state changes to the state positive or negative (N or P) for lung cancer, the cancerous state persists until additional treatment is administered or new nodules are discovered. Therefore we define them as the absorbing state a patient finally reaches:

$$\mathcal{S}_{\mathrm{F}} \in \Big\{\mathrm{Negative(N)}, \mathrm{Positive(P)}\Big\} \subset \mathcal{S} \quad (2)$$

We simplify the set of possible actions into two types, the first type $\mathcal{A}_{\mathrm{C}}$ is requiring more follow-up exams, and another type $\mathcal{A}_{\mathrm{D}}$ is terminal and confirms cancer-positive (P) or cancer-negative (N) for patients. Since these actions are diagnosis actions, not the treatments given to the patient, they will only affect the observation distribution and not the evolution of the lung nodules.

The discrete-time state transition model $\mathcal{T}$ is shown in Figure 2 with λ_1, λ_2, and λ_3 as transition probability:

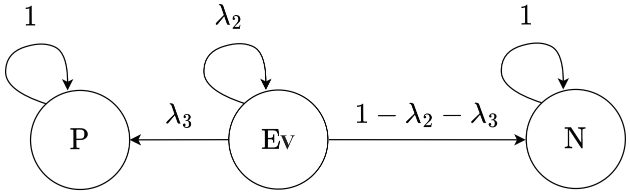

Figure 2: The state transition model.

Observation Model

Whenever an action from the follow-up category $\mathcal{A}_{\mathrm{C}}$ is selected, a corresponding observation will be observed. The observation is continuous, which depends on the previous time-points observation and the real state. Let $z_t \in \Omega$ represent the observation (medical exam's result) at time point t, and define the probabilistic observation model $\mathcal{O}$ through the definition of observation probabilities O:

$$O(z_{t+1}|z_t, a_t, s_{t+1}) = \Pr(z_{t+1}|z_t, a_t, s_{t+1}). \quad (3)$$

Based on the formulation of the state space, to simplify, we define:

$$\begin{aligned}
&\alpha_{a_t}(z_{t+1}|z_t) = O(z_{t+1}|z_t, a_t, s_{t+1} = \mathrm{Ev}),\\
&\beta_{a_t}(z_{t+1}|z_t) = O(z_{t+1}|z_t, a_t, s_{t+1} = \mathrm{N}),\\
&\gamma_{a_t}(z_{t+1}|z_t) = O(z_{t+1}|z_t, a_t, s_{t+1} = \mathrm{P}).
\end{aligned} \tag{4}$$

In the early diagnosis of lung cancer problem we consider, the observation refers to the nodule size since it remains the most widely used predictor to assess the probability of nodule malignancy and to determine nodule management according to the international guidelines (Larici et al. 2017). The comprehensive expression of the above transition probability relies on a clinical lung nodule dynamic model (Sarapata and De Pillis 2014; Vaghi et al. 2020), which will be thoroughly explicated in the appendix.

Stopping Time

Let $T \geq 1$ denote a stopping time at which a diagnosis action from $\mathcal{A}_{\mathrm{D}}$ is selected and the screening process ends.

Cost Function

Let $\delta_T \in \mathcal{A}_{\mathrm{D}}$ denote the diagnosis action given at the stopping time T, and $a_t \in \mathcal{A}_{\mathrm{C}}$ $(t < T)$ denote the action chosen before stopping time T. The overall cost will be the summation of the following two types of costs:

Misdiagnosis Cost: with $c_{\mathrm{d1}}, \cdots, c_{\mathrm{d4}}$ being the cost value weight, the misdiagnosis cost at stopping time T is given by:

$$\begin{aligned}
C_{\mathrm{D}}(\delta_T) = c_{\mathrm{d1}}\mathbb{1}_{\{\delta_T=\mathrm{N}, s_T=\mathrm{P}\}} + c_{\mathrm{d2}}\mathbb{1}_{\{\delta_T=\mathrm{N}, s_T=\mathrm{Ev}\}}\\
+ c_{\mathrm{d3}}\mathbb{1}_{\{\delta_T=\mathrm{P}, s_T=\mathrm{N}\}} + c_{\mathrm{d4}}\mathbb{1}_{\{\delta_T=\mathrm{P}, s_T=\mathrm{Ev}\}}.
\end{aligned} \tag{5}$$

This cost will penalize the occurrence of false-positive and false-negative outcomes in the diagnosis process.

Delay-diagnosis Cost (Inter-step Cost): let $\lambda(z, k)$ be the hazard function in clinical with z as the current observation, and k as an index (how many follow-ups have already been processed). At time t, we have:

$$C_{\mathrm{I}}(z_t, a_t, k_t) = c_{a_t} + c_{\mathrm{m}}\lambda(z_t, k_t)\mathbb{1}_{\{s_{\mathrm{F}}=\mathrm{P}\}}, \tag{6}$$

where c_{a_t} is a cost for the inter-step actions, c_{m} is a weighting parameter and $\lambda(z_t, k_t)\mathbb{1}_{\{s_{\mathrm{F}}=\mathrm{P}\}}$ is a penalty for late diagnosis of lung cancer.

Methodology

In the problem formulation section, we model the early diagnosis of lung cancer problem as a POMDP with the goal to minimize the overall cost. In this section, we will begin by introducing the belief update component and the theorem for converting the POMDP into a fully observable belief MDP, as we previously discussed in the introduction section. Following this, we will discuss our optimal stopping approach for the early diagnosis of lung cancer.

Belief Update and Fully Observable Belief MDP

We establish the notions of the belief that the state is positive π_t^{P}, negative π_t^{N} and evolving π_t^{Ev} at time t as follows:

$$\begin{aligned}
\pi_t^{\mathrm{P}} &:= \Pr(s_t = \mathrm{P} \mid \mathcal{F}_t),\\
\pi_t^{\mathrm{N}} &:= \Pr(s_t = \mathrm{N} \mid \mathcal{F}_t),\\
\pi_t^{\mathrm{Ev}} &:= \Pr(s_t = \mathrm{Ev} \mid \mathcal{F}_t),
\end{aligned} \tag{7}$$

where $\mathcal{F}_t$ is the σ-algebra that contains all past observations until time t and actions until time $t-1$: $\mathcal{F}_t = \{z_1, k_1, a_1, z_2, \cdots, z_{t-1}, k_{t-1}, a_{t-1}, z_t, k_t\}$.

Due to limited space, we only illustrate the posterior update formulation related to π_t^{P} as an example. The update rules for π_t^{N} and $\pi_t^{\mathrm{Ev}} = 1 - \pi_t^{\mathrm{P}} - \pi_t^{\mathrm{N}}$ are similar and will be detailed described in the appendix. At the beginning of time step $t+1$ and before receiving observation z_{t+1}, the posterior changes to $\hat{\pi}_{t+1}^{\mathrm{P}}$ based on the transition matrix:

$$\begin{aligned}
\hat{\pi}_{t+1}^{\mathrm{P}} &:= \Pr(s_{t+1} = \mathrm{P} \mid \mathcal{F}_t)\\
&= \sum_{s\in\{\mathrm{P},\mathrm{N},\mathrm{Ev}\}} \Pr(s_{t+1} = \mathrm{P}|s_t = s, \mathcal{F}_t)\Pr(s_t = s|\mathcal{F}_t)\\
&= \sum_{s\in\{\mathrm{P},\mathrm{N},\mathrm{Ev}\}} \Pr(s_{t+1} = \mathrm{P}|s_t = s)\Pr(s_t = s|\mathcal{F}_t)\\
&= \pi_t^{\mathrm{P}} + \lambda_3\pi_t^{\mathrm{Ev}}.
\end{aligned} \tag{8}$$

After adding the information of observation z_{t+1}, the updated belief is:

$$\begin{aligned}
&\pi_{t+1,a_t}^{\mathrm{P}} = \Pr(s_{t+1} = \mathrm{P} \mid \mathcal{F}_{t+1}, a_t)\\
&= \frac{\Pr(s_{t+1} = \mathrm{P} \mid \mathcal{F}_t)\Pr(z_{t+1} \mid s_{t+1} = \mathrm{P}, \mathcal{F}_t, a_t)}{\Pr(z_{t+1} \mid \mathcal{F}_t, a_t)}\\
&= \frac{\hat{\pi}_{t+1}^{\mathrm{P}}\gamma_{a_t}(z_{t+1} \mid z_t)}{\Pr(z_{t+1} \mid \mathcal{F}_t, a_t)}\\
&= \frac{\hat{\pi}_{t+1}^{\mathrm{P}}\gamma_{a_t}(z_{t+1}|z_t)}{\hat{\pi}_{t+1}^{\mathrm{P}}\gamma_{a_t}(z_{t+1}|z_t) + \hat{\pi}_{t+1}^{\mathrm{N}}\beta_{a_t}(z_{t+1}|z_t) + \hat{\pi}_{t+1}^{\mathrm{Ev}}\alpha_{a_t}(z_{t+1}|z_t)}\\
&:= B_{a_t}^{\mathrm{P}}(\pi_t^{\mathrm{P}}, \pi_t^{\mathrm{N}}, z_{t+1}, z_t).
\end{aligned} \tag{9}$$

Eq. (9) and the updated belief for $\pi_{t+1,a_t}^{\mathrm{Ev}}$ and $\pi_{t+1,a_t}^{\mathrm{N}}$ will be proved in the appendix based on the Hidden Markov Model filter (Krishnamurthy 2016). With the help of the update equations and replacing the unobserved patient real state s with the belief of the state, we have the following theorem:

Theorem 1 *The early diagnosis of lung cancer problem defined based on POMDP is equivalent to solving the problem on a fully observed MDP with state $\theta_t = (\pi_t^{\mathrm{N}}, \pi_t^{\mathrm{P}}, z_t, k_t)$ where $\pi_t^{\mathrm{N}}, \pi_t^{\mathrm{P}}$ is the belief to replace the unknown state s_t, z_t and k_t are current observation and index related to real-time step. The action space remains the same and the state transition probability is based on the belief update formulations. The new cost function is as follows:*

Misdiagnosis Cost:

$$\begin{aligned}
C_{\mathrm{D}}(\theta_T, \delta_T) = \big((c_{\mathrm{d1}} - c_{\mathrm{d2}})\pi_T^{\mathrm{P}} + c_{\mathrm{d2}}(1 - \pi_T^{\mathrm{N}})\big)\mathbb{1}_{\{\delta_T=\mathrm{N}\}}\\
+ \big((c_{\mathrm{d3}} - c_{\mathrm{d4}})\pi_T^{\mathrm{N}} + c_{\mathrm{d4}}(1 - \pi_T^{\mathrm{P}})\big)\mathbb{1}_{\{\delta_T=\mathrm{P}\}}.
\end{aligned} \tag{10}$$

Delay-diagnosis Cost:

$$C_{\mathrm{I}}(\theta_t, a_t) = c_{a_t} + c_{\mathrm{m}}\lambda(z_t, k_t)\pi_t^{\mathrm{P}}. \tag{11}$$

The proof of Theorem 1 is provided in the appendix. Based on this theorem, our original early diagnosis of lung cancer problem is converted to an MDP with the belief state $\pi_t^{\mathrm{N}}, \pi_t^{\mathrm{P}}$ and the z_t, k_t since they appear in the cost function.

Markov Optimal Stopping Formulation

In this section, based on the Theorem 1 above, our POMDP model is converted to an MDP model with state $\theta_t = (\pi_t^{\mathrm{N}}, \pi_t^{\mathrm{P}}, z_t, k_t)$. Thus, the issue of diagnosing lung cancer at an early stage can be reframed as a Markov optimal stopping problem, as shown below:

Theorem 2 *The early diagnosis of lung cancer problem can be reformed as a Markov optimal stopping problem with respect to the belief MDP defined in Theorem 1. Formally,*

$$\begin{aligned}
&\inf_{T,\delta_T\in\mathcal{A}_{\mathrm{D}},a_t\in\mathcal{A}_{\mathrm{C}}} \mathbb{E}\left[C_{\mathrm{D}}(\theta_T,\delta_T)+\sum_{t=1}^{T-1}C_{\mathrm{I}}(\theta_t,a_t)\right]\\
&=\inf_{T,a_t\in\mathcal{A}_{\mathrm{C}}}\mathbb{E}\Bigg[\inf\Big\{(c_{\mathrm{d1}}-c_{\mathrm{d2}})\pi_T^{\mathrm{P}}+c_{\mathrm{d2}}\big(1-\pi_T^{\mathrm{N}}\big),\\
&\quad(c_{\mathrm{d3}}-c_{\mathrm{d4}})\pi_T^{\mathrm{N}}+c_{\mathrm{d4}}\big(1-\pi_T^{\mathrm{P}}\big)\Big\}+\sum_{t=1}^{T-1}C_{\mathrm{I}}(\theta_t,a_t)\Bigg]\\
&:=\inf_{T,a_t\in\mathcal{A}_{\mathrm{C}}}\mathbb{E}\left[g(\pi_T^{\mathrm{P}},\pi_T^{\mathrm{N}})+\sum_{t=1}^{T-1}C_{\mathrm{I}}(\theta_t,a_t)\right].
\end{aligned} \tag{12}$$

The proof of Theorem 2 is provided in the appendix. The analytical solution of the Markov optimal stopping problem can be obtained using the Snell envelope process M_n:

$$M_n=\inf_{T\geq n,a_t\in\mathcal{A}_{\mathrm{C}}}\mathbb{E}\left[g\Big(\pi_T^{\mathrm{P}},\pi_T^{\mathrm{N}}\Big)+\sum_{t=1}^{T-1}\Big(C_{\mathrm{I}}(\theta_t,a_t)\Big)\right], \tag{13}$$

which has the same formulation as what we get from Theorem 2 except the condition of $T \geq n$. Below are several important propositions that aid in solving the Markov optimal stopping problem:

Proposition 1 *The associated Snell envelop process M_n satisfies the backward recursion $M_n = \inf\big\{g\big(\pi_T^{\mathrm{P}},\pi_T^{\mathrm{N}}\big), C_{\mathrm{I}}(\theta_n,a_n)+\mathbb{E}[M_{n+1}\mid\mathcal{F}_n]\big\}$.*

Proposition 2 *The optimal stopping time $T^* = \inf\big\{n : M_n \geq g\big(\pi_T^{\mathrm{P}},\pi_T^{\mathrm{N}}\big)\big\}$ is the optimal solution for the Markov stopping problem in Theorem 2.*

The above propositions are well-known results and a standard proof can be found in (Karatzas et al. 1991).

Reinforcement Learning for Optimal Stopping

Various modeling approaches have been proposed to estimate the optimal value of the objective function in Theorem 2 (Ery and Michel 2021). In our study, based on proposition 1, we propose to define the optimal state-action function Q^* as:

$$Q^*(\theta_t,a_t)=\begin{cases}g\big(\pi_t^{\mathrm{P}},\pi_t^{\mathrm{N}}\big) & \text{if } a_t\in\mathcal{A}_{\mathrm{D}}\\ C_{\mathrm{I}}(\theta_t,a_t)+\mathbb{E}[M_{t+1}\mid\mathcal{F}_t] & \text{if } a_t\in\mathcal{A}_{\mathrm{C}}\end{cases} \tag{14}$$

where M_{t+1} is defined in Eq. (13) and

$$a_t=\arg\min_{a_t}Q^*(\theta_t,a_t). \tag{15}$$

According to Proposition 2, we can recover the result in the Markov optimal stopping problem through the acknowledgment of the optimal state-action function (14). This modeling approach is very similar to the Q-learning machinery. Formulated in reinforcement learning notation, the state-action function Q can be rewritten as:

$$Q(\theta_t,a_t)=\begin{cases}g\big(\pi_t^{\mathrm{P}},\pi_t^{\mathrm{N}}\big) & \text{if } a_t\in\mathcal{A}_{\mathrm{D}}\\ \mathcal{BT}(Q(\theta_t,a_t)) & \text{if } a_t\in\mathcal{A}_{\mathrm{C}}\end{cases} \tag{16}$$

where $\mathcal{BT}$ is the Bellman operator:

$$\mathcal{BT}(Q(\theta_t,a_t))=C_{\mathrm{I}}(\theta_t,a_t)+\sum_{z_{t+1}}\Pr\big(z_{t+1}|\pi_t^{\mathrm{P}},\pi_t^{\mathrm{N}},a_t,z_t\big)V^{'}, \tag{17}$$

where

$$\begin{aligned}V'=V\Bigg(&B_{a_t}^{\mathrm{P}}\big(\pi_t^{\mathrm{P}},\pi_t^{\mathrm{N}},z_{t+1},z_t\big),\\ &B_{a_t}^{\mathrm{N}}\big(\pi_t^{\mathrm{P}},\pi_t^{\mathrm{N}},z_{t+1},z_t\big),z_{t+1},k_{t+1}\Bigg).\end{aligned}$$

The value function $V(\theta_t)$ above is given by:

$$V(\theta_t)=\min_{a_t}Q(\theta_t,a_t). \tag{18}$$

Define the belief space $\Pi(X)$:

$$\Pi(X)=\Big\{\pi\in\mathcal{R}^3:\pi^{\mathrm{P}}+\pi^{\mathrm{Ev}}+\pi^{\mathrm{N}}=1,0\leq\pi^{\{\mathrm{P,Ev,N}\}}\leq1\Big\} \tag{19}$$

as a 2-dimensional unit simplex. Define set $\mathcal{R}_{\mathrm{P}}$ as the set of belief states for which the diagnosis action $\delta_T = \mathrm{P}$ is optimal. Similarly, define $\mathcal{R}_{\mathrm{N}}$ as the set of belief states that $\delta_T = \mathrm{N}$ is optimal and $\mathcal{R}_{\mathrm{C}}$ for belief states with continuous action (more follow-up exams) as optimal. The following is our key theorem which gives the structural result for the optimal policy and enables us to obtain a comprehensible stopping criterion.

Theorem 3 *The value function $V(\theta)$ and the state-action function $Q(\theta, a)$ are concave functions with respect to $\pi \in \Pi(X)$ (recall that $\theta = (\pi^{\mathrm{N}}, \pi^{\mathrm{P}}, z, k)$) and the stopping regions $\mathcal{R}_{\mathrm{N}}$ and $\mathcal{R}_{\mathrm{P}}$ are convex and individually connected in the belief space.*

The proof of Theorem 3 is provided in the appendix.

Summary: Starting from framing the early diagnosis of lung cancer problem as a POMDP in the problem formulation section, we converted the POMDP into a fully observable belief MDP following the principles outlined in Theorem 1. Then, we further reframed it as a Markov optimal stopping problem that can be addressed by employing the Snell envelope, as outlined in Theorem 2 and Eq. (13). In the next step, we made a connection between the Snell envelope and the reinforcement learning framework. By solving Eq. (16), in turn, we could derive the optimal stopping problem associated with the Snell envelope and facilitated the determination of the optimal policy for managing lung cancer screening. Finally in Theorem 3, we demonstrated that the stopping regions are both convex and individually connected within the belief space. This indicates the interpretability of the optimal policy which will be elucidated and used in the following EarlyStop-RL algorithm design.

EarlyStop-RL

Based on the formulations elucidated in the preceding sections, we introduce our EarlyStop-RL algorithm, aimed at facilitating the early diagnosis of lung cancer.

A simplified pseudo-code of our EarlyStop-RL can be found in Algorithm 1. During the implementation, we utilize a three-layer "linear-batch normalization-ReLu" neural network structure to represent the Q-function, which is trained using the value iteration method based on the update equation (16) under the model-free reinforcement learning framework. Detailed information such as computing infrastructure and training process will be explained in the appendix. Once the optimal state-action function $Q(\theta, a)$ is obtained, the determination of the stopping time and diagnosis result becomes feasible by relying on the proposition 2.

Recall that our belief space $\Pi(X)$ is a 2-dimensional unit simplex (triangle) and the three vertices are distinctly associated with three different regions ($\mathcal{R}_{\mathrm{N}}$, $\mathcal{R}_{\mathrm{C}}$, and $\mathcal{R}_{\mathrm{P}}$). Given the individually connected and convex property of the stopping regions $\mathcal{R}_{\mathrm{N}}$ and $\mathcal{R}_{\mathrm{P}}$, it can be readily shown that two threshold stitching curves exist, effectively partitioning the belief space into three distinct regions: $\mathcal{R}_{\mathrm{N}}$, $\mathcal{R}_{\mathrm{C}}$, and $\mathcal{R}_{\mathrm{P}}$. (This is an extension of (Krishnamurthy 2016, Theorem 12.3.4 2(a))). In general, any user-defined basis function approximation can be used to parameterize this stitching cure (Krishnamurthy 2016). In light of the particular contextual inherent in our problem, including the linear nature of stopping cost (Misdiagnosis cost) and the outcomes obtained through simulation, we proceed to establish the linear function approximation for the switching curves within the belief space $\Pi(X)$. In our experiments, we employ a support vector machine with a linear kernel to establish the threshold boundary to divide the belief space into three regions ($\mathcal{R}_{\mathrm{N}}$, $\mathcal{R}_{\mathrm{C}}$, and $\mathcal{R}_{\mathrm{P}}$). Please refer to the appendix where you can find additional information. In this manner, despite the potential occurrence of performance degradation and computational complexity during the quantization and fitting process, it obviates the necessity of recurrently executing the neural network at each stage during the implementation phase. This attribute renders it more conducive for employment in clinical settings.

Experiments and Results

National Lung Screening Trial

The performance assessment of EarlyStop-RL is conducted through its implementation on the National Lung Screening Trial (NLST) (https://cdas.cancer.gov/nlst/) (Team 2011a), a multicenter trial involving 33 centers across the United States. NLST has the largest available dataset to date in terms of the number of patients and is the most representative dataset in terms of its diversity and the potential for practice over a wide range of populations. More importantly, it's a crucial clinical trial widely recognized for its significance in facilitating the reduction of lung cancer mortality. The NLST dataset is publicly accessible, but obtaining permission from the NLST research team is required.

Algorithm 1: EarlyStop-RL (offline training)

```
Input: A dataset including M patients' lung cancer screening traces:
    $\hat{D} = \{s_1^i, z_1^i, a_1^i, s_2^i, z_2^i, a_2^i, \cdots\}_{i=1}^{M}$.
Belief Update: Calculate belief traces for each patient based on update equation (9):
    $\pi_{t+1}^{i,\mathrm{P}} = B_{a_t^i}^{\mathrm{P}}(\pi_t^{i,\mathrm{P}}, \pi_t^{i,\mathrm{N}}, z_{t+1}, z_t)$,
    $\pi_{t+1}^{i,\mathrm{N}} = B_{a_t^i}^{\mathrm{N}}(\pi_t^{i,\mathrm{P}}, \pi_t^{i,\mathrm{N}}, z_{t+1}, z_t)$,
  and reformulate the dataset as:
    $D = \left\{(\pi_1^{i,\mathrm{P}}, \pi_1^{i,\mathrm{N}}, z_1^i, k_1^i), a_1^i, \cdots, \delta_T\right\}_{i=1}^{M}$.
Initialize state-action function $Q(\theta, a)$, parameters $c_{\mathrm{d}1}, \cdots, c_{\mathrm{d}4}, c_a, c_{\mathrm{m}}$ in the cost function, and convergence tolerance $\varepsilon$.
for Epoch $j = 1, 2, \cdots$ do
    Sample bathes $\{\theta_t, a_t, \theta_{t+1}\}$ from the dataset $D$.
    Update $Q(\theta, a)^j$ based on Eq. (16):
    $Q(\theta_t, a_t)^j = \begin{cases} g(\pi_t^{\mathrm{N}}, \pi_t^{\mathrm{P}}) & \text{if } a_t \in \mathcal{A}_{\mathrm{D}} \\ \mathcal{BT}(Q(\theta_t, a_t)^{j-1}) & \text{if } a_t \in \mathcal{A}_{\mathrm{C}} \end{cases}$
    if $||Q(\theta_t, a_t)^j - Q(\theta_t, a_t)^{j-1}|| < \varepsilon$ then
        Break
    end if
end for
Fitting the threshold boundary between continue and stop regions based on proposition 2:
    $Q(\theta_t, a_t) = g(\pi_t^{\mathrm{N}}, \pi_t^{\mathrm{P}})$
  and calculate the region $\mathcal{R}_{\mathrm{N}}, \mathcal{R}_{\mathrm{P}}, \mathcal{R}_{\mathrm{C}}$.
Output: Parameters for the threshold of regions $\mathcal{R}_{\mathrm{N}}$, $\mathcal{R}_{\mathrm{P}}$, $\mathcal{R}_{\mathrm{C}}$, and $Q(\theta_t, a_t)$ if needed.
```

Cohort

With access permission, we collected low-dose CT (LDCT) scans from 2500 patients who underwent annual screening for up to 3 years. Among the 2500 patients, 1951 patients with at least one 4 to 30 mm non-calcified nodule found on their baseline year are included in this study. In NLST, patients with positive lung cancer diagnosis were all confirmed through biopsy, and all patients negative for lung cancer were confirmed based on three years of screening LDCTs and/or up to seven years of subsequent non-CT follow-up. Following a random partitioning procedure, the training/validation set includes 1404 patients, among which 372 were diagnosed with lung cancer, while the remaining 1032 were cancer-free. The ratio of positive to negative in the test set is comparable, comprising 150 patients with positive results and 397 patients with negative results for lung cancer.

Baseline Algorithms

To make a comparison between EarlyStop-RL and widely used clinical models, we first implemented two clinical models for the diagnosis of lung cancer as baseline algorithms:

- Lung CT Screening Reporting and Data System (Lung-RADS) is a widely employed clinical model that serves as the foundation for developing related models and policies. (McKee et al. 2016; Ardila, Kiraly et al. 2019).

- The Brock model (McWilliams et al. 2013) holds a high level of esteem and has been endorsed by respected organizations such as the British Thoracic Society. Furthermore, it has been tested in many external validation studies.

The above models can classify patients into low-risk, medium-risk, or high-risk categories for lung cancer (Huang et al. 2019; Ardila, Kiraly et al. 2019). These outputs correspond to negative, requiring further follow-up, and positive for lung cancer in our EarlyStop-RL, respectively. Additional explanations of these models can be found in the appendix.

There exist several other AI models have been created to comprehend the lung cancer screening (LCS) process and diagnosis. In this study, we compare the reported results from the Google AI model (Ardila, Kiraly et al. 2019), given its status as one of the highly cited and influential studies in recent years. Google AI model is an end-to-end convolutional neural network, with the nodules' region of interest (ROI) as input and the risk of malignancy as the output. The risk is then categorized into low-risk, medium-risk, or high-risk categories based on the Lung-RADS. It is important to note that while this approach also utilizes the NLST dataset, there are variations in the number of patients and the criteria for patient inclusion compared to our study. For example, it includes much more patients with cancer-free diagnoses in the testing set (the ratio of positive to negative is 0.01, and ours is 0.38).

Evaluation Metrics

Initially, we evaluate our EarlyStop-RL algorithm and baseline algorithms using standard evaluation metrics including false-positive rate, false-negative rate, F1 score, and Matthews Correction Coefficient (MCC), as shown in Table **??**. Furthermore, we incorporate two clinically relevant metrics: the Net Reclassification Index and the early diagnosis rate, as below:

Early Diagnosis Rate We define the percentage of patients who would have earlier correct diagnoses (fewer follow-up LDCT examinations) than the NLST clinical trial if the EarlyStop-RL algorithm is incorporated into the medical management process as the early diagnosis rate. The clinical significance of early diagnosis rates is multifaceted, impacting individual patient outcomes, the overall health of populations, and the efficiency of healthcare systems. In this context, achieving a higher rate of early diagnosis signifies an improved level of performance.

Net Reclassification Index We conducte the reclassification analysis to assess the impact of our EarlyStop-RL in improving early diagnosis of lung cancer. The Net Reclassification Index (NRI) (Leening et al. 2014) is a popular metric in clinical practice that attempts to quantify how well a new model reclassifies subjects, either appropriately or inappropriately, as compared to an old model and the ability to lead to better-informed clinical decisions. The NRI is defined as follows:

$$\begin{aligned}\mathrm{NB_P} &= \Pr(\text{up}|\text{event}) - \Pr(\text{down}|\text{event}),\\ \mathrm{NB_N} &= \Pr(\text{down}|\text{non-event}) - \Pr(\text{up}|\text{non-event}),\\ \text{NRI} &= \mathrm{NB_P} + \mathrm{NB_N}.\end{aligned} \tag{20}$$

In this context, the term "event" refers to the true positive cancer state of a patient, and "up" indicates that our EarlyStop-RL has assigned the patient to a higher risk category (such as changing from negative for lung cancer to requiring follow-up or from requiring follow-up to positive for lung cancer) than the compared model. Conversely, "non-event" and "down" refer to negative cancer state and lower risk categories, respectively. An ideal model would assign a higher risk category for events and a lower risk category for non-events. Therefore, a larger NRI indicates improved performance. The statistical significance is determined using the Z-statistic following McNamara's test (McNemar 1947) with a threshold of $p < 0.05$.

Results and Discussion

Early Diagnosis Rate Within our test cohort comprising 547 patients, 450 individuals have already undergone follow-up examinations as part of the clinical trial, with an early diagnosis rate of only 17.73%, which means more than 80% of patients in clinical will take the risk of delay-diagnosis. This outcome underscores the significance of developing and employing more robust models to enhance the early detection of lung cancer.

Table 1 shows the outcomes of early diagnosis rates for each model, along with their respective false-positive and false-negative rates, F1 score, and MCC. Notably, our EarlyStop-RL algorithm achieves a significantly higher early diagnosis rate (60.88%) compared to clinical models (Lung-RADS and Brock model), while simultaneously posing a lower risk of false positives and false negatives. Despite the Brock model achieving a comparable false-positive and false-negative rate, it tends to classify more nodules as indeterminate, recommending additional follow-ups and resulting in a low early diagnosis rate. Maybe this tendency is attributed to the Brock model's lack of consideration for the delay-diagnosis risk during its development. As for the Google AI model, incorporating a testing set with 100 times more negative cases than positive ones yields an imbalanced dataset. This is reflected in a low false-positive rate but a higher false-negative rate, along with very low MCC and F1 scores when compared with our EarlyStop-RL, as the last two metrics also account for robust and balanced performance across positive and negative subjects.

Net Reclassification Index Table 2 shows the Net Benefit for lung cancer patients ($\mathrm{NB_P}$) and for cancer-free patients ($\mathrm{NB_N}$) separately and also the overall NRI. A higher Net Benefit could translate to a net benefit of fewer unnecessary follow-up procedures and fewer missed cancers in clinical practice. The overall NRI is 0.24 when compared to Lung-RADS, with a Z-statistic of 4.20 and a p-value of 0.00. Similarly, when compared to the Brock model, the overall NRI is

	False-positive rate ↓	False-negative rate ↓	Early diagnosis rate ↑	F1 score ↑	MCC ↑
Lung-RADS	29.47%	12.00%	35.47%	0.66	0.52
Brock Model	13.10%	4.67%	33.64%	0.83	0.77
Google AI Model	9.00%	4.40%	72.5%	0.22	0.33
EarlyStop-RL	12.85%	1.33%	60.88%	0.85	0.80

Table 1: Comparison of our EarlyStop-RL with clinical methods regarding rates of false positives, false negatives, F1 Score, Matthews Correction Coefficient (MCC), and early diagnosis rate. ↑ represents a higher value signifies an improved performance and ↓ is the opposite.

		EarlyStop-RL		$\mathrm{NB_P}\big/\mathrm{NB_N}$
		UP/patient	Down/patient	
Lung cancer 150 patients	Lung-RADS	31	15	0.11
	Brock Model	29	15	0.10
Cancer-free 397 patients	Lung-RADS	78	129	0.13
	Brock Model	81	78	0.00
Overall NRI compare with Lung-RADs: 0.24, $p = 0.00$; with Brock Model: 0.10, $p = 0.03$				

Table 2: Reclassification of our EarlyStop-RL when compared with the clinical Lung-RADS and Brock model for the early diagnosis of lung nodule.

0.10, with a Z-statistic of 1.88 and a p-value of 0.03. These results present a statistically significant performance gap between our EarlyStop-RL algorithm and other models, which demonstrates the superiority of our EarlyStop-RL.

Conclusion

In this paper, we proposed an algorithm called EarlyStop-RL, which effectively improved early diagnosis of lung cancer. By formulating the natural history of lung cancer as a POMDP and converting it into a belief MDP with the imperfect state information setting, we established the early diagnosis of lung cancer as an optimal stopping problem under the belief MDP. Using a deep reinforcement learning approach, we solved the optimal stopping problem and leveraged the property of Snell envelop to derive an interpretable stopping rule utilizing the convexity of the stopping region. Our numerical results on a real-world lung cancer clinical trial NLST demonstrated the superior performance of EarlyStop-RL compared to widely employed clinical models, individuals could use our mathematically validated model to inform personalized decision-making as a second opinion about LCS, and health systems could run the model at the population level to streamline the diagnostic process.

There are a few limitations of our method, which will be our future work. In order to smoothly integrate our Earlystop-RL into the current lung cancer screening process without significant disruptions to the standard clinical workflow, we utilized the clinical observation model based solely on clinical risk factors provided by radiologists, such as the nodule diameter, as input. In future research, we plan to explore representation learning or deep-learning techniques to extract more representative features and use the more powerfully dynamic Bayesian networks as the observation model while maintaining the model's interpretability. Additionally, conducting prospective studies, external validations, and assessing generalizability/robustness is crucial for AI methods in healthcare. These aspects will be the focus of our future work.

Acknowledgements

This work of Yifan Wang and Chuan Zhou is supported in part by the NIH grant U01CA216459.

References

Ajdari, A.; Niyazi, M.; Nicolay, N. H.; Thieke, C.; Jeraj, R.; and Bortfeld, T. 2019. Towards optimal stopping in radiation therapy. *Radiotherapy and Oncology*, 134: 96–100.

Ardila, D.; Kiraly, A. P.; et al. 2019. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature medicine*, 25(6): 954–961.

Bertsekas, D. P.; and Tsitsiklis, J. N. 1995. Neuro-dynamic programming: an overview. In *Proceedings of 1995 34th IEEE conference on decision and control*, volume 1, 560–564. IEEE.

Braca, P.; Gaglione, D.; Marano, S.; Millefiori, L. M.; Willett, P.; and Pattipati, K. 2021. Decision support for the quickest detection of critical COVID-19 phases. *Scientific reports*, 11(1): 8558.

Ery, J.; and Michel, L. 2021. Solving optimal stopping problems with Deep Q-Learning. *arXiv preprint arXiv:2101.09682*.

Fathan, A.; and Delage, E. 2021. Deep reinforcement learning for optimal stopping with application in financial engineering. *arXiv preprint arXiv:2105.08877*.

Huang, P.; Lin, C. T.; Li, Y.; Tammemagi, M. C.; Brock, M. V.; Atkar-Khattra, S.; Xu, Y.; Hu, P.; Mayo, J. R.; Schmidt, H.; et al. 2019. Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *The Lancet Digital Health*, 1(7): e353–e362.

Karatzas, I.; Karatzas, I.; Shreve, S.; and Shreve, S. E. 1991. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media.

Krishnamurthy, V. 2016. *Partially observed Markov decision processes*. Cambridge university press.

Larici, A. R.; Farchione, A.; Franchi, P.; Ciliberto, M.; Cicchetti, G.; Calandriello, L.; Del Ciello, A.; and Bonomo, L. 2017. Lung nodules: size still matters. *European respiratory review*, 26(146).

Leening, M. J.; Vedder, M. M.; Witteman, J. C.; Pencina, M. J.; and Steyerberg, E. W. 2014. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Annals of internal medicine*, 160(2): 122–131.

Liu, Z.; Yao, C.; Yu, H.; and Wu, T. 2019. Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things. *Future Generation Computer Systems*, 97: 1–9.

McKee, B. J.; Regis, S. M.; McKee, A. B.; Flacke, S.; and Wald, C. 2016. Performance of ACR Lung-RADS in a clinical CT lung screening program. *Journal of the American College of Radiology*, 13(2): R25–R29.

McNemar, Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2): 153–157.

McWilliams, A.; Tammemagi, M. C.; Mayo, J. R.; Roberts, H.; Liu, G.; Soghrati, K.; Yasufuku, K.; Martel, S.; Laberge, F.; Gingras, M.; et al. 2013. Probability of cancer in pulmonary nodules detected on first screening CT. *New England Journal of Medicine*, 369(10): 910–919.

Mehta, H. J.; Mohammed, T.-L.; and Jantz, M. A. 2017. The American College of Radiology lung imaging reporting and data system: potential drawbacks and need for revision. *Chest*, 151(3): 539–543.

Mridha, M.; Prodeep, A. R.; Hoque, A.; Islam, M.; Lima, A. A.; Kabir, M. M.; Hamid, M.; Watanobe, Y.; et al. 2022. A Comprehensive Survey on the Progress, Process, and Challenges of Lung Cancer Detection and Classification. *Journal of Healthcare Engineering*, 2022.

Nasrollahzadeh, A. A.; and Khademi, A. 2020. Optimal stopping of adaptive dose-finding trials. *Service Science*, 12(2-3): 80–99.

Ning, J.; Ge, T.; Jiang, M.; Jia, K.; Wang, L.; Li, W.; Chen, B.; Liu, Y.; Wang, H.; Zhao, S.; et al. 2021. Early diagnosis of lung cancer: which is the optimal choice? *Aging (Albany NY)*, 13(4): 6214.

Poor, H. V.; and Hadjiliadis, O. 2009. *Quickest detection*, volume 40. Cambridge University Press Cambridge.

Ritov, Y. 1990. Decision theoretic optimality of the CUSUM procedure. *The Annals of Statistics*, 1464–1469.

Robbins, H. A.; Cheung, L. C.; Chaturvedi, A. K.; Baldwin, D. R.; Berg, C. D.; and Katki, H. A. 2022. Management of lung cancer screening results based on individual prediction of current and future lung cancer risks. *Journal of Thoracic Oncology*, 17(2): 252–263.

Sarapata, E. A.; and De Pillis, L. 2014. A comparison and catalog of intrinsic tumor growth models. *Bulletin of mathematical biology*, 76(8): 2010–2024.

Shiryaev, A. N. 1963. On optimum methods in quickest detection problems. *Theory of Probability & Its Applications*, 8(1): 22–46.

Siegel, R. L.; Miller, K. D.; Fuchs, H. E.; and Jemal, A. 2022. Cancer statistics, 2022. *CA: a cancer journal for clinicians*, 72(1): 7–33.

Team, N. L. S. T. R. 2011a. The national lung screening trial: overview and study design. *Radiology*, 258(1): 243–253.

Team, N. L. S. T. R. 2011b. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5): 395–409.

Vaghi, C.; Rodallec, A.; Fanciullino, R.; Ciccolini, J.; Mochel, J. P.; Mastri, M.; Poignard, C.; Ebos, J. M.; and Benzekry, S. 2020. Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors. *PLoS computational biology*, 16(2): e1007178.

Wang, Y.; Zhou, C.; Ying, L.; Chan, H.-P.; Hadjiiski, L. M.; Chughtai, A.; and Kazerooni, E. A. 2021. Reinforced learning from serial CT to improve the early diagnosis of lung cancer in screening. In *Medical Imaging 2021: Computer-Aided Diagnosis*, volume 11597, 412–417. SPIE.

Wei, H.; Kang, X.; Wang, W.; and Ying, L. 2019. QuickStop: A Markov optimal stopping approach for quickest misinformation detection. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 3(2): 1–25.

Willsky, A.; and Jones, H. 1976. A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic control*, 21(1): 108–112.

Xie, L.; Zou, S.; Xie, Y.; and Veeravalli, V. V. 2021. Sequential (quickest) change detection: Classical results and new directions. *IEEE Journal on Selected Areas in Information Theory*, 2(2): 494–514.

Zhang, Q.; Wei, H.; Wang, W.; and Ying, L. 2022. On low-complexity quickest intervention of mutated diffusion processes through local approximation. In *Proceedings of the*

Twenty-Third International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing, 141–150.