Production and Operations Management

Reducing the Price of Naïveté in Return-to-Play from Sports-related Concussion

Gian-Gabriel P. Garcia*
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA giangarcia@gatech.edu

Mariel S. Lavieri
Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI lavieri@umich.edu

Thomas W. McAllister
Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN twmcalli@iupui.edu

Michael A. McCrea
Departments of Neurosurgery and Neurology, Medical College of Wisconsin, Milwaukee, WI mmccrea@mcw.edu

Steven P. Broglio
School of Kinesiology and Michigan Concussion Center, University of Michigan, Ann Arbor, MI broglio@umich.edu

CARE Consortium Investigators

*Corresponding author

Patient-reported outcomes play an increasingly important role in medical decision-making. Yet, patients whose objectives differ from their physician's may strategically report symptoms to alter treatment decisions. For example, athletes may under-report symptoms to expedite return-to-play from sports-related concussion. Thus, clinicians must implement treatment policies that mitigate the *Price of Naïveté*, i.e., the reduction in health outcomes due to naïvely believing strategically reported symptoms. In this study, we analyze dynamic treatment cessation decisions with strategic patients. Specifically, we formulate the Behavior-Aware Partially Observable Markov Decision Process (BA-POMDP), which optimizes the timing of treatment cessation decisions while accounting for known symptom-reporting behaviors. We then analytically characterize the BA-POMDP's optimal policy, leading to several practical insights. Next, we formulate the Behavior-Learning Partially Observable Markov Decision Process (BL-POMDP), which extends the BA-POMDP by learning a patient's symptom-reporting behavior over time. We show that the BL-POMDP is decomposable into several BA-POMDPs, allowing us to leverage the BA-POMDP's structural properties for solving the BL-POMDP. Then, we apply the BL-POMDP to return-to-play from sports-related concussion using data from 29 institutions across the United States. We estimate the Price of Naïveté by comparing the BL-POMDP to naïve benchmark policies. Accordingly, the BL-POMDP reduces premature return-to-play by over 44% and provides up to 3.63 additional health-adjusted athletic exposures per athlete compared to current practice. Overall, changing the interpretation of reported symptoms can better reduce the Price of Naïveté over adjusting treatment cessation thresholds. Therefore, to improve patients' health outcomes, clinicians must understand how strategic behavior manifests in patient-reported outcomes.

Key words: Stochastic dynamic programming, medical decision-making, behavioral modeling, patient-centered care, concussion management

History: Received: May 2022; accepted: May 2023 by Sergei Savin after two revisions.
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/poms.14024

1 Introduction

Clinical experts have been increasingly advocating for patient-centered care (Epstein and Street 2011). Patient-reported outcomes (PROs), such as symptom presentation, are a critical component of patient-centered care. PROs help physicians (hereafter referred to as doctors) assess a patient's health status and make treatment decisions. One major challenge in incorporating PROs is the potential for patients to purposely present misleading information (Lohr and Zebrack 2009). For example, in a recent study, between 60% and 80% of participants reported that they had intentionally withheld clinically relevant information from their doctors (Levy et al. 2018).

Patients may inaccurately report PROs for several reasons including poor memory, poor questionnaire or survey design, lack of knowledge, or the desire to respond in socially acceptable ways (Levy et al. 2018). However, patients may also be *strategic*. That is, strategic patients have different objectives than their doctors and consequently, they deliberately misrepresent PROs to influence the doctor's health assessment or treatment decisions. As such, strategic patients might intentionally under-report symptoms. For example, high school and collegiate athletes who are assessed for sports-related concussion (SRC) may under-report symptoms in an effort to return-to-play (RTP) more quickly (Meier et al. 2015, Conway et al. 2018). Strategic patients may also malinger, i.e., exaggerate their symptoms, for the purpose of "avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (Bass and Halligan 2014). For doctors, strategically reported PROs can complicate the process of accurately assessing a patient's health, leading to erroneous treatment decisions or wasted clinical resources. As such, doctors may pay the *Price of Naïveté* if they naïvely take PROs at face value, rather than interpreting potentially strategic PROs cautiously and adjusting treatment decisions accordingly.

Despite the potential issues associated with strategically reported PROs, assuming all patients exhibit strategic reporting behavior can backfire; biases about strategic or deceptive patient behavior can hamper open communication between the doctor and patient, leading to a damaged doctor-patient relationship and negative health outcomes. Furthermore, these issues disproportionately affect patients from marginalized groups (Perloff et al. 2006, Dehon et al. 2017), potentially due to biased expectations or negative stereotypes associated with these patients. Such consequences can be potentially mitigated when treatment planning occurs over several visits, wherein doctors have an opportunity to update their beliefs about the patient's PRO-reporting behavior as they interact over time. To this end, the operations research community has explored sequential treatment planning in many settings including chronic disease management and cancer treatment, among others (Keskinocak and Savva 2020, Capan et al. 2017). Yet, many of these approaches do not account for the potential of strategic PRO-reporting behavior over time. To improve patient health outcomes, these models must incorporate patient behavior while optimizing treatment decisions.

The goals of this research are three-fold. **Our first goal is to formulate and analyze an optimal sequential decision-making model which jointly captures uncertainty in a patient's health and PRO-reporting behavior**. To address this goal, we formulate an extension to the Partially Observable Markov Decision Process (POMDP), termed the *Behavior-Aware POMDP* (BA-POMDP), for treatment cessation problems. In this framework, the doctor dynamically optimizes the timing of treatment cessation and learns the patient's evolving health state through a combination of (potentially strategically) reported symptoms and objective clinical measures under the assumption that the patient's behavior type is known. Then, we formulate an extension of the BA-POMDP called the Behavior-Learning POMDP (BL-POMDP), wherein the patient's PRO-reporting behavior must be learned over time. For both models, we conduct an analytical study to identify key structural characteristics of each model and related managerial insights. **Our second goal is to analyze the Price of Naïveté.** To address this goal, we conduct an analytical study on a stylized treatment cessation example, comparing the BL-POMDP with the POMDP. This analysis leads to managerial insights on the cost of ignoring patients' symptom-reporting behavior. Finally, **our third goal is to analyze the performance of our modeling framework in return-to-play from sports-related concussion**. To address this goal, we parameterize the BL-POMDP using a large multi-site dataset on SRC among collegiate athletes. Drawing on recent estimates for symptom-reporting behavior across collegiate athletes, we estimate lower and upper bounds on the Price of Naïveté in terms of total discounted health utility by comparing our BL-POMDP with a POMDP and practice-based approaches, none of which account for symptom-reporting behavior. We further contextualize the Price of Naïveté in terms of (1) reducing the likelihood of premature RTP and (2) increasing each athlete's total health-adjusted athletic exposures after RTP.

This research builds upon the medical decision-making literature in which modeling frameworks incorporate patient behavior. Previous work in this area primarily focused on patients' adherence to optimized treatment decisions. Yet, few have explicitly modeled how a patient's objectives influences their adherence. For example, Shechter et al. (2008) and Barnett et al. (2017) assumed that patient adherences were known problem parameters, whereas Mason et al. (2012) and Ayer et al. (2016) assumed that adherence was unknown but unaffected by a patient's health state or treatment decisions. Patient adherence has also been modeled as a function of medication cost (Schell et al. 2019) or the use of adherence-improving interventions (Lobo et al. 2017, Suen et al. 2022). While adherence ultimately affects treatment decisions, it does not change how doctors interpret results of clinical assessments. In contrast, PRO-reporting behavior modifies the interpretation of clinical assessments, necessitating extensions of established modeling frameworks such as POMDPs.

Beyond adherence, others have developed modeling frameworks where the patient's behavior is derived from his or her own objectives. Schottmüller (2013) models a patient-doctor interaction in which the patient knows that the doctor has financial incentives to recommend certain treatment decisions. Specifically, the author develops and analyzes a single-period cheap talk model to derive the patient's symptom-reporting behavior. In contrast, our modeling framework considers a multi-period treatment decision problem. To this end, Aswani et al. (2018) and Mintz et al. (2020) develop multi-period frameworks in which the decision-maker must determine appropriate weight-loss interventions or incentives for patients with obesity. The patient's behavior (i.e., exercise and diet levels) is determined by the solution to his or her own utility-maximization problem which is solved independently of the doctor's response to their actions. In contrast, our case study considers the case in which the patient accounts for the doctor's potential response in determining a symptom-reporting strategy. Finally, work by Zhang et al. (2018) presents a stochastic game analysis in which a patient and doctor jointly determine chronic disease management activities. While we model patient-doctor interactions in a multi-period stochastic setting, their model considers a perfectly observable state space and "switching" structure in their actions. That is, only the patient can perform actions in some states and only the doctor can perform actions in other states. In contrast, our model considers a partially observable state space and leader-follower structure in the actions where both parties perform actions in every period.

Overall, this research makes the following contributions. First, **we formulate and apply the BA-POMDP and BL-POMDP frameworks.** In comparison to the classic POMDP, the BA-POMDP and its multiple-behavior extension, the BL-POMDP, explicitly model both objective clinical assessments and potentially strategically reported symptoms into the observation space. As a result, our models account for how these two types of observations impact a doctor's assessment of a patient's health and their subsequent treatment decisions. These frameworks extend POMDPs in a way that naturally fits sequential patient-doctor interactions in many healthcare settings. Second, **we analytically characterize the structure of the BA-POMDP and BL-POMDP.** Classic results for POMDPs (i.e., piecewise linearity and convexity of the value function) do not hold, in general, for the BA-POMDP. As such, we leverage Blackwell dominance to establish connections between the POMDP and BA-POMDP. These structural results lead to insights for managing strategic patients and approximating the optimal policy. Then, we prove that the BL-POMDP can be decomposed into several BA-POMDPs, allowing us to leverage our analysis of the BA-POMDP to solve the BL-POMDP. Third, **we are one of the first to approach return-to-play from concussion with a decision-theoretic approach.** With concussion identified as a major public health issue, management of SRC has undergone significant changes in the last decade. Both national and international guidelines are shifting from primarily expert consensus-based best practices to data-driven guidelines grounded in evidence-based research. Recent developments in the application of decision-theoretic methods for concussion management have largely focused on concussion diagnosis decisions (Garcia et al. 2019, 2020a,b). Critically, decision-theoretic approaches to optimizing the timing of RTP from SRC are scarce; to our knowledge, only the simulation-optimization model by Garcia et al. (2022) considers a mathematical optimization framework for RTP decisions. Notably, their model optimizes the length of the *symptom-free waiting period* to minimize post-RTP injury rates. However, their simulation-optimization modeling approach does not consider the impact of athletes' symptom-reporting behavior. Since RTP decisions are often left to clinical judgment, symptom over- and under-reporting could significantly impact the timing of RTP. As such, we extend this literature by incorporating symptom-reporting behavior in our model for optimizing the timing of RTP decisions. Moreover, we contribute to the development of data-driven RTP guidelines by applying and evaluating the BL-POMDP to RTP from SRC among collegiate athletes. Finally, **we provide a data-driven framework to tailor athlete-specific return-to-play criteria and demonstrate its improvement over current practice.** By leveraging multi-center data from the CARE Consortium, we tailor optimal RTP criteria for male collegiate football players with different concussion histories. This framework can generalize across a broader range of athlete types and account for differences in health utilities, symptom-reporting behavior, injury presentation, pre-RTP recovery trajectories, and sport-specific injury risks. On various clinically relevant performance measures, our numerical analysis demonstrates a marked improvement of these tailored RTP criteria over existing approaches.

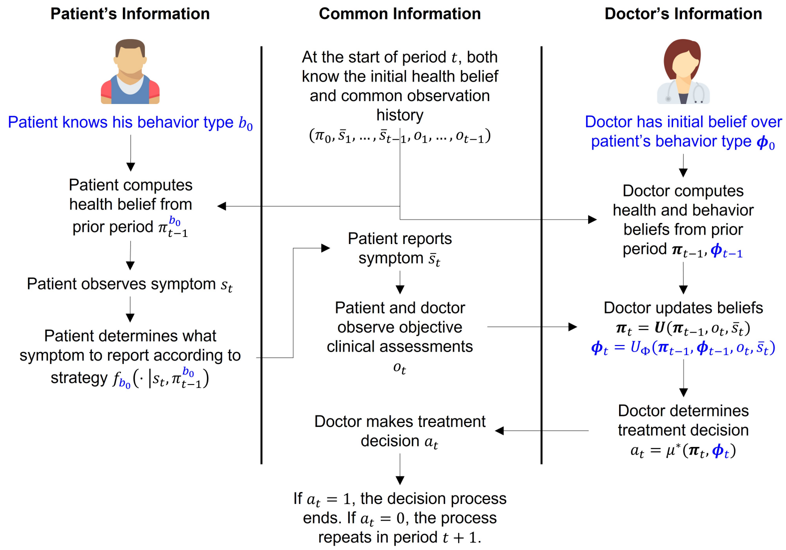
The remainder of this paper is organized as follows. In §2, we present the BA-POMDP modeling framework to optimize the timing of treatment cessation for strategic patients and derive structural results as well as related managerial insights. Then, in §3, we formulate and characterize the BL-POMDP, which extends the BA-POMDP to multiple behavior types. In §4, we formally define the Price of Naïveté and generate some key managerial insights. Next, in §5, we apply the BL-POMDP to optimize the timing of RTP from SRC. Finally, in §6, we discuss our findings.

2 Modeling Framework

In this section, we describe our model formulation for optimizing the timing of treatment cessation with potentially strategic patients. Our model is formulated from the perspective of a doctor who aims to anticipate potentially strategic reporting of PROs. This framework can be modified easily to fit the more commonly studied treatment initiation context. We illustrate our framework in Figure 1 and summarize all modeling notation in EC.1 in the e-companion.

In our model, a patient is recovering from a diagnosed disease and is undergoing a fixed treatment regimen. As part of this treatment regimen, the patient meets periodically with his doctor over several periods $t = 1, 2, \ldots$ until his doctor decides to cease treatment. In each meeting, the

Figure 1 Illustration of BA-POMDP and BL-POMDP modeling frameworks for treatment cessation decisions. Black text specifies key components of the BA-POMDP framework (see §2) with one behavior type. Blue text specifies additional model components for the BL-POMDP framework (see §3) with multiple behavior types.



patient observes a measure s_t, and based on this measure, reports a PRO $\bar{s}_t$ to his doctor. To simplify exposition, we refer to this measure and PRO as the observed symptom level and reported symptom level, respectively. The patient, who may be strategic, acts according to his own objectives. That is, he may report a symptom-level different from his observed symptom, i.e., $\bar{s}_t \neq s_t$, in an effort to manipulate the doctor's treatment cessation decision a_t. After observing the PRO $\bar{s}_t$, the doctor performs separate clinical evaluations on the patient and observes the results of these objective assessments, o_t. Based on the reported symptoms and the objective assessments, the doctor assesses the patient's health π_t, updates her belief about the patient's behavior type ϕ_t, and determines whether to cease treatment. If treatment ceases, the decision process ends. Otherwise, the patient's health evolves stochastically and the decision process repeats in the next meeting. This decision process continues indefinitely until treatment ceases or the patient dies. Note that the patient's recovery depends on receiving treatment and thus, premature treatment cessation could result in adverse health outcomes for the patient. Conversely, delayed treatment cessation may also have consequences, e.g., side effects and costs related to unnecessary treatment. Throughout this decision process, we assume that the patient and doctor have the same understanding of the disease dynamics and that the patient has limitations on their memory of experienced symptoms. Further, we assume that the doctor can reasonably estimate of the patient's symptom-reporting behavior — either by knowing it exactly (see §2.2) or by learning it over time (see §3). We do not assume that the patient has any information about the doctor's treatment cessation policy, but our model can be adapted to account for this setting (see §5.1). We now describe the model in more detail, starting with the case in which the patient's behavior type is known.

2.1 Patient's Health Dynamics

For each decision period $t = 1, 2, \ldots$ we model the patient's health, its progression over time, and its corresponding observations as a stationary discrete-time Hidden Markov Model (HMM). For clarity, we use a subscript t to denote the status of an HMM component during a specific decision period. The components of this HMM are as follows.

Core health states. Let $\mathcal{H} := \{0, \ldots, H\}$ be the ordered set of unobservable core health states. The health state $h = 0$ represents the state in which the patient has fully recovered from the disease while health states $h = 1, \ldots, H$ represent diseased states in increasing severity. For example, in the case of SRC, the health states can be modeled such that $h = 0$ represents no concussion, $h = 1$ represents asymptomatic concussion, and $h = 2$ represents symptomatic concussion. Without loss of generality, death can be modeled by setting H as an observable absorbing state.

State transition probabilities. The initial state distribution is given by the vector π_0 with components $\pi_0(h) = \mathbb{P}(h_0 = h)$. The patient's health state evolves at the start of each period t according to a transition probability matrix P with components $P(h'|h) := \mathbb{P}(h_{t+1} = h'|h_t = h)$ for all $h, h' \in \mathcal{H}$. Because h is not directly observable by the patient or doctor, it must be inferred from symptoms and objective health assessments (i.e., objective observations).

Symptom observations. After the patient's health evolves to h at the start of period t, the patient (and only the patient) observes a symptom level $s \in \mathcal{S} := \{0, \ldots, S\}$ where all $s \in \mathcal{S}$ are ordered in increasing severity. The likelihood of observing s_t is summarized in an observation probability matrix $Q^{\mathcal{S}}$ with entries $Q^{\mathcal{S}}(s|h) := \mathbb{P}(s_t = s|h_t = h)$ for all $s \in \mathcal{S}$ and $h \in \mathcal{H}$. After observing s, the patient reports $\bar{s} \in \mathcal{S}$ to the doctor according to his strategy (see §2.2).

Objective observations. After the patient reports $\bar{s}$ to the doctor, she performs additional clinical assessments (e.g., neurocognitive exam) which cannot be strategically manipulated. The results of this assessment are denoted by the objective observation $o \in \mathcal{O} := \{0, \ldots, O\}$, where all $o \in \mathcal{O}$ are ordered in increasing severity. The likelihood of observing o_t is summarized in an observation probability matrix $Q^{\mathcal{O}}$ with components $Q^{\mathcal{O}}(o|h) := \mathbb{P}(o_t = o|h_t = h)$ for all $o \in \mathcal{O}$ and $h \in \mathcal{H}$.

Assumption 1. *We assume that symptom observations and objective observations are independent conditional on the patient's health state. That is,* $\mathbb{P}(o_t = o, s_t = s|h_t = h) = Q^{\mathcal{O}}(o|h)Q^{\mathcal{S}}(s|h)$.

Assumption 1 implies that s and o test different functional domains with minimal overlap. We show in our case study (§5) that this assumption is reasonable for concussion recovery dynamics.

2.2 Doctor's Treatment Cessation Problem

We now formulate the doctor's treatment cessation problem under strategic symptom-reporting behavior. As our base model, we consider the case in which the patient's symptom-reporting strategy is known exactly by the doctor. We formulate this decision problem as an extension to the classic POMDP framework derived in Krishnamurthy (2012). We term this model the Behavior-Aware POMDP (BA-POMDP), which has the following components.

Actions. After observing $\bar{s}_t$ and o_t, the doctor must choose an action $a \in \mathcal{A} := \{0, 1\}$. If she chooses $a = 0$, the decision process continues in the next period. Otherwise, choosing $a = 1$ ceases treatment and the decision process ends.

Rewards. After performing an action, the doctor immediately receives a reward based on her reward function $r^d : \mathcal{H} \times \mathcal{A} \to \mathbb{R}_+$, where $r^d(\cdot, 0)$ describes the reward received for continuing treatment and $r^d(\cdot, 1)$ describes a large lump sum reward received for ceasing treatment. We use r^d_a to denote the reward vector with components $r^d_a(h) = r^d(h, a)$.

Health state beliefs. Prior to beginning the decision process, the doctor has an initial health state belief $\pi_0 \in \Pi$, where $\Pi := \{\pi \in \mathbb{R}^{|\mathcal{H}|}_+ : \sum_{h\in\mathcal{H}} \pi(h) = 1\}$ is the $|\mathcal{H}| - 1$ probability simplex. Before making a decision in period t, the doctor's information state is given by the tuple $(\pi_0, \bar{s}_1, ..., \bar{s}_t, o_1, ..., o_t)$.

Honest health state belief update. If the patient reports symptoms honestly (i.e., $\bar{s}_t = s_t$ for all t), a sufficient statistic for the doctor's information state is the health belief state $\pi_t \in \Pi$ with components $\pi_t(h) = \mathbb{P}(h_t = h | \pi_0, \bar{s}_1, ..., \bar{s}_t, o_1, ..., o_t)$ for all $h \in \mathcal{H}$ (Åström 1965). Given $\pi_{t-1} = \pi, o_t = o$, and $\bar{s}_t = s_t = s$,

$$\pi_t(h) = \frac{Q(o|h)Q(s|h)\sum_{h'\in\mathcal{H}} P(h|h')\pi(h')}{\dot{C}(\pi, o, s)} \text{ for all } h \in \mathcal{H}, \tag{1}$$
$$\dot{C}(\pi, o, s) := \mathbb{P}(o_t = o, s_t = s | \pi_{t-1} = \pi) = \sum_{h\in\mathcal{H}} Q(o|h)Q(s|h) \sum_{h'\in\mathcal{H}} P(h|h')\pi(h').$$

We denote the updated belief vector using the belief update function $\dot{U}(\pi, o, s) = \pi_t$. However, as we show in our analytical and numerical analysis (see §2.3 and §5, respectively), this type of health state belief update has dire consequences when used with strategic patients.

Patient's symptom-reporting strategy. The patient's symptom-reporting strategy f maps his observation history (i.e., sequence of past symptom observations, objective observations, and reported symptoms) to his reported symptom $\bar{s}_t$. The primary impact of the patient's symptom-reporting strategy is in how the doctor updates her belief about the patient's health by modifying observation probabilities for symptoms. We soon make this connection more precise, but first, we make the following assumption to facilitate our analysis of treatment cessation policies.

Assumption 2. *The patient has limited private memory, i.e., at the beginning of period t, his information state is given by $(\pi_0, \bar{s}_1, ..., \bar{s}_{t-1}, o_1, ..., o_{t-1}, s_t)$.*

We remark that Assumption 2 is necessary to ensure the tractability of our modeling approach (see EC.2 in the e-companion). Assumption 2 implies that the patient's symptom-reporting strategy depends only on his current symptoms s_t, history of previously reported symptoms $\bar{s}_1, \ldots, \bar{s}_{t-1}$, and objective observations $o_1, \ldots, o_t$. In other words, it is easier for the patient to remember what symptoms he reported (e.g., by reviewing past medical records) than the symptoms he actually experienced. This assumption applies when a patient's recall of medical history has low reliability or accuracy, as was found in athletes with SRC (Arends et al. 2019, Kerr et al. 2015). We test the practical implications of this assumption in EC.6.6 in the e-companion.

An important consequence of Assumption 2 is that the patient's information state in period t, prior to reporting a symptom, can be summarized by the tuple (π_{t-1}, s_t), i.e., the doctor's health state belief at the end of period $t-1$ and his current symptom observation. Given this information state, the patient's strategy is a mapping $f : \Pi \times \mathcal{S} \times \mathcal{S} \to [0,1]$, where $f(\bar{s}_t|\pi_{t-1}, s_t)$ denotes the probability that the patient reports $\bar{s}_t$ given his information state (π_{t-1}, s_t). We next describe how the health state belief π_t is updated recursively when the patient reports symptoms strategically.

Strategic health state belief update. If the patient reports symptoms strategically, the doctor's health state belief update can only be computed recursively via Bayes' rule if she knows the patient's strategy f and can recover the mapping between $\bar{s}$ and his information state based on her own information state. Under Assumption 2, the doctor's information state at the end of period t and the patient's information state at the start of period $t+1$ prior to seeing s_{t+1} are equal. Thus, we satisfy this requirement. Consequently, we update the doctor's health state belief using Bayes Rule; given $\pi_{t-1} = \pi$, $o_t = o$, and $\bar{s}_t = \bar{s}$, the health state belief can be updated via:

$$\pi_t(h) := \frac{Q(o|h)\sum_{s\in\mathcal{S}} f(\bar{s}|\pi, s) Q^{\mathcal{S}}(s|h) \sum_{h'\in\mathcal{H}} P(h|h')\pi(h')}{C(\pi, o, \bar{s})} \text{ for all } h \in \mathcal{H}, \tag{2}$$

$$C(\pi, o, \bar{s}) := \mathbb{P}(o_t = o, \bar{s}_t = \bar{s}|\pi_{t-1} = \pi) = \sum_{h\in\mathcal{H}} Q(o|h) \sum_{s\in\mathcal{S}} f(\bar{s}|\pi, s) Q^{\mathcal{S}}(s|h) \sum_{h'\in\mathcal{H}} P(h|h')\pi(h').$$

We denote by $U(\pi, o, \bar{s}) = \pi_t$ the patient's updated belief vector. This interdependence between the doctor's beliefs, the patient's beliefs, and the patient's strategy link the doctor's decisions with the patient's reported symptoms.

Optimality equations. The doctor aims to determine a stationary policy $\mu : \Pi \to \mathcal{A}$ which maximizes her expected total discounted reward. The optimal value function V can be determined by solving the following Bellman equation for all $\pi \in \Pi$

$$V(\pi) = \max_{a\in\mathcal{A}} \begin{cases} \pi^\top r_0^d + \rho \sum_{o\in\mathcal{O}} \sum_{\bar{s}\in\mathcal{S}} C(\pi, o, \bar{s}) V(U(\pi, o, \bar{s})) & a = 0 \\ \pi^\top r_1^d & a = 1 \end{cases}. \tag{3}$$

The optimal policy $\mu^*(\pi)$ corresponds to the maximizing action for $V(\pi)$.

2.3 Analytical Results

In this section, we analyze structural properties of the BA-POMDP by leveraging comparisons to the POMDP. All proofs are detailed in EC.3 of the e-companion. We apply the following definitions and assumptions throughout this section.

Definition 1 (Submodular). A function of two variables $f(x,y)$ is said to be submodular if for every $x' > x$ and $y' > y$, $f(x,y') + f(x',y) \geq f(x,y) + f(x',y')$.

Assumption 3. *The doctor's reward function $r^d(h,a)$ is submodular in (h,a) and non-increasing in h for all $a \in \mathcal{A}$.*

Assumption 3 implies that the "opportunity cost" of delaying treatment cessation increases as the patient's health improves. For SRC management, these assumptions could imply that an athlete who is further along their recovery trajectory may not reduce their post-RTP injury risk by waiting an additional day to RTP as much as an athlete whose SRC is still acute.

Definition 2 (Totally Positive of Order 2 (TP2)). A matrix M is TP2 if all of its second-order minors are non-negative. Equivalently, a transition kernel M is TP2 if $M_i(j)M_{i+1}(j') \geq M_i(j')M_{i+1}(j)$ for all $j < j'$, where M_i denotes the i^{th} row of M.

Assumption 4. *The matrices P, $Q^{\mathcal{O}}$, and $Q^{\mathcal{S}}$ are TP2.*

Assumption 4 implies that sicker patients are more likely to remain sick or die than healthier patients. Furthermore, sicker patients are more likely to present more "severe" assessments and symptoms than healthier patients. In the context of SRC, this assumption means that athletes who are still in the acute, symptomatic stages of the injury are less likely to reach full recovery in the same amount of time as athletes in the asymptomatic stage of the injury.

In this section, $F^\pi \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{S}|}$ denotes a matrix with components $F^\pi(\bar{s}|s) = f(\bar{s}|\pi,s)$. We express the observation probability matrix for reported symptoms as $\bar{Q}^{\mathcal{S}}(\pi) = Q^{\mathcal{S}}F^\pi$ with entries $\mathbb{P}(\bar{s}_t = \bar{s}|h_t = h, \pi_t = \pi, b_0 = b)$. If patients are honest(i.e., $F^\pi = I$ for all $\pi \in \Pi$) or the strategy f does not depend on π(i.e., $F^\pi = F$ for all $\pi \in \Pi$), the BA-POMDP reduces to a POMDP. Let $\dot{V}$ and $\dot{\mu}^*$ be the optimal value function and policy associated with this POMDP. We now establish Theorem 1, which characterizes the structure of the BA-POMDP when patients report symptoms truthfully.

Theorem 1. *If the patient is honest, the following properties hold.*

1. *The set $\dot{\Pi}_1 := \{\pi \in \Pi : \dot{\mu}^*(\pi) = 1\}$ is convex.*
2. *Define $r^d_L := r^d_1 - (r^d_0 + \rho P r^d_1)$ and $\Pi_L := \left\{\pi \in \Pi : \pi^\top r^d_L \geq 0\right\}$. For each $h \in \mathcal{H}$, let $\nu_h := \left\{\pi \in \Pi : \pi(0) = \frac{-r^d_L(h)}{r^d_L(0) - r^d_L(h)}, \pi(h) = 1 - \pi(0)\right\}$ define the vertices in which the boundary of Π_L intersects*

the faces of Π*. Additionally, define the diagonal matrices* $D_O^{\mathcal{O}} := diag(Q^{\mathcal{O}}(O|0), ..., Q^{\mathcal{O}}(O|H))$ *and* $D_S^{\mathcal{S}} := diag(Q^{\mathcal{S}}(S|0), ..., Q^{\mathcal{S}}(S|H))$*. If*

$$\left(D_O^{\mathcal{O}} D_S^{\mathcal{S}} P^\top \nu_i\right)^\top r_L^d \geq 0 \text{ for all } i \in \mathcal{H}, \tag{4}$$

then $\dot{\Pi}_1 = \Pi_L$*, i.e.,* $\dot{\mu}^*(\pi) = 1$ *if* $\pi^\top r_L^d \geq 0$ *and* $\dot{\mu}^*(\pi) = 0$ *otherwise.*

Theorem 1 shows that the BA-POMDP's optimal policy for honest patients has an interpretable structure, i.e., there is a (potentially non-linear) decision threshold on the belief space such that treatment cessation is optimal for patients who are healthier than that threshold. In some cases (Property 2), the myopic one-step look ahead policy is optimal so the decision threshold is linear (Yasuda 1988). However, for strategic patients, the BA-POMDP's value function is not generally piecewise linear and convex in π as it is for POMDPs.

Proposition 1. *In general,* $V(\pi)$ *is not piecewise linear and convex in* π*.*

Proposition 1 arises because symptom observations depend on the patient's symptom-reporting strategy f, which can cause belief updates to be non-linear in π unlike in POMDPs. For example, f may be discontinuous in π (see Figure EC.1 in the e-companion). Hence, exact solution approaches which rely on this property (e.g., use of α-vectors) cannot be applied for BA-POMDPs and typical approaches to establishing structural results for POMDPs cannot be taken for the BA-POMDP. Thus, we begin our analysis by establishing the connection between treatment cessation for honest and strategic patients. In Proposition 2, we show that the doctor's performance is no better with a strategic patient than it would be with an honest patient.

Proposition 2. *For any patient strategy* f*, the following inequalities hold:*

$$\sum_{\bar{s}\in\mathcal{S}} C(\pi, o, \bar{s})\dot{V}(U(\pi, o, \bar{s})) \leq \sum_{\bar{s}\in\mathcal{S}} \dot{C}(\pi, o, \bar{s})\dot{V}(\dot{U}(\pi, o, \bar{s})) \text{ for any } o \in \mathcal{O} \text{ and } \pi \in \Pi \tag{5}$$

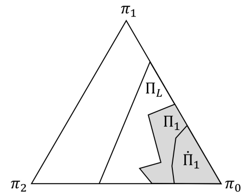
$$V(\pi) \leq \dot{V}(\pi) \text{ for all } \pi \in \Pi. \tag{6}$$

Proposition 2 reflects Blackwell dominance (Blackwell 1953); since $\bar{Q}^{\mathcal{S}}(\pi) = Q^{\mathcal{S}} F^\pi$, symptoms reported by a strategic patient are less informative than those reported by an honest patient, making the doctor's decision process more challenging. Furthermore, the bounds in (5)-(6) allow us to leverage the structure of $\dot{\mu}^*$ to characterize μ^*. We state these connections in Theorem 2.

Theorem 2. *The BA-POMDP's optimal policy* μ^* *satisfies the following properties.*

1. *For any* $\pi \in \Pi$, $\pi^\top r_L^d < 0$ *implies* $\mu^*(\pi) = 0$.
2. *For any* $\pi \in \Pi$, $\dot{\mu}^*(\pi) = 1$ *implies* $\mu^*(\pi) = 1$.

Figure 2 Theorem 2 implies that $\dot{\Pi}_1 \subseteq \Pi_1 \subseteq \Pi_L$ with $\mu^*(\pi) = 1$ in all shaded regions and $\mu^*(\pi) = 0$ in all non-shaded regions.



Theorem 2 reveals several practical insights. In particular, Property 1 implies that it is optimal to wait one more period if it is better to cease treatment in the next period than in the current period. Moreover, Property 2 implies that it is optimal to cease treatment for a strategic patient if it would also be optimal to cease treatment assuming they are honest in the next period. This latter case may be most useful, for example, if patients only report symptoms strategically early in the decision process and become more honest in reporting as they get healthier. Such behavior may be exhibited by athletes with SRC who aim to expedite RTP by under-reporting symptoms; however, as they become healthier, their symptom levels will start to become low enough that they do not need to under-report symptoms to signal that they are healthy.

Theorem 2 also leads to geometric insights (see Figure 2). Define the sets $\dot{\Pi}_1 = \{\pi \in \Pi : \dot{\mu}^*(\pi) = 1\}$, $\Pi_1 = \{\pi \in \Pi : \mu^*(\pi) = 1\}$, and $\Pi_L = \{\pi \in \Pi : \pi^\top r_L^d \geq 0\}$. Theorem 2 implies that $\dot{\Pi}_1 \subseteq \Pi_1 \subseteq \Pi_L$. This characterization is useful for guiding the allocation of grid points when approximating μ^* (see EC.4 in the e-companion). Furthermore, μ^* inherits important structural properties of $\dot{\mu}^*(\pi)$, e.g., Π_1 contains a non-linear threshold for which it is optimal to cease treatment for all patients believed to be healthier than that threshold. In Corollary 1, we establish necessary conditions which guarantee that μ^* is characterized by an interpretable linear decision threshold.

Corollary 1. *If (4) holds, $\mu^*(\pi) = 1$ if $\pi^\top r_L^d \geq 0$ and $\mu^* = 0$ otherwise.*

Theorem 2 and Corollary 1 hold regardless of the patient's strategy f. Furthermore, the conditions required for these results can be verified using only the problem data. However, by ignoring symptom-reporting behavior, the doctor's belief about the patient's health will not be correctly adjusted and could lead to erroneous treatment cessation decisions. In Proposition 3, we show that a doctor who anticipates symptom under-reporting (resp., over-reporting) from the patient will update her belief about the patient's health more conservatively (resp., aggressively) than a doctor who believes that the patient is reporting symptoms truthfully.

Proposition 3. *If F^π is lower triangular, then $U(\pi, o, \bar{s}) \succeq_R \dot{U}(\pi, o, \bar{s})$. If F^π is upper triangular, then $\dot{U}(\pi, o, \bar{s}) \succeq_R U(\pi, o, \bar{s})$.*

Proposition 3 implies that if a patient under-reports (resp., over-reports) symptoms, a doctor who does not account for strategic behavior would think that the patient is healthier (resp., sicker) than he actually is. Consequently, failing to account for strategic behavior can result in premature (resp., delayed) treatment cessation for patients who under-report (resp., over-report) symptoms. Theorem 2 and Proposition 3 illustrate the important role played by the health state belief update for strategic patients. We strengthen these findings with our numerical analysis in §5.

3 A Multi-behavior Extension to the BA-POMDP

In practice, the patient's symptom-reporting behavior is initially unknown to the doctor. Instead, the doctor must learn the patient's behavior dynamically during their interactions. In Figure 1, we highlight the novel aspects of the Behavior-Learning POMDP (BL-POMDP), which generalizes the BA-POMDP in this regard and includes the following additional model elements.

Patient behavior types. We denote by $\mathcal{B} := \{0, 1, ..., B\}$ the set of all patient behavior types known to the doctor. In practice, doctors may have a sense of what types of symptom-reporting behaviors to expect based on their past experience. Further, it is not necessary for a doctor to know *every possible* behavior type so long as all behavior types she encounters can be explained by a combination of behavior types that she knows (see Assumption 5 and Proposition 4). We assume that each patient knows his own behavior type $b_0 \in \mathcal{B}$ and each $b \in \mathcal{B}$ dictates a symptom-reporting strategy f_b. As such, each patient type would generate its own belief state update function. We add a subscript b to the notation in §2.2 to reflect this behavior-specific change. For example, $U_b(\pi, o, \bar{s})$ denotes the belief update function if the patient reports symptoms according to f_b.

Health state and behavior type beliefs. When the patient's behavior type is unknown to the doctor, her information state at time t is given by $(\pi_0, \phi_0, o_1, ..., o_t, \bar{s}_1, ..., \bar{s}_t)$. Here, $\phi_0 \in \Phi$ is an initial behavior belief vector with components $\phi_0(b) = \mathbb{P}(b_0 = b)$ and $\Phi := \{\phi \in \mathbb{R}_+^{|\mathcal{B}|} : \sum_{b \in \mathcal{B}} \phi(b) = 1\}$ represents the $|\mathcal{B}| - 1$ probability simplex. Let π_t^b denote the doctor's belief about the patient's health if his behavior type was b. We can summarize her information state by the tuple $(\boldsymbol{\pi}_t, \phi_t)$, where $\boldsymbol{\pi}_t = (\pi_t^0, ..., \pi_t^B)$ is a tuple of health belief states for all $b \in \mathcal{B}$ and the vector ϕ_t is the doctor's behavior belief vector in period t with components $\phi_t(b) = \mathbb{P}(b_0 = b | \pi_0, \phi_0, o_1, ..., o_t, \bar{s}_1, ..., \bar{s}_t)$. Given the previous state $(\boldsymbol{\pi}_{t-1} = \boldsymbol{\pi}, \phi_{t-1} = \phi)$ and current observations $o_t = o, \bar{s}_t = s$, the behavior belief vector is updated according to

$$\phi_t(b) = \frac{C_b(\pi^b, o, \bar{s})\phi(b)}{C_\Phi(\boldsymbol{\pi}, \phi, o, \bar{s})} \text{ for all } b \in \mathcal{B}, \tag{7}$$

$$C_\Phi(\boldsymbol{\pi}, \phi, o, \bar{s}) := \mathbb{P}(o_t = o, \bar{s}_t = \bar{s} | \boldsymbol{\pi}_t = \boldsymbol{\pi}, \phi_t = \phi) = \sum_{b' \in \mathcal{B}} C_{b'}(\pi^{b'}, o, \bar{s})\phi(b').$$

We denote the behavior belief update function as $U_\Phi(\boldsymbol{\pi}, \phi, o, \bar{s}) = \phi_t$. Additionally, the tuple of health state belief vectors is updated according to $\boldsymbol{\pi}_t = \mathbf{U}(\boldsymbol{\pi}, o, \bar{s}) = (U_0(\pi^0, o, \bar{s}), ..., U_B(\pi^B, o, \bar{s}))$. We remark that the behavior belief update function $U_\Phi(\boldsymbol{\pi}, \phi, o, \bar{s})$ essentially updates the doctor's belief about the patient's symptom-reporting behavior by comparing his reported symptoms to the what she would expect to see from patients of different behavior types. In principal, this process is similar to how a doctor may update their belief in practice based on their past experience dealing with different types of patients.

Optimality equations. The optimal value function V is obtained by solving the following Bellman Equations for all $\boldsymbol{\pi}$ and ϕ

$$V(\boldsymbol{\pi},\phi)=\max_{a\in\mathcal{A}}\begin{cases}\sum_{b\in\mathcal{B}}\phi(b)(\pi^b)^\top r_0^d \\ \quad +\rho\sum_{o\in\mathcal{O}}\sum_{\bar{s}\in\mathcal{S}}C_\Phi(\boldsymbol{\pi},\phi,o,\bar{s})V(\mathbf{U}(\boldsymbol{\pi},o,\bar{s}),U_\Phi(\boldsymbol{\pi},\phi,o,\bar{s})) & a=0 \\ \sum_{b\in\mathcal{B}}\phi(b)(\pi^b)^\top r_1^d & a=1\end{cases}. \tag{8}$$

The optimal policy $\mu^*(\boldsymbol{\pi},\phi)$ corresponds to the maximizing action for $V(\boldsymbol{\pi},\phi)$.

Let $V_b(\cdot)$ denote the BA-POMDP value function for a patient of behavior b. If ϕ satisfies $\phi(b)=1$ for some $b\in\mathcal{B}$, we have $V(\boldsymbol{\pi},\phi)=V_b(\pi^b)$ implying that the BA-POMDP is a special case of the BL-POMDP. Given that the BL-POMDP's state space, $\Pi^{|\mathcal{B}|}\times\Phi$, is the Cartesian product of several probability simplexes, the BL-POMDP suffers from the curse of dimensionality. To this end, our analytical study in §3.1 gives way to a numerical solution approach for approximating μ^*.

3.1 Analysis of the BL-POMDP

In this section, we characterize important analytical properties of the BL-POMDP model. First, we study properties of the behavior belief state ϕ_t. A potential practical concern is that a patient's behavior type may not be contained within the set $\mathcal{B}$. Hence, a particular sequence of observations may lead to statistical inconsistency issues during belief updating. Consider the following example.

Example 1. Let $\mathcal{B}=\{0,1,...,S\}$ and suppose that patients with behavior type b report symptom $\bar{s}_t=b$ for all t. That is, $f_b(\bar{s}|s,\pi)=1$ if $\bar{s}=s=b$ and $f_b(\bar{s}|s,\pi)=0$ otherwise. In period $t=1$, a reported symptom of $\bar{s}_1=0$ would produce an updated behavior belief $\phi_1=[1\ 0\ \cdots\ 0]^T$. That is, the doctor believes that the patient is of type 0. However, if in period $t=2$, the patient reports $\bar{s}_2=1$, then it is not possible to update ϕ_2 since $f_0(1|\pi,s)=0$ for all $\pi\in\Pi$ and $s\in\mathcal{S}$.

To avoid the issue highlighted by Example 1, we make the following assumption.

Assumption 5. *There exists at least one behavior type $b\in\mathcal{B}$ for which $\phi_0(b)>0$ and $f_b(\bar{s}|\pi,s)>0$ for all $\bar{s},s\in\mathcal{S}$ and $\pi\in\Pi$.*

Assumption 5 is known as the "grain of truth" condition (Doshi and Gmytrasiewicz 2006), i.e., "any sequence of reported symptoms is possible," thereby avoiding any inconsistency issues when updating belief states. If all entries of the observation matrix $Q^{\mathcal{S}}$ are non-zero and there exists some $b\in\mathcal{B}$ describing an honest patient, then Assumption 5 is satisfied. Assumption 5 could also be satisfied by a simple scaling procedure to ensure that all entries of $Q^{\mathcal{S}}$ are non-zero (see EC.5.2 for details) or including an *uninformative* behavior type with strategy $f_b(\bar{s}|s,\pi) = 1/|\mathcal{S}|$ for all $\bar{s}, s \in \mathcal{S}$ and $\pi \in \Pi$. We can now establish the limiting characteristics of the behavior belief state ϕ_t.

PROPOSITION 4. *The limit $\phi_\infty = \lim_{t\to\infty} \phi_t$ exists and ϕ_t converges to ϕ_∞ almost surely in L^2.*

Critically, Proposition 4 implies that the BL-POMDP model can be applied even when a patient's behavior is not captured by any behavior type $b \in \mathcal{B}$. As such, reasonable approximations to patient behavior types may suffice to apply the model in practice. We numerically test the performance of the BL-POMDP model for unaccounted behavior types in EC.6.6 in the e-companion.

REMARK 1. The doctor could have non-zero beliefs over multiple behavior types in the limit. This phenomenon occurs when two behavior types generate identical probability measures over all possible future observation sequences. For example, suppose that $\mathcal{B} = \{0,1,2\}$, where type $b = 0$ satisfies Assumption 5 and types $b = 1,2$ report symptoms honestly once $\pi(0)$ exceeds a threshold τ_b, with $\tau_1 \le \tau_2$. Then, for any π such that $\pi(0) \ge \tau_2$, we have $f_1(\bar{s}|\pi,s) = f_2(\bar{s}|\pi,s)$.

Now, we study properties of the optimal value function for the BL-POMDP. Recall that $V_b(\cdot)$ corresponds to the BA-POMDP value function for a patient of known behavior type b. In Theorem 3, we facilitate the solution of the BL-POMDP by showing that the value function $V(\boldsymbol{\pi},\phi)$ can be restated in terms of BA-POMDP models corresponding to each $b \in \mathcal{B}$.

THEOREM 3. *The value function for the BL-POMDP can be reformulated as*

$$V(\boldsymbol{\pi},\phi) = \max_{a\in\mathcal{A}} \begin{cases} \sum_{b\in\mathcal{B}} \phi(b)\Big((\pi^b)^\top r_0^d + \rho \sum_{o\in\mathcal{O}} \sum_{\bar{s}\in\mathcal{S}} C_b(\pi^b,o,\bar{s}) V_b(U_b(\pi^b,o,\bar{s}))\Big) & a=0 \\ \sum_{b\in\mathcal{B}} \phi(b)(\pi^b)^\top r_1^d & a=1 \end{cases}. \tag{9}$$

Theorem 3 implies that the BL-POMDP's value function and optimal policy follow from the value functions of $|\mathcal{B}|$ BA-POMDPs. Further, the BA-POMDP's structural properties can be exploited when solving the BL-POMDP. Theorem 3 also suggests that doctors need not consider behavior types that have been ruled out. For example, if a patient begins to report symptoms that are too low for an over-reporting type of patient, then the doctor can treat the patient as if he were under-reporting symptoms or honest. Moreover, if $\mu_b^*(\pi^b) = 1$ (resp., $\mu_b^*(\pi^b) = 0$) for all b such that $\phi(b) > 0$, then $\mu^*(\boldsymbol{\pi},\phi) = 1$ (resp., $\mu^*(\boldsymbol{\pi},\phi) = 0$). Hence, it is optimal to cease treatment if it is optimal to cease treatment for all plausible behavior types (i.e., $\phi(b) > 0$).

4 Price of Naïveté

We now summarize key managerial insights for doctors from our analytical study through the Price of Naïveté. The Price of Naïveté quantifies the "cost" incurred by naïve policy p' (i.e., does not explicitly account for symptom-reporting behavior) compared to the BA-POMDP's optimal policy (for a single behavior type) or BL-POMDP's optimal policy (for multiple behavior types), denoted μ^*. For a single behavior type, we mathematically define the Price of Naïveté as

$$\mathtt{PoN}(\pi_0, p', f) := v(\pi_0, \mu^*, f) - v(\pi_0, p', f), \tag{10}$$

where $v(\pi_0, p, f) = \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^{t-1} r(h_t, a_t) \middle| h_0 \sim \pi_0, a_t \sim p, \bar{s}_t \sim f\right]$ denotes the total expected discounted reward given that the patient's initial state probabilities are π_0, the actions are taken using the policy p, and the patient reports symptoms according to f. We now highlight the following key managerial insights, taking the naïve policy to be the POMDP's optimal policy, i.e., $p' = \dot{\mu}^*$.

Even if the policies μ^* and $\dot{\mu}^*$ are equal, the Price of Naïveté may be incurred if symptom-reporting behavior is not considered in belief updates. For example, if (4) holds, then by Corollary 1 and Theorem 1, both μ^* and $\dot{\mu}^*$ are the same one-step look-ahead policies. Yet, if the patient strategically under-reports (resp., over-reports) symptoms (see Proposition 3), a doctor who does not account for symptom-reporting behavior in their belief updates may prematurely cease treatment (resp., delay treatment cessation). As such, we may have $\mathtt{PoN}(\pi_0, \dot{\mu}^*, f) > 0$.

Depending on the patient's initial health assessment, no Price of Naïveté may be paid, even if the patient plans to report symptoms strategically. Because $\mathtt{PoN}(\pi_0, \dot{\mu}^*, f)$ depends on both π_0 and f, a patient may report symptoms strategically only for some subset of health belief states and report symptoms truthfully otherwise. If the patient's belief about his health remains only in the subset where he reports honestly, then he will report symptoms truthfully throughout the entire decision process. For example, suppose that the patient prefers to cease treatment early, so his strategy is to under-report symptoms unless he is sufficiently healthy, e.g., $\pi(0) \geq \tau$ for some threshold τ. If the initial belief satisfies $\pi_t(0) \geq \tau$ for all t, then $\pi_t = \dot{\pi}_t$ for all t. That is, health belief states will not differ between the BA-POMDP and the POMDP. Moreover, if $\mu^* = \dot{\mu}^*$, the time of treatment cessation will be the same for both policies leading to no Price of Naïveté , i.e., $\mathtt{PoN}(\pi_0, p', f) = 0$.

By combining the previous two insights, *the Price of Naïveté can be mitigated by knowing when to believe and when to ignore reported symptoms.* Consider a patient who under-reports symptoms to expedite treatment cessation. The *worst-case* behavior would be if he always reported "healthy". By Proposition 3, a doctor who assumes this worst-case behavior would believe that the patient is *sicker* than he actually is, which would lead to a delayed treatment cessation compared to a doctor who naïvely believed that the patient was honest. If the price of premature treatment cessation is greater than the price of delayed treatment cessation, then the former doctor would pay a lower Price of Naïveté. If patterns in symptom-reporting behavior are known (e.g., patients are likely to under-report early and be honest later), the Price of Naïveté could be further reduced by combining these two approaches and ignoring reported symptoms when they are likely to be under-reported and believing reported symptoms when they are likely to be reported honestly.

If patients exhibit one of several behavior types, we define the Multi-behavior Price of Naïveté

$$\texttt{MPoN}(\pi_0, p', [f_b]_{b\in\mathcal{B}}, \phi_0) := \sum_{b\in\mathcal{B}} \phi_0(b)\Big(v(\pi_0, \mu^*, f_b) - v(\pi_0, p', f_b)\Big). \tag{11}$$

Equation (11) measures the Price of Naïveté paid on average across all behavior types by a naïve policy p' compared to the BL-POMDP's optimal policy μ^*. This expression shows that *small reductions in the Price of Naïveté for the most probable behavior types can greatly reduce the Multi-behavior Price of Naïveté, even if the Price of Naïveté increases slightly for other behavior types.* Consider, for example, two naïve policies p' and p'', where p'' pays a higher Multi-behavior Price of Naïveté, i.e., $\texttt{MPoN}(\pi_0, p'', [f_b]_{b\in\mathcal{B}}, \phi_0) > \texttt{MPoN}(\pi_0, p', [f_b]_{b\in\mathcal{B}}, \phi_0)$. By rearranging terms, we have

$$\sum_{b\in\mathcal{B}'} \phi_0(b)\Big(v(\pi_0, p', f_b) - v(\pi_0, p'', f_b)\Big) > \sum_{b\notin\mathcal{B}'} \phi_0(b)\Big(v(\pi_0, p'', f_b) - v(\pi_0, p', f_b)\Big),$$

where $\mathcal{B}'$ is the set of behavior types for which p' pays a lower Price of Naïveté than p''. As such, p' need not pay a lower Price of Naïveté on *every* behavior type — just the ones that are most likely to appear. For example, if patients are likely to under-report symptoms to expedite treatment cessation, then strategies to reduce the Price of Naïveté for these patients, like a mandatory waiting period, can improve population-wide outcomes. Likewise, *the Multi-behavior Price of Naïveté could also be reduced by a policy p' over p'' if p' drastically reduces the Price of Naïveté for a few improbable behavior types without increasing the Price of Naïveté too much for the most likely behavior types.* For example, if a new policy reduces the Price of Naïveté (compared to an existing policy) for under-reporting patients but does not greatly affect outcomes for honest or over-reporting patients, then the new policy may reduce the Multi-behavior Price of Naïveté.

5 Return-to-play from Sports-related Concussion

Concussion, the most common type of traumatic brain injury, is a major public health issue (McCrory et al. 2017). In the short-term, concussion is associated with the alteration of neurologic function and a wide-ranging set of symptoms, including confusion and memory loss. Further, while the exact relationship is unclear, concussion may be associated with long-term consequences such as cognitive impairment, neurodegenerative disease, and depression (Kerr et al. 2014a, 2018). Improving concussion management is critical to improving patient health outcomes.

For SRC, a key component of the management protocol is determining when the patient may RTP, i.e., return to unrestricted participation in sport. During the RTP decision process, current guidelines recommend a multi-faceted approach which combines objective clinical measures (e.g., neurocognitive assessments) and self-reported symptom measures to estimate an athlete's health status. While a multidimensional approach to SRC assessment is best practice, the most widely used method in practice is based on self-reported symptoms since symptoms are typically the most indicative measure of SRC (Garcia et al. 2018). Yet, assessing SRC through self-reported symptoms invites influence from strategic symptom-reporting. For example, athletes may strategically under-report or over-report symptoms to expedite or delay RTP, among other reasons (Meier et al. 2015, Conway et al. 2018). Given the consequences associated with premature RTP (e.g., increased risk of injury (McCrea et al. 2020)) and delayed RTP (e.g., reduced health benefits from physical activity and exercise), understanding the role of symptom self-reporting behavior and its impact in the RTP decision process is critical to improving health outcomes.

In our numerical analysis, we apply the BL-POMDP to RTP from SRC for male collegiate football players. We focus on collegiate football since it has one of the greatest risks of injury among all collegiate sports and is a focal point for SRC in the media (Baugh and Kroshus 2016). We describe our BL-POMDP model for RTP from SRC in §5.1 and the RTP policies we evaluate in §5.2. We then analyze the optimal policies in §5.3, the effect of symptom-reporting behavior on each RTP policy in §5.4, and estimate the Price of Naïveté in §5.5. This analysis aims to (1) provide insights on how to optimize the timing of RTP for potentially strategic athletes and (2) quantify the cost of ignoring symptom-reporting behavior in RTP decisions.

5.1 BL-POMDP for RTP from SRC

We augment all BL-POMDP modeling components with a subgroup-specific variable $\theta \in \Theta = \{0, 1+\}$, where θ represents the athlete's number of previous concussions. In general, we could set Θ to account for several modifying factors such as sex, concussion history, and sport. Additionally, we model the time between decision epochs $t = 1, 2, ...$ as one day since daily or near-daily assessments may be used in collegiate sports at the discretion of the local medical staff (Broglio et al. 2017).

Patient's Health Dynamics. We model the set of health states as $\mathcal{H} = \{0, 1, 2\}$ where $h = 0, 1$ and 2 represent recovered, asymptomatic concussion, and symptomatic concussion, respectively. We assume that the state transition probabilities take the form $P_\theta = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon_1(\theta) & 1 - \epsilon_1(\theta) & 0 \\ 0 & \epsilon_2(\theta) & 1 - \epsilon_2(\theta) \end{bmatrix}$. Imposing this structure on P_θ implies that the athlete's cannot directly recover from concussion if they are still symptomatic. Furthermore, the concussion cannot worsen. Instead, it can either improve towards recovery or remain the same. This structure is consistent with most concussion recovery trajectories (Broglio et al. 2022) and satisfies Assumption 4 for P_θ.

The objective measures are given by $\mathcal{O} = \{30+, 29, 28, 27, 0\text{-}26\}$, where each $o \in \mathcal{O}$ represents a range of scores on the Standard Assessment of Concussion (SAC) — a neurocognitive exam used to measure impairment after SRC. The subjective measures are given by $\mathcal{S} = \{0, 1, 2\text{-}4, 5\text{-}13, 14\text{-}32, 33+\}$, where each $s \in \mathcal{S}$ represents total symptom severity scores on the Sport Concussion

Assessment Tool (SCAT) graded symptom checklist. Since there is no gold standard for SRC severity, we could not directly test the conditional independence between symptoms and objective measures (i.e., Assumption 1). However, we used the χ^2 tests for independence based on SRC assessments taken within 48 hours of the injury ($P < 0.001$), at the time at which the athlete began the RTP protocol ($P = 0.13$), and the time at which the athlete was permitted to RTP ($P = 0.20$), and we conclude that Assumption 1 may be reasonable since symptoms and objective observations only seem to be conditionally dependent during the acute stage (typically lasting 1-3 days). In light of these results, we jointly derived the HMM parameters P_θ, $Q_\theta^{\mathcal{O}}$, and $Q_\theta^{\mathcal{S}}$ by applying the Baum-Welch algorithm (Rabiner 1989) on data from the National Collegiate Athletic Association and Department of Defense (NCAA-DoD) Concussion Assessment, Research, and Education (CARE) Consortium (Broglio et al. 2017). This dataset combines SRC assessment data from 29 NCAA universities and military service academies throughout the United States. To our knowledge, this dataset is the largest available on concussion among collegiate athletes. While these data do not explicitly account for over-reporting or under-reporting behavior, the size of the dataset allows our models to extract meaningful information on average performance on the SCAT and SAC assessments across this athlete population. Accordingly, our models of symptom-reporting behavior can be interpreted as how an athlete may over or under-report symptoms *relative* to what a doctor expects to see. We validated our HMMs using held-out testing data (see EC.5.2 in the e-companion).

Rewards. The doctor's reward function incorporates daily physical and mental health-related quality of life (HRQoL) for the patient at each health state along with potential gains and losses in HRQoL from increased exercise and potential injury risks associated with RTP. Specifically, we set the doctor's reward function as

$$r_\theta^d(h,a) = \begin{cases} \zeta\Big(\mathtt{PCS}(h,\theta), \mathtt{MCS}(h,\theta)\Big) & a = 0 \\ \sum_{t=1}^{120} \mathbb{E}[\rho^{t-1} r^{\mathtt{MRP}}(\omega_t) | \omega_1 = h] & a = 1 \end{cases}, \tag{12}$$

where $\zeta(\cdot)$ is a function which maps physical and mental HRQoL scores to health utilities, $\mathtt{PCS}(h,\theta)$ and $\mathtt{MCS}(h,\theta)$ are physical and mental HRQoL composite scores, respectively, when the athlete's health state is h and demographic group is θ, and $\sum_{t=1}^{120} \mathbb{E}[\rho^{t-1} r^{\mathtt{MRP}}(\omega_t) | \omega_1 = h]$ is the athlete's expected 120-day total discounted reward when they RTP in health state h based on a post-RTP Markov Reward Process. This 120-day period captures a similiar duration considered in other post-RTP injury studies (Herman et al. 2017, Cross et al. 2016). This Markov Reward Process has state space $\Omega = \mathcal{H} \cup \{3\}$, where $\omega = 3$ is an absorbing state representing time-loss injury such as a repeat concussion or lower extremity injury (see EC.5.2 in the e-companion for details). We used HRQoL estimates from the sports medicine literature (McAllister et al. 2001, Weber et al. 2019, Cowee and Simon 2019) combined with a previously published health utility estimating function (Lawrence and Fleishman 2004) to estimate these rewards.

Symptom-reporting behavior. In RTP from sports-related concussion, symptom-reporting behavior can be complex and depend on a wide range of factors, including the athlete's perception of his own concussion severity and risks related to premature RTP (Kerr et al. 2014b). As such, we model symptom-reporting behavior as the solution to an optimization problem. For any patient of fixed behavior type $b_0 = b \in \mathcal{B}$, we first define the patient's reward function as

$$r^p_\theta(h,a,b) = \begin{cases} r^d(h,0) & a=0 \\ \zeta\Big(\texttt{PCS}(\omega,\theta) + b\texttt{PCS}^+(\omega,\theta) - \texttt{PCS}^-(\omega,\theta), \texttt{MCS}(\omega,\theta) + b\texttt{MCS}^+(\omega,\theta) - \texttt{MCS}^-(\omega,\theta)\Big) & a=1 \end{cases}, \tag{13}$$

for all $h \in \mathcal{H}$, $a \in \mathcal{A}$, and $b \in \mathcal{B}$, where $\text{PCS}^+(h,\theta)$ and $\text{MCS}^+(h,\theta)$ (resp., $\text{PCS}^-(h,\theta)$ and $\text{MCS}^-(h,\theta)$) are the post-RTP boost (resp., decline) in physical and mental HRQoL composite scores due to increased exercise and activity levels. Note that in (13), $r^p(h,a,b)$ is increasing in b, where b can be interpreted as a ratio indicating how much the athlete values RTP relative to the doctor. As such, an athlete with behavior type $b=1$ values RTP as much as the doctor (and hence reports symptoms honestly), $b<1$ values RTP *less* than the doctor (and hence, tends to over-report symptoms to prevent RTP), and $b>1$ values RTP *more* than the doctor (and hence, tends to under-report symptoms to achieve RTP), respectively. In our analysis, we consider a set of behavior types given by $\mathcal{B} = \{0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2\}$. We chose the minimum value of this set to be 0 because b is a ratio and we set the maximum to 2 because we found in our analysis that values of $b \geq 2$ yielded similar symptom-reporting behaviors.

Based on the reward function in (13), the patient's symptom-reporting strategy is determined in two steps. First, after observing symptom $s_t = s$ and recalling his prior health belief $\pi_{t-1} = \pi$, the patient determines the set of symptoms he can report which maximize his expected reward

$$\mathcal{S}^*(\pi,s) := \arg\max_{\bar{s}\in\mathcal{S}} \mathbb{E}[r^p_\theta(h,a,b)|\pi,s,\bar{s},b]. \tag{14}$$

In the right-hand side of (14), the expectation is taken over h and a since the patient is uncertain about his own health and the doctor's ensuing action. That is,

$$\mathbb{E}[r^p_\theta(h,a,b)|\pi,s,\bar{s},b] = \sum_{h\in\mathcal{H}} \Big(r^p_\theta(h,1,b)\mathbb{P}(a_t=1|\pi,s,\bar{s}) + r^p_\theta(h,0,b)\mathbb{P}(a_t=0|\pi,s,\bar{s}) \Big) \mathbb{P}(h_t=h|\pi,s), \tag{15}$$

where

$$\mathbb{P}(h_t=h|\pi,s) = \frac{Q^{\mathcal{S}}(s|h)\sum_{h'\in\mathcal{H}} P(h|h')\pi(h')}{\sum_{h\in\mathcal{H}} Q^{\mathcal{S}}(s|h)\sum_{h'\in\mathcal{H}} P(h|h')\pi(h')} \tag{16}$$

$$\mathbb{P}(a_t=1|\pi,s,\bar{s}) = \sum_{o\in\mathcal{O}} \mu(U(\pi,o,\bar{s})) \sum_{h\in\mathcal{H}} Q^{\mathcal{O}}(o|h)\mathbb{P}(h_t=h|\pi,s) \tag{17}$$

$$\mathbb{P}(a_t=0|\pi,s,\bar{s}) = 1 - \mathbb{P}(a_t=1|\pi,s,\bar{s}). \tag{18}$$

In (15), the components (16), (17), and (18) are interpreted as the patient's posterior health belief state after observing symptom $s_t = s$, the probability with which the patient expects the doctor to allow RTP after he reports $\bar{s}$, and the probability with which the patient expects the doctor to wait one more period after he reports $\bar{s}$, respectively. After determining $\mathcal{S}^*(\pi, s)$, we model the patient's symptom-reporting behavior according to

$$f_{b,\theta}(\bar{s}|s,\pi) = \begin{cases} \frac{\sum_{h\in\mathcal{H}} Q^{\mathcal{S}}(\bar{s}|h)\mathbb{P}(h_t=h|\pi,s)}{\sum_{\bar{s}'\geq s}\sum_{h\in\mathcal{H}} Q^{\mathcal{S}}(\bar{s}'|h)\mathbb{P}(h_t=h|\pi,s)} & \bar{s}\in\mathcal{S}^*, \bar{s}\geq s, b<1 \\ \frac{\sum_{h\in\mathcal{H}} Q^{\mathcal{S}}(\bar{s}|h)\mathbb{P}(h_t=h|\pi,s)}{\sum_{\bar{s}'\leq s}\sum_{h\in\mathcal{H}} Q^{\mathcal{S}}(\bar{s}'|h)\mathbb{P}(h_t=h|\pi,s)} & \bar{s}\in\mathcal{S}^*, \bar{s}\leq s, b>1 \\ 1 & s=\bar{s}, b=1 \\ 0 & \text{otherwise.} \end{cases} \qquad (19)$$

By design, $b = 1$ indicates honest symptom-reporting while $b < 1$ indicates over-reporting (i.e., $\bar{s} \geq s$) and $b > 1$ indicates under-reporting (i.e., $\bar{s} \leq s$). That is, in (19), the first (resp., second) case reflects the conditional probability that the doctor would expect to see $s_t = \bar{s}$ conditioned on the patient's posterior health belief after seeing his own symptoms (16) and that s_t is being over-reported (resp., under-reported).

Remark 2. Both (17) and (18) depend on the patient's knowledge of the doctor's policy and belief updating function. However, both the doctor's policy and the belief updating function require the patient's symptom-reporting behavior to be computed. To retain computational tractability, we approximate $\mu(U(\pi, o, \bar{s})) \approx \mathbb{1}\{\dot{U}(\pi, o, \bar{s})^\top r_L^p \geq 0\}$, where r_L^p is a vector with components $r_L^p(h) = r_\theta^p(h, 1, 1) - r_\theta^p(h, 0, 1)$ for all $h \in \mathcal{H}$. That is, we assume that the patient estimates the doctor's action based on a one-step look-ahead policy with an honest patient's reward function. We also assume that the patient believes that the doctor's belief updating occurs according to $\dot{U}(\cdot)$. In practice, athletes can develop a sense of how doctors form their decision thresholds either through prior experience or interactions with teammates who have experienced the RTP process (Moreau et al. 2014). In EC.6.6 in the e-companion, we relax this approximation and design an Oracle patient who knows the doctor's policy μ and belief update function $U(\cdot)$ exactly.

5.2 RTP Policies and Evaluation

For the BL-POMDP, we set $\pi_0 = e_2$, i.e., all athletes are assumed to have symptomatic concussions at the start of the RTP decision process, and $\phi_0 = [0.025\ 0.025\ 0.05\ 0.15\ 0.5\ 0.15\ 0.05\ 0.025\ 0.025]^\top$, i.e., the doctor assumes that the patient is likely to be honest. The BL-POMDP is not sensitive to ϕ_0 in this case study (see EC.6.4 in the e-companion). We compared the BL-POMDP to an Oracle POMDP which knows the patient's true symptom observations, a POMDP which does not account for symptom-reporting behavior, and two practice-based policies which assume that the athlete reports symptoms honestly. The first practice-based policy, denoted *Myopic*, is a simple myopic policy which permits RTP once SAC and symptom scores are normal (Broglio et al. 2009), i.e., in period t' where $t' = \inf\{t : o_t = \bar{s}_t = 0\}$. The second policy, denoted *CurrPrac*, more closely mimics current practice by permitting RTP 7 days after the athlete presents normal SAC and symptom scores at least once (McCrea et al. 2020), i.e., in period $t'+7$ where period $t' = \sup\{\inf_t\{t : \bar{s}_t = 0\}, \inf\{t : o_t = o\}\}$.

We evaluated the performance by each RTP policy using agent-based simulation with 1,000 replications for each $\theta \in \Theta$ and $b_0 \in \mathcal{B}$. In each iteration, we assumed that the athlete initially had acute concussion, i.e., $h_0 = 2$. Although we formulated the BL-POMDP for an infinite horizon, we evaluated the policy over 90 days since the majority of RTP decisions are made well within that timeframe (McCrea et al. 2020). All problem data follow the models and parameters defined in §5.1. Once the patient is permitted to RTP, we simulated the post-RTP Markov Reward Process which is briefly described in §5.1 and whose parameters are detailed in EC.5.1 in the e-companion. We remark that the total number of days simulated in each replication differs depending on the timing of RTP. However, our reward structure ensures that premature RTP is penalized by a reduced lump sum RTP reward while delayed RTP is penalized by reward discounting in each period.

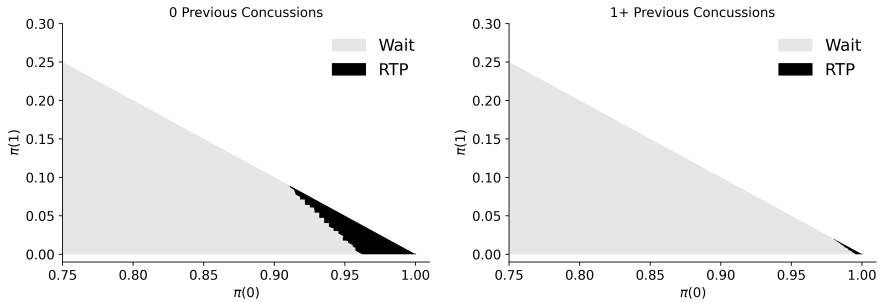
5.3 Analysis of BA-POMDP RTP Policies

In this section, we analyze how concussion history and symptom-reporting behavior modify optimal RTP policies, providing some insights on how optimal RTP policies might change for athletes not included in our study. To construct the BL-POMDP optimal policy, we solved a BA-POMDP for each $\theta \in \Theta$ and $b_0 \in \mathcal{B}$ using a grid-based approximation (see EC.4 in the e-companion) with 400, 150, and 550 grid points in regions $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$, respectively.

The BA-POMDP policies for honest athletes are illustrated in Figure 3. By Theorem 1, these policies are equivalent to the optimal POMDP policies. The optimal policy is more conservative for athletes with 1+ (vs. 0) previous concussions. This difference is likely due to the increased post-RTP injury risk incurred by athletes with a greater concussion history. Extrapolating beyond our analysis, we expect that athletes at higher risk of injury after RTP require a more conservative RTP policy than those with lower risk (e.g., football vs. golf). Furthermore, we expect that if the risk of injury post-RTP is sufficiently high, then an optimal policy would never permit RTP.

We present the BA-POMDP policies for each behavior type $b_0 \neq 1$ in EC.6 in the e-companion. For athletes with 0 previous concussions, RTP policies are slightly more aggressive for athletes who greatly over-report symptoms (i.e., $b_1 = 0$) and for athletes who under-report symptoms (i.e., $b_0 > 1$) compared to honest athletes. However, there are only minor differences between the BA-POMDP policy for honest athletes and slightly over-reporting athletes (i.e., $b_0 \in \{0.25, 0.5, 0.75\}$). For athletes with 1+ previous concussions, we also find few differences between the BA-POMDP

Figure 3 BA-POMDP policies for honest athletes. Black indicates that it is optimal to RTP and gray indicates that it is optimal to wait. For belief states not shown, it is optimal to wait.



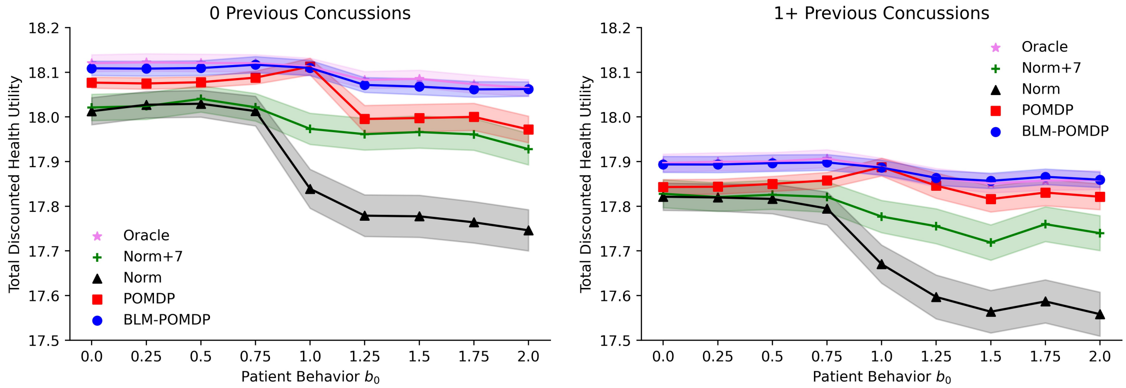
policies for honest athletes and all other behavior types. These results suggest several insights. First, adaptations to symptom-reporting behavior in the health state belief update will sufficiently modify RTP decisions for "less extreme" symptom-reporting behaviors or conservative RTP policies, requiring fewer compensations within the RTP policy itself. Second, for "extreme" symptom-reporting behaviors, the RTP decision thresholds may require slight adjustments to account for conservative belief-updating when patients strategically report symptoms. Finally, when the BA-POMDP for honest patients is similar to patients who report symptoms strategically, the RTP policy cannot become any more conservative than it already is for honest patients (see Theorem 2). Hence, potential differences in performance between the BL-POMDP and POMDP are likely driven by differences in belief updating and not the RTP policy (see Figure EC.6 in the e-companion).

5.4 Effect of Symptom-reporting Behavior on RTP Policy Performance

To analyze how symptom-reporting behavior affects each RTP policy, we first estimate the optimality gap associated with the approximate BL-POMDP policy (see EC.6.3 in the e-companion). These gaps are no larger than 0.38% across all policies, behavior types, and concussion histories. Given the magnitude of these estimates, the BL-POMDP policies are likely close to optimal.

We next present the expected total discounted health utilities achieved by each policy in Figure 4. The BL-POMDP outperforms all benchmark policies except for the Oracle POMDP and the POMDP for honest patients. Notably, the BL-POMDP performs similarly to the Oracle POMDP across all behavior types, indicating that the BL-POMDP — which must infer the patient's true symptoms — performs nearly as well as a POMDP that *actually knows* the patient's true symptom observations. Additionally, the POMDP performs slightly better among honest athletes, although the difference may be negligible. Nonetheless, this finding implies that the simpler POMDP suffices for this group. Among strategic athletes, the BL-POMDP most outperforms the practice-based RTP policies when athletes are under-reporting symptoms (i.e., $b_0 > 1$). This finding also holds when comparing the BL-POMDP to the POMDP for athletes with 0 previous concussions, but less

Figure 4 Total discounted health utility-adjusted days for each RTP policy and behavior type $b_0 \in \mathcal{B}$. Markers indicate mean values and shaded areas indicate 95% confidence intervals of the mean.



so for athletes with 1+ previous concussions. We suspect this difference may arise because of how conservative the POMDP policy is for athletes with 1+ previous concussions (see Figure 3).

We also investigated the timing of RTP decisions relative to the timing of recovery from concussion (see EC.6.5 in the e-companion). When patients are honest, both the BL-POMDP and POMDP are take longer to RTP than the practice-based policies. Across all behavior types, the time to RTP is consistent for the BL-POMDP, with only a small proportion of athletes permitted to RTP before a full recovery. In alignment with Proposition 3, the POMDP and practice-based policies permit RTP sooner (resp., later) for under-reporting (resp., over-reporting) athletes compared to honest athletes. These results suggest that failing to account for symptom-reporting behavior permits the athlete to manipulate the time to RTP alignment with his own objectives. In contrast, the BL-POMDP is more consistent across behavior types and generally adjusts in the opposite way — it becomes more aggressive when athletes over-report symptoms and more conservative when athletes under-report symptoms compared to the POMDP, Myopic, and CurrPrac policies.

Unfortunately, the timing of full recovery from SRC is unobservable in practice. To develop practically implementable insights, we also compared the BL-POMDP and CurrPrac policies with regard to the *symptom-free waiting period* (SFWP) (see EC.6.5.2 in the e-companion). The SFWP is defined by the length of time between RTP and the first period in which the athlete presents normal symptoms and SAC measurements at least once, i.e., $t_{\texttt{SFWP}} = \sup\{\inf_t\{t : \bar{s}_t = 0\}, \inf\{t : o_t = o\}\}$. Recall that CurrPrac applies a SFWP of 7 days. Generally speaking, the BL-POMDP takes a shorter (resp., longer) SFWP when athletes are over-reporting (resp., under-reporting) symptoms. Compared to CurrPrac, the BL-POMDP generally applies a SFWP longer than 7 days, except when $\theta = 0$ and the athlete is over-reporting symptoms. Furthermore, the BL-POMDP takes a longer SFWP for athletes with 1+ (vs. 0) previous concussions. These results suggest that (1) clinicians can improve collegiate football players' health outcomes by considering a SFWP longer than 7 days and (2) clinicians should tailor the SFWP based on the athlete's post-RTP injury risk.

5.5 Estimating and Contextualizing the Price of Naïveté

We now aim to estimate the Price of Naïveté in RTP from SRC by adapting (11) to our numerical analysis. Since the true distribution across behavior types ϕ_0 is difficult to estimate in practice, we conservatively estimate lower (`MPoN-LB`) and upper (`MPoN-UB`) bounds on (11), respectively, by

$$\texttt{MPoN-LB}(\beta_0, \beta_1, \pi_0, [f_b]_{b\in\mathcal{B}}, p') := \min_{\phi_0 \in \Phi'(\beta_0,\beta_1)} \texttt{MPoN}(\pi_0, p', [f_b]_{b\in\mathcal{B}}, \phi_0), \tag{20}$$

$$\texttt{MPoN-UB}(\beta_0, \beta_1, \pi_0, [f_b]_{b\in\mathcal{B}}, p') := \max_{\phi_0 \in \Phi'(\beta_0,\beta_1)} \texttt{MPoN}(\pi_0, p', [f_b]_{b\in\mathcal{B}}, \phi_0), \tag{21}$$

where the set $\Phi'(\beta_0, \beta_1)$ is defined by

$$\Phi'(\beta_0, \beta_1) := \left\{ \phi \in \Phi : \begin{array}{cc} \phi(0) = \beta_0 & \sum_{b>0} \phi(b) = \beta_1 \\ \phi(b+1) \geq \phi(b) \text{ for all } b > 0 & \phi(b) \geq \phi(b-1) \text{ for all } b < 0 \end{array} \right\}. \tag{22}$$

In our computation of the bounds in (20) and (21), we estimate `MPoN-LB` and `MPoN-UB` by using our numerical estimates for the value of all benchmark policies (except the Oracle POMDP) in §5.4 instead of the true value. For example, we replace $v(\pi_0, \mu^*, f_b)$ for the BL-POMDP and $v(\pi_0, \dot{\mu}^*, f_b)$ for the POMDP with $\hat{v}(\pi_0, \mu^*, f_b)$ and $\hat{v}(\pi_0, \dot{\mu}^*, f_b)$, respectively, for all $b \in \mathcal{B}$. As such, the optimization problem in (20) and (21) reduces to a linear program. In (22), β_0 and β_1 denote the proportion of honest and under-reporting athletes, respectively. For example, $\Phi'(0.5, 0.4)$ defines the set of all distributions over symptom-reporting behavior when 50% of athletes are honest, 40% under-report symptoms, and 10% over-report symptoms. Finally, notice that $\Phi'(\beta_0, \beta_1)$ contains only distributions with monotone tails, i.e., the proportion of extremely strategic athletes is no greater than the proportion who subtly over- or under-report symptoms.

Bounds on the Price of Naïveté over the POMDP are shown in Figure 5. The Price of Naïveté is relatively low when the proportion of honest patients is high and there are very few under-reporting athletes (e.g., $\beta_0 = 0.75$, $\beta_1 = 0$ for $\theta = 0$). In some cases, the Price of Naïveté is slightly negative (difference is $<$0.00017%) when nearly all athletes are honest and very few athletes under-report (e.g., $\beta_0 = 0.925, \beta_1 = 0.05$ for $\theta = 0$). We suspect that the POMDP slightly outperforms the BL-POMDP in these cases because most athletes are honest, and the BL-POMDP has to learn that the athlete is honest whereas the POMDP assumes so. However, such settings are unrealistic given that an estimated 50%-60% of athletes under-report symptoms (Meier et al. 2015, Conway et al. 2018). To this end, both `MPoN-LB` and `MPoN-UB` for the POMDP are increasing as the proportion of under-reporting athletes increases, implying that the Price of Naïveté is greatest among athletes who under-report (vs. over-report) symptoms. These trends also hold for the Myopic and CurrPrac policies. Furthermore, $\texttt{MPoN-LB}(\beta_0, \beta_1, \cdot) > 0$ for all β_0, β_1 except when $\beta_0 \geq 0.925, \beta_1 \leq 0.25$ for $\theta = 0$ and $\beta_0 \geq 0.95, \beta_1 \leq 0.05$ for $\theta = 1+$. This finding implies that there is significant benefit in incorporating patient-behavior and tailoring RTP policies over current practice.

Figure 5 Lower (`MPoN-LB`) and upper (`MPoN-UB`) bounds on the Price of Naïveté for the POMDP.

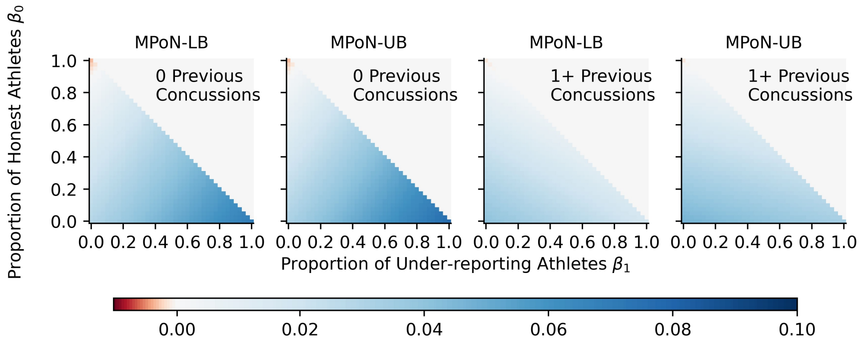


Table 1 BL-POMDP reduction in PoPRTP and gain in THAEs over benchmark policies

	$\theta = 0$				$\theta = 1+$			
	PoPRTP		**THAEs**		**PoPRTP**		**THAEs**	
Policy	*Reduction*	*Relative Risk*	*Gain*	*% Improvement*	*Reduction*	*Relative Risk*	*Gain*	*% Improvement*
POMDP	0.17-0.25	1.95-2.34	1.21-2.23	0.03-0.05	0.14-0.17	2.67-2.96	0.79-2.82	0.02-0.08
Myopic	0.59-0.68	4.30-4.79	4.39-5.05	0.11-0.13	0.61-0.70	8.17-10.02	3.88-4.59	0.11-0.13
CurrPrac	0.44-0.47	3.26-3.63	2.81-3.36	0.07-0.08	0.46-0.49	6.19-7.36	2.72-3.34	0.08-0.09

PoPRTP: Probability of Premature Return-to-Play; THAEs: Total Health-adjusted Athletic Exposures; Ranges computed over distributions $\phi_{LB}(\beta_0, \beta_1, \pi_0, [f_b]_{b\in\mathcal{B}}, p')$ and $\phi_{UB}(\beta_0, \beta_1, \pi_0, [f_b]_{b\in\mathcal{B}}, p')$ with $\beta_0 \in [0.3, 0.5]$ and $\beta_1 \in [0.5, 0.6]$. Relative Risk = $\text{PoPRTP}_{p'}/\text{PoPRTP}_{\text{BL-POMDP}}$; % Improvement = $100\times(\text{THAEs}_{\text{BL-POMDP}}/\text{THAEs}_{p'}$ - 1).

To contextualize the Price of Naïveté for clinicians, we quantify the Probability of Premature RTP (PoPRTP) and Total Health-adjusted Athletic Exposures (THAEs). Minimizing PoPRTP, i.e., $\mathbb{P}(a_t = 1, h_t > 0)$, is critical given the purported consequences of premature RTP, e.g., late-life neurocognitive impairment (Cantu and Gean 2010). THAEs measure health-weighted participation in sport after RTP and are calculated as 0 for athletes who do not RTP and as $\mathbb{E}[\sum_{t=1}^{120} r_\theta^\Omega(\omega_t)\mathbb{1}\{\omega_t \neq 3\}]$ using the post-RTP Markov Reward Process (see EC.5.1 in the e-companion) for athletes who do RTP. For the POMDP, Myopic, and CurrPrac policies, we evaluated PoPRTP and THAEs (see Table 1) under distributions which minimize `MPoN-LB` and maximize `MPoN-UB`. We assume $\beta_0 \in [0.3, 0.5]$ and $\beta_1 \in [0.5, 0.6]$, reflecting the estimated 50%-60% of athletes who under-report symptoms (Meier et al. 2015, Conway et al. 2018). Accounting for patient behavior can greatly reduce the PoPRTP and increase the THAEs; compared to CurrPrac, the BL-POMDP reduces the PoPRTP by at least 44% and increases post-RTP participation by up to 3.36 THAEs. These improvements can greatly impact athletes with SRC, with the potential benefits extending beyond what is measurable in this analysis (e.g., reduced risk of catastrophic events from premature RTP).

6 Conclusion

The potential for strategically reported PROs in patient-centered care can make health assessments and treatment decisions challenging. In this research, we formulated and analyzed the BA-POMDP and BL-POMDP models — novel stochastic dynamic programming models which incorporate uncertainty around the patient's health and PRO-reporting behavior. We applied the BL-POMDP to optimize the timing of RTP from concussion by leveraging data from over 29 sites to parameterize and validate our model. The BL-POMDP outperformed several benchmark RTP policies based on several markers of health outcomes — especially in the presence of symptom under-reporting.

In our application to RTP from concussion, the BL-POMDP generally takes a longer SFWP than current practice, implying that clinicians may be able to improve health outcomes for collegiate football players by waiting longer than 7 days after normal symptom and neurocognitive presentation before RTP. Moreover, RTP decisions should be tailored by athlete specific features, e.g., taking a more conservative approach for athletes with 1+ vs. 0 previous concussions. These differences in RTP policy by post-RTP risk can be implemented, in part, by tailoring the SFWP. Finally, the Price of Naïveté is greatest for athletes who are under-reporting symptoms. Specifically, the health-related costs associated with premature RTP are far greater than delayed RTP. Given the high rate of symptom under-reporting, clinicians should cautiously interpret signals of rapid recovery in the RTP decision process, while learning symptom-reporting behavior over time and adjusting interpretation of reported symptoms accordingly.

Our findings lead to several insights for doctors who aim to make treatment decisions under potentially strategic symptom-reporting behavior. In general, our main insight is that *using knowledge about patient behavior to modify interpretation of PROs can be at least as important as optimizing the treatment decision policy.* In our analysis of the Price of Naïveté, we found that the Price of Naïveté can still be paid if patient behavior is neglected in health state belief updates. Further, knowing *when* patients are likely to report symptoms honestly can help to mitigate the Price of Naïveté. Our numerical study supports this finding; specifically, we found that even when the BL-POMDP and POMDP policies were similar, differences in belief updates led to large differences in the timing of the RTP decision. Hence, understanding the drivers of strategic behavior and how this behavior manifests in PROs is critical to improving clinical practice.

This work can be extended in several ways. First, future research can incorporate larger action spaces in the BL-POMDP framework, which may be more appropriate for other medical decision-making contexts such as hypertension treatment planning. Second, our model of patient behavior was influenced by the types of behavior which are reasonably exhibited in medical decision-making contexts. Future research can also consider alternative models of patient behavior, including those for which our modeling assumptions no longer hold. For example, Assumption 2 can be relaxed by considering a setting wherein the patient's recall of his own symptoms includes the k previous periods instead of just the current period. Finally, our models assumed that the parameters of the BA-POMDP and BL-POMDP models are accurately estimated. However, parameter misestimation could have potentially significant implications on how the doctor interprets the patient's reported symptoms. For example, if symptom-reporting severity is systematically under-estimated in our concussion application, then the doctor might be more prone to allowing premature RTP. Future research can consider extensions that are robust to parameter uncertainty in the transition probabilities, observation probabilities, and estimates of symptom-reporting strategies.

As medicine continues shifting toward patient-centered care, the importance of incorporating PROs and understanding patient behavior in medical decision-making will continue rise. Our research provides a modeling framework and detailed numerical analysis which illustrate the importance of adaptively learning and accounting for a patient's PRO-reporting behavior throughout a long treatment planning process. By understanding patients' objectives and expectations, doctors can better account for individual differences in patient behavior and tailor their treatment decisions accordingly, ultimately improving each patient's health outcomes.

Acknowledgments

The authors would like to thank the departmental editor Sergei Savin, the anonymous senior editor, and the anonymous referees for their constructive and detailed comments which have helped significantly improve the content and exposition of this manuscript. The authors would also like to thank Dr. Michelle Weber Rawlins, who kindly provided SF-12 values for our case study.

The material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1256260. We acknowledge the research team members and clinical athletic trainers who assisted in the data collection for this project. This publication was made possible, in part, with support from the Grand Alliance Concussion Assessment, Research and Education (CARE) Consortium, funded, in part, by the National College Athletic Association (NCAA) and the Department of Defense (DOD). The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office. This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Psychological Health and Traumatic Brain Injury Program under Award No. W81XWH-14-2-0151. Opinions, interpretations, conclusions and recommendations are those of the author(s) and are not necessarily endorsed by the Department of Defense (DHP funds).

Contributing CARE Consortium Investigators are listed alphabetically by institution: April Marie (Reed) Hoy, MS, ATC (Azusa Pacific University); Joseph B Hazzard Jr, EdD, ATC (Bloomsburg University); Louise A Kelly, PhD (California Lutheran University); Justus Ortega, PhD (Humboldt State University); Nicholas Port, PhD (Indiana University); Margot Putukian, MD (Princeton University); Jonathan C Jackson, MD (US Air Force Academy); Carlos Estevez, DPT

(US Coast Guard Academy); Kenneth Cameron, MPH, PhD, ATC (US Military Academy); Adam J Susmarski, DO (US Naval Academy); Christopher C Giza, MD (University of California-Los Angeles); Holly J Benjamin MD (University of Chicago); Thomas Buckley, EdD, ATC, and Thomas W Kaminski, PhD, ATC (University of Delaware); James R Clugston, MD, MS (University of Florida); Julianne D Schmidt, PhD, ATC (University of Georgia); Luis A Feigenbaum, PT, DPT, ATC, LAT (University of Miami); James T Eckner, MD, MS (University of Michigan); Jason P Mihalik, PhD, CAT(C), ATC (University of North Carolina at Chapel Hill); Jessica Miles, PhD, ATC (University of North Georgia); Scott Anderson, ATC (University of Oklahoma); Christina L Master, MD (University of Pennsylvania); Micky Collins, PhD, and Anthony P Kontos, PhD (University of Pittsburgh Medical Center); Sara PD Chrisman, MD, MPH (University of Washington); M Alison Brooks, MD, MPH (University of Wisconsin); Stefan Duma, PhD and Steven Rowson, PhD (Virginia Tech); Christopher M Miles, MD, Christopher T Whitlow, MD, PhD, and Laura J Lintner, DO (Wake Forest University).

References

Arends, P, R J S Wagner, M Mrazik. 2019. SCAT Symptom Recall Bias in Concussed Athletes. *Archives of Clinical Neuropsychology*, 34 (5), 733-733. doi:10.1093/arclin/acz026.03.

Åström, K. J. 1965. Optimal control of Markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10 (1), 174-205. doi:10.1016/0022-247X(65)90154-X.

Aswani, Anil, Philip Kaminsky, Yonatan Mintz, Elena Flowers, Yoshimi Fukuoka. 2018. Behavioral modeling in weight loss interventions. *European Journal of Operational Research*, 0 1-15. doi:10.1016/j.ejor.2018.07.011.

Ayer, Turgay, Oguzhan Alagoz, Natasha K. Stout, Elizabeth S. Burnside. 2016. Heterogeneity in Women's Adherence and Its Role in Optimal Breast Cancer Screening Policies. *Management Science*, 62 (5), 1339-1362. doi:10.1287/mnsc.2015.2180.

Barnett, Christine L, Scott A Tomlins, Daniel J Underwood, John T Wei, Todd M Morgan, James E Montie, Brian T Denton. 2017. Two-Stage Biomarker Protocols for Improving the Precision of Early Detection of Prostate Cancer. *Medical Decision Making*, 37 (7), 815-826. doi:10.1177/0272989X17696996.

Bass, Christopher, Peter Halligan. 2014. Factitious disorders and malingering: Challenges for clinical assessment and management. *The Lancet*, 383 (9926), 1422-1432. doi:10.1016/S0140-6736(13)62186-8.

Baugh, Christine M, Emily Kroshus. 2016. Concussion management in US college football: Progress and pitfalls. *Concussion*, 1 (1), cnc.15.6. doi:10.2217/cnc.15.6.

Blackwell, David. 1953. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, 24 (2), 265-272.

Broglio, Steven P, Stephen N Macciocchi, Michael S Ferrara. 2009. Neurocognitive performance of concussed athletes when symptom free. *Journal of athletic training*, 42 (4), 504-8. doi:10.1016/S0162-0908(09)79463-2.

Broglio, Steven P, Thomas McAllister, Barry P. Katz, Michelle LaPradd, Wenxian Zhou, Michael A. McCrea, April Hoy, Joseph B. Hazzard, Louise A. Kelly, John DiFiori, Justus D. Ortega, Nicholas Port, Margot Putukian, Dianne Langford, Jane McDevitt, Darren Campbell, Jonathan C. Jackson, Gerald McGinty, Carlos Estevez, Kenneth L. Cameron, Megan N. Houston, Steven J. Svoboda, Adam James Susmarski, Chris Giza, Holly J. Benjamin, Thomas W. Kaminski, Thomas Buckley, James R. Clugston, Julianne Schmidt, Luis A. Feigenbaum, J. T. Eckner, Jason Mihalik, Jessica Dysart Miles, Scott Anderson, Kristy Arbogast, Christina L. Master, Anthony P. Kontos, Sara P. D. Chrisman, M. Alison Brooks, Steve Rowson, Stefan M. Duma, Chris Miles. 2022. The Natural History of Sport-Related Concussion in Collegiate Athletes: Findings from the NCAA-DoD CARE Consortium. *Sports Medicine*, 52 (2), 403-415. doi:10.1007/s40279-021-01541-7.

Broglio, Steven P, Michael McCrea, Thomas McAllister, Jaroslaw Harezlak, Barry Katz, Dallas Hack, Brian Hainline. 2017. A National Study on the Effects of Concussion in Collegiate Athletes and US Military Service Academy Members: The NCAA–DoD Concussion Assessment, Research and Education (CARE) Consortium Structure and Methods. *Sports Medicine*, 47 (7), 1437-1451. doi: 10.1007/s40279-017-0707-1.

Cantu, Robert C, Alisa D Gean. 2010. Second-Impact Syndrome and a Small Subdural Hematoma: An Uncommon Catastrophic Result of Repetitive Head Injury with a Characteristic Imaging Appearance. *Journal of Neurotrauma*, 27 (9), 1557-1564. doi:10.1089/neu.2010.1334.

Capan, Muge, Anahita Khojandi, Brian T. Denton, Kimberly D. Williams, Turgay Ayer, Jagpreet Chhatwal, Murat Kurt, Jennifer Mason Lobo, Mark S. Roberts, Greg Zaric, Shengfan Zhang, J. Sanford Schwartz. 2017. From Data to Improved Decisions: Operations Research in Healthcare Delivery. *Medical Decision Making*, 0272989X1770563doi:10.1177/0272989X17705636.

Conway, Fiona N., Marianne Domingues, Robert Monaco, Laura M. Lesnewich, Anne E. Ray, Brandon L. Alderman, Sabrina M. Todaro, Jennifer F. Buckman. 2018. Concussion Symptom Underreporting Among Incoming National Collegiate Athletic Association Division I College Athletes. *Clinical Journal of Sport Medicine*, 0 (0), 1. doi:10.1097/JSM.0000000000000557.

Cowee, Katlyn, Janet E. Simon. 2019. A history of previous severe injury and health-related quality of life among former collegiate athletes. *Journal of Athletic Training*, 54 (1), 64-69. doi: 10.4085/1062-6050-377-17.

Cross, Matthew, Simon Kemp, Andrew Smith, Grant Trewartha, Keith Stokes. 2016. Professional Rugby Union players have a 60% greater risk of time loss injury after concussion: A 2-season prospective study of clinical outcomes. *British Journal of Sports Medicine*, 50 (15), 926-931. doi: 10.1136/bjsports-2015-094982.

Dehon, Erin, Nicole Weiss, Jonathan Jones, Whitney Faulconer, Elizabeth Hinton, Sarah Sterling. 2017. A Systematic Review of the Impact of Physician Implicit Racial Bias on Clinical Decision Making. *Academic Emergency Medicine*, 24 (8), 895-904. doi:10.1111/acem.13214.

Doshi, Prashant, Piotr J. Gmytrasiewicz. 2006. On the difficulty of achieving equilibrium in interactive POMDPs. *9th International Symposium on Artificial Intelligence and Mathematics, ISAIM 2006*, 1131-1136.

Epstein, Ronald M, Richard L Street. 2011. The Values and Value of Patient-Centered Care. *The Annals of Family Medicine*, 9 (2), 100-103. doi:10.1370/afm.1239.

Garcia, Gian-Gabriel P., Steven P. Broglio, Mariel S. Lavieri, Michael McCrea, Thomas McAllister. 2018. Quantifying the Value of Multidimensional Assessment Models for Acute Concussion: An Analysis of Data from the NCAA-DoD Care Consortium. *Sports Medicine*, 48 (7), 1739-1749. doi:10.1007/s40279-018-0880-x.

Garcia, Gian-Gabriel P., Lauren L. Czerniak, Mariel S. Lavieri, Spencer W. Liebel, Michael A. McCrea, Thomas W. McAllister, Paul F. Pasquina, Steven P. Broglio. 2022. Simulation-Optimization to Distinguish Optimal Symptom Free Waiting Period for Return-to-Play Decisions in Sport-Related Concussion. *2022 Winter Simulation Conference (WSC)*. IEEE, Singapore, 1021-1032. doi:10.1109/WSC57314.2022.10015285.

Garcia, Gian-Gabriel P., Mariel S. Lavieri, Ruiwei Jiang, Thomas W. McAllister, Michael A. McCrea, Steven P. Broglio. 2019. A Data-Driven Approach to Unlikely, Possible, Probable, and Definite Acute Concussion Assessment. *Journal of Neurotrauma*, 36 (10), 1571-1583. doi:10.1089/neu.2018.6098.

Garcia, Gian-Gabriel P., Mariel S. Lavieri, Ruiwei Jiang, Michael A. McCrea, Thomas W. McAllister, Steven P. Broglio. 2020a. Data-driven stochastic optimization approaches to determine decision thresholds for risk estimation models. *IISE Transactions*, 52 (10), 1098-1121. doi:10.1080/24725854.2020.1725254.

Garcia, Gian-Gabriel P., Jing Yang, Mariel S. Lavieri, Thomas W. McAllister, Michael A. McCrea, Steven P. Broglio. 2020b. Optimizing Components of the Sport Concussion Assessment Tool for Acute Concussion Assessment. *Neurosurgery*, 87 (5), 971-981. doi:10.1093/neuros/nyaa150.

Herman, Daniel C., Debi Jones, Ashley Harrison, Michael Moser, Susan Tillman, Kevin Farmer, Anthony Pass, James R. Clugston, Jorge Hernandez, Terese L. Chmielewski. 2017. Concussion May Increase the Risk of Subsequent Lower Extremity Musculoskeletal Injury in Collegiate Athletes. *Sports Medicine*, 47 (5), 1003-1010. doi:10.1007/s40279-016-0607-9.

Kerr, Zachary Y, Kelly R Evenson, Wayne D Rosamond, Jason P Mihalik, Kevin M Guskiewicz, Stephen W Marshall. 2014a. Association between concussion and mental health in former collegiate athletes. *Injury Epidemiology*, 1 (1), 28. doi:10.1186/s40621-014-0028-x.

Kerr, Zachary Y., Jason P. Mihalik, Kevin M. Guskiewicz, Wayne D. Rosamond, Kelly R. Evenson, Stephen W. Marshall. 2015. Agreement Between Athlete-Recalled and Clinically Documented Concussion Histories in Former Collegiate Athletes. *The American Journal of Sports Medicine*, 43 (3), 606-613. doi:10.1177/0363546514562180.

Kerr, Zachary Y, Johna K Register-Mihalik, Stephen W Marshall, Kelly R Evenson, Jason P Mihalik, Kevin M Guskiewicz. 2014b. Disclosure and non-disclosure of concussion and concussion symptoms in athletes: Review and application of the socio-ecological framework. *Brain injury*, 28 (8), 1009-1021. doi:10.3109/02699052.2014.904049.

Kerr, Zachary Y., Leah C. Thomas, Janet E. Simon, Michael McCrea, Kevin M. Guskiewicz. 2018. Association Between History of Multiple Concussions and Health Outcomes Among Former College Football Players: 15-Year Follow-up From the NCAA Concussion Study (1999-2001). *The American Journal of*

Sports Medicine, 46 (7), 1733-1741. doi:10.1177/0363546518765121.

Keskinocak, Pinar, Nicos Savva. 2020. A Review of the Healthcare-Management (Modeling) Literature Published in Manufacturing & Service Operations Management. *Manufacturing & Service Operations Management*, 22 (1), 59-72. doi:10.1287/msom.2019.0817.

Krishnamurthy, Vikram. 2012. Quickest detection pomdps with social learning: Interaction of local and global decision makers. *IEEE Transactions on Information Theory*, 58 (8), 5563-5587. doi:10.1109/TIT.2012.2201372.

Lawrence, William F., John A. Fleishman. 2004. Predicting EuroQoL EQ-5D Preference Scores from the SF-12 Health Survey in a Nationally Representative Sample. *Medical Decision Making*, 24 (2), 160-169. doi:10.1177/0272989X04264015.

Levy, Andrea Gurmankin, Aaron M. Scherer, Brian J. Zikmund-Fisher, Knoll Larkin, Geoffrey D. Barnes, Angela Fagerlin. 2018. Prevalence of and Factors Associated With Patient Nondisclosure of Medically Relevant Information to Clinicians. *JAMA Network Open*, 1 (7), e185293. doi:10.1001/jamanetworkopen.2018.5293.

Lobo, J. M., B. T. Denton, J. R. Wilson, N. D. Shah, S. A. Smith. 2017. Using claims data linked with electronic health records to monitor and improve adherence to medication. *IISE Transactions on Healthcare Systems Engineering*, 7 (4), 194-214. doi:10.1080/24725579.2017.1346728.

Lohr, Kathleen N., Bradley J. Zebrack. 2009. Using patient-reported outcomes in clinical practice: Challenges and opportunities. *Quality of Life Research*, 18 (1), 99-107. doi:10.1007/s11136-008-9413-7.

Mason, Jennifer E., Darin A. England, Brian T. Denton, Steven A. Smith, Murat Kurt, Nilay D. Shah. 2012. Optimizing Statin Treatment Decisions for Diabetes Patients in the Presence of Uncertain Future Adherence. *Medical Decision Making*, 32 (1), 154-166. doi:10.1177/0272989X11404076.

McAllister, David R., Ali R. Motamedi, Sharon L. Hame, Matthew S. Shapiro, Frederick J. Dorey. 2001. Quality of Life Assessment in Elite Collegiate Athletes. *The American Journal of Sports Medicine*, 29 (6), 806-810. doi:10.1177/03635465010290062201.

McCrea, Michael, Steven Broglio, Thomas McAllister, Wenxian Zhou, Shi Zhao, Barry Katz, Maria Kudela, Jaroslaw Harezlak, Lindsay Nelson, Timothy Meier, Stephen William Marshall, Kevin M. Guskiewicz. 2020. Return to play and risk of repeat concussion in collegiate football players: Comparative analysis from the NCAA Concussion Study (1999–2001) and CARE Consortium (2014–2017). *British Journal of Sports Medicine*, 54 (2), 102-109. doi:10.1136/bjsports-2019-100579.

McCrory, Paul, Willem Meeuwisse, Jiří Dvorak, Mark Aubry, Julian Bailes, Steven Broglio, Robert C Cantu, David Cassidy, Ruben J Echemendia, Rudy J Castellani, Gavin A Davis, Richard Ellenbogen, Carolyn Emery, Lars Engebretsen, Nina Feddermann-Demont, Christopher C Giza, Kevin M Guskiewicz, Stanley Herring, Grant L Iverson, Karen M Johnston, James Kissick, Jeffrey Kutcher, John J Leddy, David Maddocks, Michael Makdissi, Geoff T Manley, Michael McCrea, William P Meehan, Sinji Nagahiro, Jon Patricios, Margot Putukian, Kathryn J Schneider, Allen Sills, Charles H Tator, Michael Turner, Pieter E Vos. 2017. Consensus statement on concussion in sport—the 5 th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, 51 838-847. doi:10.1136/bjsports-2017-097699.

Meier, Timothy B., Bradley J. Brummel, Rashmi Singh, Christopher J. Nerio, David W. Polanski, Patrick S.F. Bellgowan. 2015. The underreporting of self-reported symptoms following sports-related concussion. *Journal of Science and Medicine in Sport*, 18 (5), 507-511. doi:10.1016/j.jsams.2014.07.008.

Mintz, Yonatan, Anil Aswani, Philip Kaminsky, Elena Flowers, Yoshimi Fukuoka. 2020. Nonstationary Bandits with Habituation and Recovery Dynamics. *Operations Research*, 68 (5), 1493-1516. doi: 10.1287/opre.2019.1918.

Moreau, Matthew S, Jody Langdon, Thomas a Buckley. 2014. The lived experience of an in-season concussion amongst NCAA Division I student-athletes. *International Journal of Exercise Science*, 7 (1), 62-74.

Perloff, Richard M., Bette Bonder, George Berlin Ray, Eileen Berlin Ray, Laura A. Siminoff. 2006. Doctor-Patient Communication, Cultural Competence, and Minority Health. *American Behavioral Scientist*, 49 (6), 835-852. doi:10.1177/0002764205283804.

Rabiner, L.R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77 (2), 257-286. doi:10.1109/5.18626.

Schell, Greggory J., Gian-Gabriel P. Garcia, Mariel S. Lavieri, Jeremy B. Sussman, Rodney A. Hayward. 2019. Optimal Coinsurance Rates for a Heterogeneous Population under Inequality and Resource Constraints. *IISE Transactions*, 51 (1), 74-91. doi:10.1080/24725854.2018.1499053.

Schottmüller, Christoph. 2013. Cost incentives for doctors: A double-edged sword. *European Economic Review*, 61 43-58. doi:10.1016/j.euroecorev.2013.03.001.

Shechter, Steven M., Matthew D. Bailey, Andrew J. Schaefer, Mark S. Roberts. 2008. The Optimal Time to Initiate HIV Therapy Under Ordered Health States. *Operations Research*, 56 (1), 20-33. doi: 10.1287/opre.1070.0480.

Suen, Sze-chuan, Diana Negoescu, Joel Goh. 2022. Design of Incentive Programs for Optimal Medication Adherence in the Presence of Observable Consumption. *Operations Research*, 70 (3), 1691-1716. doi:10.1287/opre.2021.2227.

Weber, M. L., R. C. Lynall, N. L. Hoffman, E. H. Miller, T. W. Kaminski, T. A. Buckley, H. J. Benjamin, C. M. Miles, C. T. Whitlow, L. Lintner, S. P. Broglio, M. McCrea, T. McAllister, J. D. Schmidt. 2019. Health-Related Quality of Life Following Concussion in Collegiate Student-Athletes With and Without Concussion History. *Annals of Biomedical Engineering*, 47 (10), 2136-2146. doi:10.1007/s10439-018-02151-7.

Yasuda, Masami. 1988. The optimal value of markov stopping problems with one-step look ahead policy. *Journal of Applied Probability*, 25 (3), 544-552. doi:10.2307/3213983.

Zhang, Hui, Christian Wernz, Danny R. Hughes. 2018. A Stochastic Game Analysis of Incentives and Behavioral Barriers in Chronic Disease Management. *Service Science*, 10 (3), 302-319. doi:10.1287/serv.2018.0211.