

SIMULATION-OPTIMIZATION TO DISTINGUISH OPTIMAL SYMPTOM FREE WAITING PERIOD FOR RETURN-TO-PLAY DECISIONS IN SPORT-RELATED CONCUSSION

Gian-Gabriel P. Garcia

H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Institute of Technology
765 Ferst Dr, Groseclose 416
Atlanta, GA 30332, USA

Lauren L. Czerniak
Mariel S. Lavieri

Department of Industrial and Operations Engineering
University of Michigan
1891 IOE Building 1205, Beal Ave
Ann Arbor, MI 48109, USA

Spencer W. Liebel

Department of Neurology
University of Utah School of Medicine
383 Colorow Dr, Room 352
Salt Lake City, UT 84108, USA

Michael A. McCrea

Department of Neurosurgery and Neurology
Medical College of Wisconsin
8701 Watertown Plank Rd
Milwaukee, WI 53226, USA

Thomas W. McAllister

Department of Psychiatry
Indiana University School of Medicine
340 West 10th St, Fairbanks Hall Suite 6200
Indianapolis, IN 46202, USA

Paul F. Pasquina

Department of Physical Medicine & Rehabilitation
Uniformed Services University of the Health Sciences
4301 Jones Bridge Rd
Bethesda, MD 20814, USA

Steven P. Broglio

Michigan Concussion Center
University of Michigan
830 N University Ave
Ann Arbor, MI 48109, USA

CARE Consortium Investigators

NCAA-DoD CARE Consortium
355 West 16th Street
Goodman Hall, Suite 4800
Indianapolis, IN 46202, USA

ABSTRACT

Approximately 1.6-3.8 million sport and recreation concussions occur annually. Yet, there is currently no universal agreement on when an athlete should be permitted to unrestricted play after being diagnosed with a concussion. Simulation-optimization provides a tractable method to optimize the length of the symptom-free waiting period (SFWP), i.e., the number of consecutive days after starting the return-to-play protocol that an athlete must be symptom-free before they are permitted to unrestricted play. We develop a two-part treatment initiation/cessation simulation model consisting of a (1) Controlled Hidden Markov Model [pre-return-to-play] and (2) Uncontrolled Markov Chain [post-return-to-play] and apply four

simulation-optimization methods (Crude Monte Carlo, 2-Stage Decomposition, NSGS, KN) to optimize the SFWP. For collegiate men's football and women's soccer, we find an optimal SFWP of approximately 2 and 3.5 weeks, respectively. This research provides clinical decision-support for return-to-play decisions.

1 INTRODUCTION

Treatment initiation and cessation settings arise when a physician meets sequentially with a patient and determines when to initiate a new treatment or cease an ongoing treatment. Return-to-play (RTP) from sports-related concussion naturally fits the framework of treatment cessation problems. Concussion has been defined as "a change in brain function following a force to the head, which may be accompanied by temporary loss of consciousness, but is identified in awake individuals with measures of neurologic and cognitive dysfunction" (Carney et al. 2014). Further, approximately 1.6-3.8 million sport and recreation concussions occur each year (Langlois et al. 2006). While there is currently no universal agreement on when an athlete should be permitted to unrestricted play after being diagnosed with a concussion, international and domestic organizations widely support a graduated RTP protocol that begins when the athlete reports symptom-free and ends once the athlete is permitted to unrestricted play. To this end, the optimal length of this *symptom-free waiting period* (SFWP), i.e., the number of consecutive days after starting the RTP protocol that an athlete must remain symptom-free to be permitted to unrestricted play, remains ambiguous. While Markov models provide a natural way of modeling this pre-RTP process (i.e., concussion to RTP decision) and post-RTP process (i.e., RTP decision to 30 days post-RTP), simulation-optimization provides a tractable approach to optimize the length of the SFWP.

The goal of this research is to determine the optimal length of the SFWP as a function of post-RTP injury rates. We create a two-part treatment initiation/cessation simulation model consisting of a (1) pre-RTP and (2) post-RTP phase. For the pre-RTP phase, we model the athlete's concussion dynamics using a Controlled Hidden Markov Model (HMM), which is a useful class of Markov models wherein the underlying disease state is unknown to a clinician and can be inferred only through correlated observations (e.g., findings from clinical assessments). For the post-RTP phase, the physician no longer monitors the athlete's health. As such, we model the athlete's concussion dynamics in this phase using an uncontrolled Markov Chain (UMC). This two-part treatment initiation/cessation simulation model provides a viable method to evaluate SFWP decisions in concussion management. Given a certain SFWP, the simulation model estimates the possible outcomes of the athlete (through simulation replications) where the expected outcome of the athlete is the average over the simulation replications. Simulation-optimization uses the outcomes from the simulation model to solve for the optimal treatment policy (i.e., optimal SFWP in our setting). We consider four simulation-optimization methods when solving for the optimal SFWP: Crude Monte Carlo, 2-Stage Decomposition, NSGS (Nelson et al. 2001), and KN (Kim and Nelson 2001). NSGS and KN are ranking and selection algorithms named as such because of the authors of the original research. Like the 2-Stage Decomposition method, these methods leverage the 2-stage structure of the problem. We apply our methods to a large multi-site dataset on sports-related concussion to estimate the optimal SFWP for collegiate men's football and women's soccer (i.e., the sports with the highest rates of concussions with respect to each sex; Van Pelt et al. 2021). We evaluate each simulation-optimization method by the number of simulation replications required, the optimal SFWP selected, and the variance of various metrics of interest.

In summary, this research makes the following contributions: (1) We propose a general two-part treatment initiation/cessation simulation model that is applicable to a broad class of healthcare applications. (2) We are the first to create a quantitative model that takes a simulation-based approach to optimize the SFWP in sport-related concussion. (3) We apply our model to sport-related concussion in collegiate men's football and women's soccer where the findings suggest (a) an optimal SFWP of approximately 2 weeks and 3.5 weeks, respectively, (b) NSGS and KN allocate a greater number of simulation replications to SFWPs near the optimal SFWP in contrast to Crude Monte Carlo and 2-Stage Decomposition which allocate the same number of simulation replications to all SFWPs, (c) 2-Stage Decomposition, NSGS, and KN select

similar (or the same) optimal SFWPs, and (d) leveraging the two-stage structure of the two-part treatment initiation/cessation model leads to significant variance reduction across all performance measures compared to a Crude Monte Carlo method that ignores the two-stage structure.

This paper is organized as follows: §2 describes the two-part treatment initiation/cessation simulation model. §3 describes the four simulation-optimization methods to solve for the optimal SFWP. In §4, we apply our model to two case studies: collegiate men’s football and women’s soccer. In §5, we conclude the paper with final thoughts and opportunities for future research.

2 TWO-PART TREATMENT INITIATION/CESSATION SIMULATION MODEL

Our overall framework is illustrated in Figure 1. We create a two-part treatment initiation/cessation simulation model consisting of a (1) Controlled Hidden Markov Model [pre-RTP] and (2) Uncontrolled Markov Chain [post-RTP]. We denote \mathbb{Y} as the finite set of permissible SFWPs that can be selected: $\mathbb{Y} = \{0, 1, 2, \dots, 30\}$. Here, $y = 0$ implies you immediately allow the athlete to RTP after being diagnosed with a concussion. We select $y = 30$ days as the largest SFWP that can be selected as the total time to RTP was 28 days or fewer for over 97% of athletes in a large study among collegiate athletes (Broglio et al. 2022). Before allowing the two-part simulation model to evolve, the decision-maker chooses a SFWP $y \in \mathbb{Y}$. We now proceed to describe the two-part simulation model in 2.1 and 2.2, respectively.

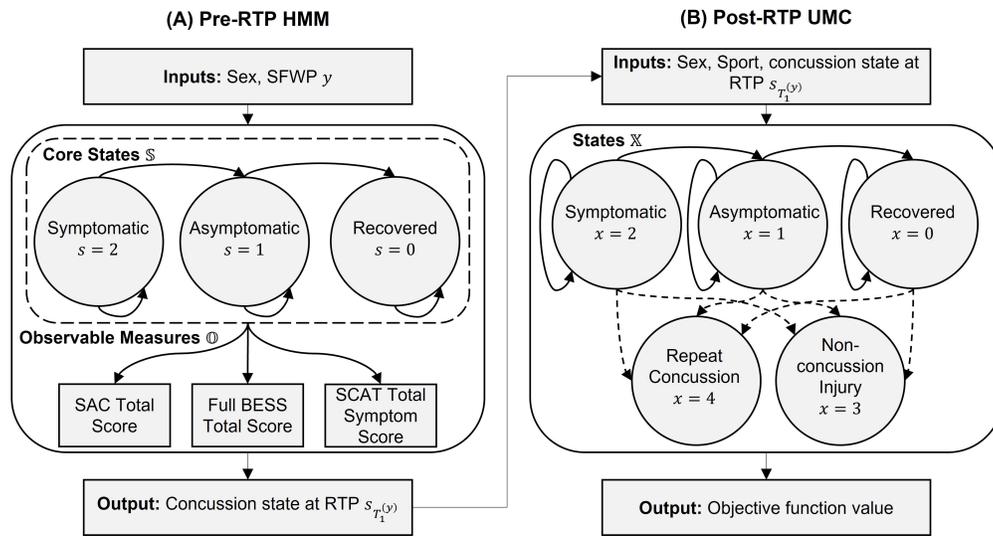


Figure 1: Illustration of two-part return-to-play (RTP) simulation framework composed of (A) Pre-RTP controlled hidden Markov model (HMM) including inputs, core states, observations, and output, and (B) Post-RTP uncontrolled Markov chain (UMC) including inputs, state space, and output.

2.1 Part 1: HMM (pre-RTP)

The HMM is defined by a hidden state space (\mathbb{S}), observation space (\mathbb{O}), stationary state transition probability matrix (\mathbf{P}_1), and stationary observation probability matrix (\mathbf{Q}_1). The hidden state space (i.e., core state space) consists of three states which indicate the severity of the athlete’s concussion: recovered ($s = 0$), asymptomatic ($s = 1$), symptomatic ($s = 2$). For the observation space, we consider a multi-dimensional observation (as recommended by the international consensus statement; McCrory et al. 2017) consisting of the (i) Standard Assessment of Concussion (SAC) total score, (ii) Balance Error Scoring System (BESS) total score, and (iii) the Sport Concussion Assessment Tool (SCAT) total symptom severity score. The SAC, BESS, and SCAT symptom survey assess an athlete’s neurocognitive status, postural control, and symptom presentation, respectively. We denote observation $o \in \mathbb{O}$ as the tuple ((i), (ii), (iii)). We let $t_1 = 0, 1, 2, \dots$

represent the day in the pre-RTP planning horizon and $T_1^{(y)}$ represent the number of days the HMM model evolves given policy $y \in \mathbb{Y}$ is chosen.

We use the defined HMM to simulate the pre-RTP process of the athlete. On day $t_1 = 0$, we assume the athlete is diagnosed with a concussion, *i.e.*, $s_0 = 2$, and that the decision-maker selects a SFWP $y \in \mathbb{Y}$. Based on the choice of y , the HMM evolves as a stochastic process $\{(s_t, o_t)\}$ until day $t_1 = T_1^{(y)}$ which denotes the day that the athlete has been symptom-free for y days. Here, we define symptom-free as the point at which all total scores for the SAC, BESS, and SCAT symptom severity reach previously identified normative baseline performance by sex (Katz et al. 2018). On day $t_1 = T_1^{(y)}$, the athlete is in state $s_{T_1^{(y)}}$.

2.2 Part 2: UMC (post-RTP)

The UMC is defined by a state space (\mathbb{X}) and stationary state transition probability matrix (\mathbf{P}_2). The state space consists of five states, three of which coincide with the HMMs hidden state space: recovered ($x = 0$), asymptomatic ($x = 1$), symptomatic ($x = 2$). The other two states represent absorbing states (\mathbb{X}_A) and correspond to time-loss injury states: non-concussion injury ($x = 3$) and repeat concussion ($x = 4$). We let $t_2 = 0, 1, 2, \dots$ represent the day in the post-RTP planning horizon and T_2 represent the number of days the UMC model evolves.

We use the defined UMC to simulate the post-RTP process of the athlete. Recall that on day $t_1 = T_1^{(y)}$ in the pre-RTP planning horizon, the HMM terminates and the athlete is in state $s_{T_1^{(y)}}$. Equivalently, the athlete is in state $s_{T_1^{(y)}}$ on $t_2 = 0$ in the post-RTP planning horizon (*i.e.*, $x_0 = s_{T_1^{(y)}}$). In the post-RTP planning horizon, the UMC evolves stochastically for a pre-specified T_2 periods.

3 SIMULATION-OPTIMIZATION METHODOLOGIES

Using the two-part simulation model defined in §2, we aim to solve for the optimal SFWP, $y^* \in \mathbb{Y}$, which is given by the shortest SFWP that ensures that the probabilities of non-concussion injuries and repeat concussion attributable to the choice of SFWP are at most $(1+\epsilon_n) \cdot 100\%$ and $(1+\epsilon_r) \cdot 100\%$ of the *natural* (*i.e.*, non-concussed) probability of non-concussion injuries and repeat concussion, respectively. We formally present the optimization problem as Formulation (1):

$$\min \left\{ y \in \mathbb{Y} : \begin{array}{l} \Pr(X_{T_2} = 3|y) \leq \Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \epsilon_n) \quad [\text{constraint 1}] \\ \Pr(X_{T_2} = 4|y) \leq \Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \epsilon_r) \quad [\text{constraint 2}] \end{array} \right\} . \quad (1)$$

In Formulation (1), $\Pr(X_{T_2} = 3|y)$ and $\Pr(X_{T_2} = 4|y)$ represent the probability that after T_2 days, the athlete terminates in the non-concussion injury state and repeat concussion state of the UMC, respectively, given the SFWP y is chosen. These values are calculated based on the simulation-optimization method (in 3.1-3.3). $\Pr(X_{T_2} = 3|X_0 = 0)$ and $\Pr(X_{T_2} = 4|X_0 = 0)$ in Formulation (1) represent the natural probabilities of injury, *i.e.*, the probability that the athlete experiences a non-concussion injury and repeat concussion, respectively, after T_2 days given that the athlete is initially not concussed at the start of the post-RTP planning horizon. These probabilities are given by $\Pr(X_{T_2} = 3|X_0 = 0) = \mathbf{P}_2^{T_2}(3|0)$ and $\Pr(X_{T_2} = 4|X_0 = 0) = \mathbf{P}_2^{T_2}(4|0)$, where $\mathbf{P}_2^{T_2}$ is the matrix \mathbf{P}_2 raised to the T_2^{th} power. The tolerances ϵ_n and ϵ_r in Formulation (1), represent the degree of risk the decision-maker (*i.e.*, clinician and/or athlete) are willing to take in the RTP decision.

We remark that the NSGS and KN simulation-optimization algorithms are designed for unconstrained optimization problems. Hence, we convert Formulation (1) into an unconstrained optimization problem as presented in Formulation (2) with $Z^{(y)}$ representing the objective value for the SFWP $y \in \mathbb{Y}$. Formulation (2) ensures both constraints presented in Formulation (1) are satisfied. Also, once both constraints are satisfied, it enforces the constraints to be as tight as possible which corresponds to selecting the shortest SFWP in our setting. We keep the first term in Formulation (2) which is a constant to emphasize that we enforce the constraints to be as tight as possible, but this constant term can be removed if desired. In

Formulation (2), $\mathbb{1}(\cdot)$ is the indicator function.

$$\min_y Z^{(y)} = \left(\left(\Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \varepsilon_n) \right) + \left(\Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \varepsilon_r) \right) \right) - \left(\Pr(X_{T_2} = 3|y) + \Pr(X_{T_2} = 4|y) \right) \cdot \left(\underbrace{\mathbb{1} \left(\Pr(X_{T_2} = 3|y) \leq \Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \varepsilon_n) \right)}_{\text{[constraint 1]}} \underbrace{\mathbb{1} \left(\Pr(X_{T_2} = 4|y) \leq \Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \varepsilon_r) \right)}_{\text{[constraint 2]}} \right) \quad (2)$$

Using Formulation (2), the optimal SFWP y^* is the SFWP $y \in \mathbb{Y}$ which minimizes $Z^{(y)}$. Next, we describe how to use Formulation (2) with each of the four simulation-optimization methods.

3.1 Method 1: Crude Monte Carlo

Crude Monte Carlo (CMC) is an exhaustive and aggressive approach where for each permissible SFWP $y \in \mathbb{Y}$, we simulate the pre-RTP HMM followed by the post-RTP UMC in succession for N replications. Letting $X_{T_2}^{(k)}$ denote the state at time T_2 for replication k , the CMC objective value $Z^{(y)}$ for any given $y \in \mathbb{Y}$ is given by

$$Z^{(y)} = \left(\left(\Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \varepsilon_n) \right) + \left(\Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \varepsilon_r) \right) \right) - \left(\frac{1}{N} \sum_{k=1}^N \mathbb{1}\{X_{T_2}^{(k)} = 3|y\} + \frac{1}{N} \sum_{k=1}^N \mathbb{1}\{X_{T_2}^{(k)} = 4|y\} \right) \cdot \left(\mathbb{1} \left(\frac{1}{N} \sum_{k=1}^N \mathbb{1}\{X_{T_2}^{(k)} = 3|y\} \leq \Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \varepsilon_n) \right) \mathbb{1} \left(\frac{1}{N} \sum_{k=1}^N \mathbb{1}\{X_{T_2}^{(k)} = 4|y\} \leq \Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \varepsilon_r) \right) \right) \quad (3)$$

We use bootstrapping with 1,000 samples to obtain the variance around the estimate of $Z^{(y)}$ from CMC. That is, for each bootstrap sample, we take a random sample with replacement of size N from the outputs of all CMC simulation replications and estimate $Z^{(y)}$ (see Equation (3)). Then, we compute the variance across all bootstrap samples to estimate the variance of $Z^{(y)}$ from CMC. We use common random numbers across the permissible SFWPs (i.e., \mathbb{Y}) for both the pre-RTP model and post-RTP model.

3.2 Method 2: 2-Stage Decomposition

Like CMC, 2-Stage Decomposition (2SD) is also an exhaustive approach. However, unlike CMC, 2SD relies on the fact that the post-RTP UMC is dependent on y only through the terminating state of the pre-RTP HMM, which provides an estimate of the initial state distribution $\hat{\pi}_y$ for the post-RTP UMC. Therefore, instead of simulating the pre-RTP HMM followed by the post-RTP UMC for N replications, we first simulate the pre-RTP HMM for $k = 1, \dots, N$ replications with common random numbers for each $y \in \mathbb{Y}$ to find the terminating state $s_{T_1}^{(k)}$. Then, for each replication, we set

$$\hat{\pi}_y^{(k)} = \left[\overbrace{\mathbb{1}\{s_{T_1}^{(k)} = 0|y\}}^{x=0} \quad \overbrace{\mathbb{1}\{s_{T_1}^{(k)} = 1|y\}}^{x=1} \quad \overbrace{\mathbb{1}\{s_{T_1}^{(k)} = 2|y\}}^{x=2} \quad \overbrace{0}^{x=3} \quad \overbrace{0}^{x=4} \right]. \quad (4)$$

Letting $[\cdot]_x$ denote the vector element corresponding to state x , the 2SD estimate for the terminating state distribution of the post-RTP UMC in replication k is given by $\Pr(X_{T_2} = x|y) \approx [\hat{\pi}_y^{(k)} \mathbf{P}^{T_2}]_x$ for all $k = 1, \dots, N$ and $x \in \mathbb{X}$. The objective value for any $y \in \mathbb{Y}$ is estimated by

$$Z^{(y)} = \frac{1}{N} \sum_{k=1}^N \left[\left(\left(\Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \varepsilon_n) \right) + \left(\Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \varepsilon_r) \right) \right) - \left([\hat{\pi}_y^{(k)} \mathbf{P}^{T_2}]_3 + [\hat{\pi}_y^{(k)} \mathbf{P}^{T_2}]_4 \right) \cdot \left(\mathbb{1} \left([\hat{\pi}_y^{(k)} \mathbf{P}^{T_2}]_3 \leq \Pr(X_{T_2} = 3|X_0 = 0) \cdot (1 + \varepsilon_n) \right) \mathbb{1} \left([\hat{\pi}_y^{(k)} \mathbf{P}^{T_2}]_4 \leq \Pr(X_{T_2} = 4|X_0 = 0) \cdot (1 + \varepsilon_r) \right) \right) \right]. \quad (5)$$

3.3 Methods 3 and 4: NSGS and KN

In our concussion setting, there exists a small and finite set of permissible SFWPs (i.e., $|\mathbb{Y}| = 31$). This implies it is feasible to use the simulation model defined in §2 with an exhaustive approach like CMC (§3.1) or 2SD (§3.2) which simulates all permissible SFWPs with a fixed number of simulation replications (i.e., N). However, we can more efficiently solve for the optimal SFWP by considering two ranking and selection algorithms, (1) NSGS and (2) KN, which systematically allocate simulation replications to permissible SFWPs (Fu 2015). This implies that the total number of simulation replications completed for each SFWP $y \in \mathbb{Y}$ is not necessarily the same. We denote $N_0^{(NSGS)}, \dots, N_{30}^{(NSGS)}$ and $N_0^{(KN)}, \dots, N_{30}^{(KN)}$ as the number of simulation replications allocated to each permissible SFWP $y \in \mathbb{Y}$ for NSGS and KN, respectively.

NSGS and KN ensure that the best SFWP will be selected with a certain probability (i.e., the probability of correct selection (PoCS); Nelson et al. 2001; Kim and Nelson 2001; Kim and Nelson 2003; Fu 2015). NSGS and KN guarantee PoCS through the notion of the indifference zone. Under this notion, given the difference between the “best” and second-best solution is at least $\delta > 0$, the algorithm selects the “best” solution with a probability of at least $1 - \gamma$ (Fu 2015). Both algorithms assume the outputs of the simulation model are normally distributed, but neither make assumptions on the variance of the simulation outputs. The two key differences between the algorithms are (a) NSGS is a two-stage procedure whereas KN is a sequential procedure and (b) NSGS assumes all simulation outputs are independent from one another (with respect to replication number and system design) whereas KN uses common random numbers implying independence with respect to replication number, but dependence with respect to system design (i.e., treatment policy). The algorithms make no assumptions on the structure of the objective value, which allow the algorithms to be applied to a diverse set of optimization problems.

For NSGS and KN, we define a simulation replication similar to 2SD to estimate the initial state distribution for this replication k (i.e., $\hat{\pi}_y^{(k)}$; see Equation (4)). Similar to 2SD, we calculate the objective value $Z^{(y)}$ for any given $y \in \mathbb{Y}$ using Equation (5) with the corresponding $N_y^{(NSGS/KN)}$.

4 CASE STUDY: RETURN-TO-PLAY AFTER CONCUSSION

We aim to determine the optimal SFWP for men’s American football (i.e., football) and women’s soccer using the four simulation-optimization methods (§3.1-3.3). We analyze the four simulation-optimization methods with respect to the number of simulation replications required, the optimal SFWP selected by the method, and the variance of the metrics of interest. Finally, we numerically validate our model with respect to reported outcomes in the concussion literature. We round all numbers to the nearest thousandths place.

4.1 HMM and UMC Model Parameters

We generate our stationary state transition probability matrix, \mathbf{P}_1 , and stationary observation probability matrix, \mathbf{Q}_1 , using data from the National Collegiate Athletic Association (NCAA) and United States Department of Defense (DoD) Concussion Assessment, Research, and Education (CARE) Consortium (Broglio et al. 2017). The NCAA-DoD CARE Consortium is a large multi-site study, consisting of 30 NCAA institutions and military service academies. Participants include varsity student athletes and military service academy cadets. Our study data consists of serial post-injury assessment among participants who were diagnosed with concussion by local medical staff. That is, each participant was evaluated using the SAC, BESS, and SCAT at several timepoints until they were permitted to RTP. We narrow our study sample to participants with 0 or 1 previous concussions, as past research considering concussion history found no differences among these two groups (Broglio et al. 2022). Due to previously identified sex differences in concussion recovery (Master et al. 2021), we separate our data by sex (i.e., men/women). For each sex, we categorize total scores for the SAC, BESS, and SCAT symptom severity based on pre-defined quantiles.

After data preparation, we apply the Baum-Welch algorithm (Rabiner 1989) with 100 random initializations. The Baum-Welch algorithm is an expectation-maximization procedure for parametrizing HMMs. For most athletes, concussion status does not “worsen” and can only either stay the same or improve. Hence, we impose a left-to-right structure on these initializations, guaranteeing that \mathbf{P}_1 satisfies $\mathbf{P}_1(x'|x) = 0$ for all $x' < x$. We obtain \mathbf{P}_1 and \mathbf{Q}_1 for each sex by selecting the HMMs with the greatest log likelihood.

Next, for each sex, we model our post-RTP stationary state transition probability matrix as

$$\mathbf{P}_2 = \begin{bmatrix} \mathbf{P}_1(0|0) & \mathbf{P}_1(1|0) & 1 - \sum_{x' \neq 2} \mathbf{P}_2(x'|0) & p_n(0) & p_r(0) \\ 0 & \mathbf{P}_1(2|2) & 1 - \sum_{x' \neq 2} \mathbf{P}_2(x'|1) & p_n(1) & p_r(1) \\ 0 & 0 & 1 - \sum_{x' \neq 2} \mathbf{P}_2(x'|2) & p_n(2) & p_r(2) \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where for each $x \in \{0, 1, 2\}$, $p_n(x)$ is the probability of a non-concussion injury and $p_r(x)$ is the probability of a repeat concussion given that the athlete’s concussion status is x . We estimate sex- and sport-specific concussion probabilities from a recovered state $p_r(0)$ using a previously published meta analysis (Van Pelt et al. 2021) and non-concussion injury probabilities $p_n(0)$ from sports injury epidemiology literature at the collegiate level (Roos et al. 2017; Kerr et al. 2018). Next, for $x = 1, 2$, we compute the odds of concussion and non-concussion injury as $\sigma_n(x) = \frac{p_n(0)}{1-p_n(0)} \kappa_x$ and $\sigma_r(x) = \frac{p_r(0)}{1-p_r(0)} \kappa_x$, respectively, where κ_x is an injury risk modifier representing increased risk for injury when the patient RTPs before fully recovering from concussion. We set $\kappa_1 = \kappa_2 = 1.74$ based on previous literature (McCrea et al. 2020). Finally, we convert these odds into probabilities: $p_n(x) = \frac{\sigma_n(x)}{1+\sigma_n(x)}$ and $p_r(x) = \frac{\sigma_r(x)}{1+\sigma_r(x)}$.

4.2 Simulation Parameters

We consider $T_2 = 30$ days, $\varepsilon_n = 0.25$, and $\varepsilon_r = 0.25$. For CMC and 2SD, we set $N = 5,000$. All 95% confidence interval half-widths are at most 0.005 with this number of simulation replications for both men’s football and women’s soccer. For the ranking and selection algorithms, we consider $n_0 = 500$ first-stage replications, $\gamma = 0.01$, and we select $\delta = 0.0025$ for men’s football and $\delta = 0.0005$ for women’s soccer (see §3.3; Fu 2015). We select $\gamma = 0.01$ where $1 - \gamma = 0.99$ represents the confidence in selecting the best solution. With our healthcare application area, we want a high confidence in selecting the correct (i.e., true optimum) SFWP. We select $\delta = 0.0025$ and $\delta = 0.0005$ for men’s football and women’s soccer, respectively, which represents the difference between the best and second-best solution. We choose these values because numerical studies support very small objective values (i.e., $Z^{(y)} \in (0, 0.25)$); see Formulation (2)) and very small differences between objective values across SFWPs. We select these conservative values (i.e., small values) to ensure we can distinguish the correct solution with high probability.

4.3 Numerical Results

Figures 2 and 3 illustrate the Mean Non-Concussion Injury Probability, Mean Repeat Concussion Probability, Mean Objective Value, and Number Simulation Replications against the SFWP for collegiate men’s football and women’s soccer, respectively. We analyze the results from three perspectives: number of simulation replications required, optimal SFWP selected by the method, and variance of the metrics of interest.

4.3.1 Number of Simulation Replications

CMC and 2SD allocate 5,000 simulation replications to all permissible SFWPs. NSGS and KN allocate at least 500 replications to all SFWP (i.e., $n_0 = 500$), but also systematically allocate additional simulation replications to permissible SFWPs. NSGS and KN allocate additional replications to SFWPs near the optimal solution selected by these algorithms. Across all SFWPs explored, CMC, 2SD, NSGS, and KN applies a total of 155,000, 155,000, 99,490, and 48,713 replications, respectively, for men’s football. CMC,

2SD, NSGS, and KN applies a total of 155,000, 155,000, 225,861, and 41,946 replications, respectively, for women's soccer. We suspect NSGS allocates a larger number of simulation replications in comparison to KN because NSGS is a two-stage procedure (i.e., one opportunity to add replications) whereas KN is a sequential procedure (i.e., iterative opportunity to add replications until a condition is satisfied).

4.3.2 Selection of the Optimal SFWP

For CMC, 2SD, NSGS, and KN, the optimal SFWP minimizes the objective function in Formulation (2). With an increase in the SFWP, the objective function in Formulation (2) is designed to decrease then slowly increase. However, CMC fails to capture this trend whereas 2SD, NSGS, and KN capture this trend.

For men's football, CMC selects $SFWP^* = 6$ whereas 2SD, NSGS, and KN select an optimal SFWP of about 2 weeks (i.e., $SFWP^* = 12$, $SFWP^* = 13$, and $SFWP^* = 13$, respectively). For women's soccer, CMC again selects $SFWP^* = 6$. 2SD, NSGS, and KN select an optimal SFWP of about 3.5 weeks (i.e., $SFWP^* = 23$, $SFWP^* = 23$, and $SFWP^* = 25$, respectively). For men's football, the findings suggest that after the athlete starts the RTP protocol (i.e., reports symptom-free), they should remain symptom-free for about 2 weeks before being permitted to unrestricted play. For women's soccer, the findings suggest about 3.5 weeks. These results are with respect to $\varepsilon_n = 0.25$ and $\varepsilon_r = 0.25$. Clinicians and/or athletes with a higher risk-tolerance may increase these parameters (e.g., $\varepsilon_n = 0.5$ and $\varepsilon_r = 0.5$) to shorten optimal SFWPs.

4.3.3 Variance of the Metrics of Interest

Figures 2 and 3 illustrate 95% confidence intervals around the Mean Non-Concussion Injury Probability, Mean Repeat Concussion Probability, and Mean Objective Value. CMC leads to much greater variability with respect to the mean injury probabilities in comparison to 2SD, NSGS, and KN. This extra variability in CMC may stem from the simulation of the post-RTP process, in comparison to the closed-form derivation used in 2SD, NSGS, and KN. Moreover, the absorbing probabilities for post-RTP states $x = \{3, 4\}$ are so small that they may require a large number of replications to even be reached. For the mean objective value, all methods provide tight confidence intervals around this metric except when applying NSGS and KN to small SFWP (e.g., $SFWP < 10$). However, the mean objective value for CMC can be drastically different, even for similar SFWPs, suggesting additional simulation replications may be necessary to ensure the stability of the model. To this end, we have run an auxiliary experiment with $N = 10,000$ replications for CMC and find that the estimates for mean injury probabilities and the mean objective value are still highly variable. These results suggest that CMC is not reliable for decision-making in practice.

4.4 Validation of Results

We first validate our results by comparing our estimated post-RTP repeat concussion probabilities with published same-season repeat concussion probabilities in football (McCrea et al. 2020). This study compares differences in mean SFWP and repeat concussion rates during the CARE study (2014-2017) and an NCAA study (1999-2001). We focus our validation on findings from the NCAA study since the CARE study uses a subset of the data we used to parameterize the HMM. The NCAA study had a mean SFWP of 3.25 days with a same-season repeat concussion probability of 0.0652 (12 out of 184). To generate comparable estimates, we apply 2SD to our male football model with $N = 5000$ replications, $SFWP=3$ days, and $T_2 = 90$ days. We record the repeat concussion probability of 0.0594 (≈ 297 of 5000). Using the two-sample Z-test for proportions, we find that our estimated probability of repeat concussion is not significantly different from the NCAA study ($P = 0.743$) at a significance level of $\alpha = 0.05$. We also check the face validity of our results through review with concussion experts, including some of our study's coauthors. While

future research can facilitate additional validation of our results, we suspect our simulation-based results may provide reasonable estimates for the relationship between SFWP and injury rates.

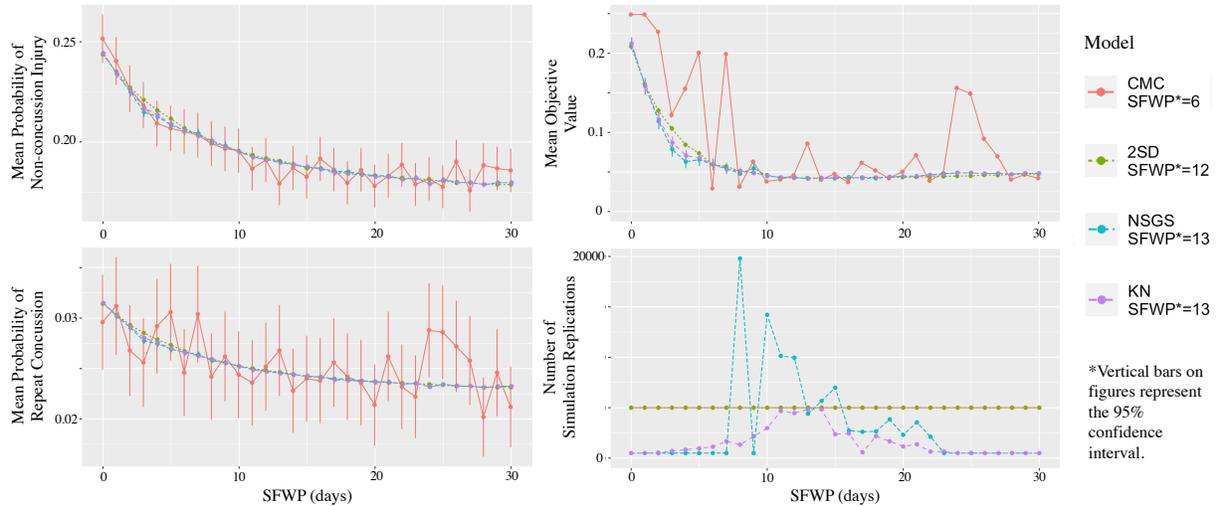


Figure 2: Results for men’s football with optimal symptom free waiting period (SFWP*) in the legend.

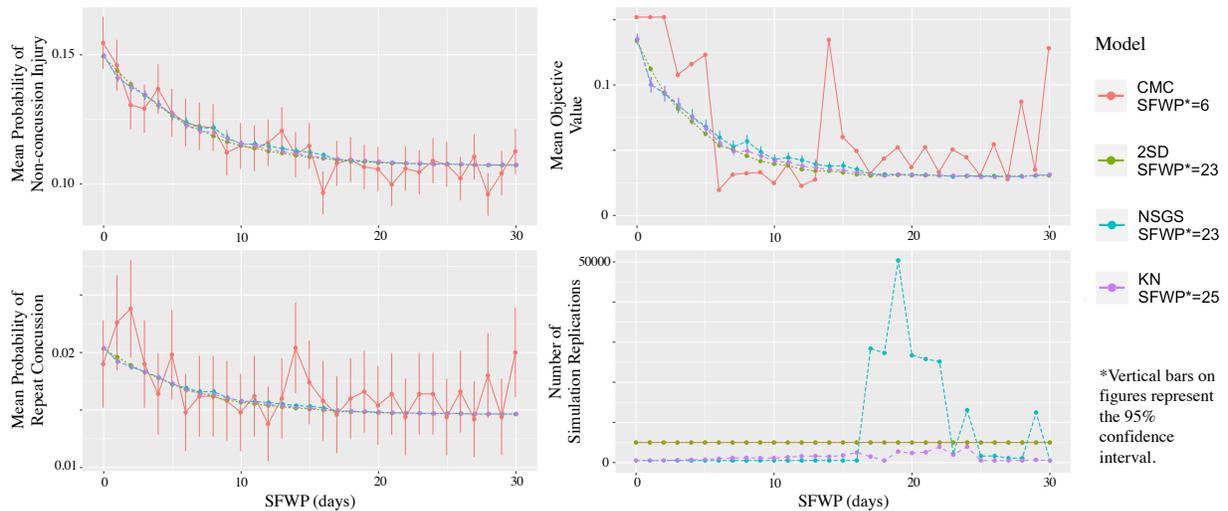


Figure 3: Results for women’s soccer with optimal symptom free waiting period (SFWP*) in the legend.

5 CONCLUSION

We design a two-part treatment initiation/cessation simulation model and apply four simulation-optimization methods to solve for the optimal SFWP in sport-related concussion. To our knowledge, this is the first quantitative model that takes a simulation-based approach to optimize the SFWP for sports-related concussion. We find that after the athlete starts the RTP protocol (i.e., reports symptom-free), the athlete should remain symptom-free for about 2-3.5 weeks before being permitted to unrestricted play. The length of this SFWP is more conservative than what is typically seen in current practice, suggesting that risk tolerances in practice may be higher than we tested in our analysis. Additionally, a 2-3.5 week SFWP may be difficult to

incorporate in practice due to the pressure faced by athletes to return to competition quickly. Nevertheless, our algorithm has the flexibility to be calibrated according to the decision makers' objectives.

For research extensions, (1) we consider one symptom-free criteria based on current practice. Defining symptom-free is challenging as some participants exhibit symptoms at baseline (i.e., prior to a concussion), so it may be beneficial to incorporate more measures (e.g., vestibular/ocular) into the symptom-free criteria. Future research can consider embedding the symptom-free criteria into the decision-making process to simultaneously optimize the symptom-free criteria and SFWP. (2) We analyze a single set of simulation parameters and collegiate athletes in men's football and women's soccer. Our findings illustrate that the SFWP for men's football and women's soccer differs by about 1.5 weeks. Exploring a broader range of simulation parameters and other sport/sex combinations (e.g., basketball) may provide insights on why this difference arises. (3) Our analysis is based on Markov Models. For applications where the necessary assumptions for these models are not met, other approaches, e.g., data mining, may be more appropriate.

Sport-related concussion management, especially RTP decision making, poses a significant challenge for physicians and athletic trainers. Despite the central role of the SFWP in sport-related concussion, guidance on how to tailor the SFWP to athletes is limited or practically non-existent. By leveraging simulation, we provide a quantitative foundation for optimizing the SFWP. Our two-part simulation model can also provide valuable decision support in other healthcare applications.

ACKNOWLEDGMENTS

Alphabetically by last name, the CARE Consortium Investigators include (USA): Scott A. Anderson, ATC (Univ of Oklahoma); Holly J. Benjamin, MD, FACSM, FAMSSM, FAAP (Univ of Chicago); Alison Brooks, MD, MPH (Univ of Wisconsin-Madison); Thomas Buckley, EdD, ATC (Univ of Delaware); Kenneth L. Cameron, PhD, MPH, ATC, FNATA (United States Military Academy); James R. Clugston, MD, MS, FAMSSM (Univ of Florida); Stefan M. Duma, PhD (Virginia Tech); Jessica Dysart Miles, PhD, LAT, ATC (Florida College); James T. Eckner, MD, MS (Univ of Michigan); Luis A Feigenbaum, PT, DPT, ATC/L (Univ of Miami); Christopher C. Giza, MD (Univ of California, Los Angeles); Kevin M. Guskiewicz, PhD, ATC (Univ of North Carolina- Chapel Hill); Thomas W. Kaminski, PhD, ATC, FNAK, FNATA, FACSM, RFSA (Univ of Delaware); Louise A. Kelly, PhD (California Lutheran Univ); Anthony P. Kontos, PhD (Univ of Pittsburgh); Christina L. Master, MD, FAAP, CAQSM, FACSM (Univ of Pennsylvania); Jason P. Mihalik, PhD, CAT(C), ATC, FACSM, FNATA (Univ of North Carolina at Chapel Hill); Christopher M. Miles, MD, FAMSSM (Wake Forest Univ); Patrick G. O'Donnell, MHA, COTA/L (UMass Memorial Health); Justus Ortega, PhD (Humboldt State Univ); Nicholas Port, PhD (Indiana Univ); Julianne Schmidt, PhD, ATC (Univ of Georgia); Adam Susmarski, DO, FAAPMR (United States Naval Academy).

We thank Jody Harland, Susan Perkins, Barry Katz, Janetta Matesan, and Larry Riggen (Indiana Univ), Ashley Rettmann, Nicole L'Heureux (Univ of Michigan), Melissa Koschnitzke (Medical College of Wisconsin), Michael Jarrett, Vibeke Brinck, and Bianca Byrne (Quesgen), Thomas Dompier, Christy Collins, Melissa Niceley Baker, and Sara Dalton (Datalys Center for Sports Injury Research and Prevention), Stacey Harcum (Uniformed Services Univ), and the research/medical staff at each participating site. We thank the three anonymous *Winter Simulation* reviewers for their constructive feedback.

This publication was made possible in part by the National Collegiate 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 Combat Casualty Care Research Program, endorsed by the Department of Defense, through the Joint Program Committee 6/ Combat Casualty Care Research Program – Psychological Health and Traumatic Brain Injury Program under Award No. W81XWH1820047. Opinions, interpretations, conclusions, and recommendations are

those of the author and are not necessarily endorsed by the Department of Defense. Lauren Czerniak was supported by the National Science Foundation Research Fellowship Program [Grant DGE 1841052]. Lauren Czerniak and Mariel Lavieri received support from the NSF CAREER Award [CMMI-1552545].

REFERENCES

- Broglio, S. P., T. McAllister, B. P. Katz, M. LaPradd, W. Zhou, M. A. McCrea, A. Hoy, J. B. Hazzard, L. A. Kelly, J. DiFiori, J. D. Ortega, N. Port, M. Putukian, D. Langford, J. McDevitt, D. Campbell, J. C. Jackson, G. McGinty, C. Estevez, K. L. Cameron, M. N. Houston, S. J. Svoboda, A. J. Susmarski, C. Giza, H. J. Benjamin, T. W. Kaminski, T. Buckley, J. R. Clugston, J. Schmidt, L. A. Feigenbaum, J. T. Eckner, J. Mihalik, J. D. Miles, S. Anderson, K. Arbogast, C. L. Master, A. P. Kontos, S. P. D. Chrisman, M. A. Brooks, S. Rowson, S. M. Duma, and C. 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.
- Broglio, S. P., M. McCrea, T. McAllister, J. Harezlak, B. Katz, D. Hack, and B. 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.
- Carney, N., J. Ghajar, A. Jagoda, S. Bedrick, C. Davis-O'Reilly, H. du Coudray, D. Hack, N. Helfand, A. Huddleston, T. Nettleton, and S. Riggio. 2014. "Concussion Guidelines Step 1: Systematic Review of Prevalent Indicators". *Neurosurgery* 75:S3–S15.
- Fu, M. 2015. *Handbook of simulation optimization*. New York: Springer.
- Katz, B. P., M. Kudela, J. Harezlak, M. McCrea, T. McAllister, S. P. Broglio, S. Anderson, J. Bazarian, C. T. Bullers, C. Giza, K. Guskiewicz, J. D. Myles, P. O'Donnell, S. J. Svoboda, R. Tierney, J. Harland, J. Matesan, L. Riggen, S. Zhao, Ashley, E. Rettmann, M. Koschnitzke, M. Jarrett, V. Brinck, B. Byrne, T. Dompier, M. N. Baker, and S. Dalton. 2018. "Baseline Performance of NCAA Athletes on a Concussion Assessment Battery: A Report from the CARE Consortium". *Sports Medicine* 48 (8): 1971–1985.
- Kerr, Z. Y., G. B. Wilkerson, S. V. Caswell, D. W. Currie, L. A. Pierpoint, E. B. Wasserman, S. B. Knowles, T. P. Dompier, R. Dawn Comstock, and S. W. Marshall. 2018. "The first decade of web-based sports injury surveillance: Descriptive epidemiology of injuries in United States high school football (2005-2006 through 2013-2014) and National collegiate athletic association football (2004-2005 through 2013-2014)". *Journal of Athletic Training* 53 (8): 738–751.
- Kim, S.-H., and B. L. Nelson. 2001. "A Fully Sequential Procedure for Indifference-Zone Selection in Simulation". *ACM Transactions on Modeling and Computer Simulation* 11 (3): 251–273.
- Kim, S.-H., and B. L. Nelson. 2003. "Selecting the best system: theory and methods". In *Proceedings of the 2003 Winter Simulation Conference*, edited by S. Chick, P. Sanchez, D. Ferrin, and D. Morrice. Dec. 7-10, New Orleans, LA, 101-112.
- Langlois, J. A., W. Rutland-Brown, and M. M. Wald. 2006. "The Epidemiology and Impact of Traumatic Brain Injury: a brief overview". *The Journal of Head Trauma Rehabilitation* 21 (5): 375–378.
- Master, C. L., B. P. Katz, K. B. Arbogast, M. A. McCrea, T. W. McAllister, P. F. Pasquina, M. Lapradd, W. Zhou, and S. P. Broglio. 2021. "Differences in sport-related concussion for female and male athletes in comparable collegiate sports: a study from the NCAA-DoD Concussion Assessment, Research and Education (CARE) Consortium". *British Journal of Sports Medicine* 55 (24): 1387–1394.
- McCrea, M., S. Broglio, T. McAllister, W. Zhou, S. Zhao, B. Katz, M. Kudela, J. Harezlak, L. Nelson, T. Meier, S. W. Marshall, and C. I. Guskiewicz, Kevin M. 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.
- McCrory, P., W. Meeuwisse, J. Dvorak, M. Aubry, J. Bailes, S. Broglio, R. C. Cantu, D. Cassidy, R. J. Echemendia, R. J. Castellani, G. A. Davis, R. Ellenbogen, C. Emery, L. Engebretsen, N. Feddermann-Demont, C. C. Giza, K. M. Guskiewicz, S. Herring, G. L. Iverson, K. M. Johnston, J. Kissick, J. Kutcher, J. J. Leddy, D. Maddocks, M. Makdissi, G. T. Manley, M. McCrea, W. P. Meehan, S. Nagahiro, J. Patricios, M. Putukian, K. J. Schneider, A. Sills, C. H. Tator, M. Turner, and P. E. Vos. 2017. "Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016". *British Journal of Sports Medicine* 51:838–847.
- Nelson, B. L., J. Swann, D. Goldsman, and W. Song. 2001. "Simple Procedures for Selecting the Best Simulated System When the Number of Alternatives is Large". *Operations Research* 49 (6): 950–963.
- Rabiner, L. 1989. "A tutorial on hidden Markov models and selected applications in speech recognition". *Proceedings of the IEEE* 77 (2): 257–286.
- Roos, K. G., E. B. Wasserman, S. L. Dalton, A. Gray, A. Djoko, T. P. Dompier, and Z. Y. Kerr. 2017. "Epidemiology of 3825 injuries sustained in six seasons of National Collegiate Athletic Association men's and women's soccer (2009/2010–2014/2015)". *British Journal of Sports Medicine* 51 (13): 1029–1034.
- Van Pelt, K. L., T. Puetz, J. Swallow, A. P. Lapointe, and S. P. Broglio. 2021. "Data-Driven Risk Classification of Concussion Rates: A Systematic Review and Meta-Analysis". *Sports Medicine* 51:1227–1244.

AUTHOR BIOGRAPHIES

GIAN-GABRIEL P. GARCIA (co-first author) is an Assistant Professor at the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. His research develops data-driven frameworks which integrate predictive and decision analytic methods for high-impact problems in health policy and medical decision-making, with a focus on health equity and social/behavioral dynamics. Prior to joining Georgia Tech, he was a postdoctoral fellow at Harvard Medical School. He received his PhD and MS in Industrial and Operations Engineering at the Univ of Michigan and BS in Industrial Engineering from the Univ of Pittsburgh. His email is giangarcia@gatech.edu. His website is <https://www.giangabrielgarcia.com>.

LAUREN L. CZERNIAK (co-first author) is a Ph.D. student in the Department of Industrial and Operations Engineering at the Univ of Michigan. She received her BS in Industrial Engineering from the Univ of Pittsburgh and received her MSE in Industrial and Operations Engineering from the Univ of Michigan. Her research focuses on developing and applying stochastic models to address current challenges in healthcare with applications in pharmaceutical drugs, glaucoma, and concussion management. Her email is czernL@umich.edu. Her website is <https://sites.google.com/umich.edu/lauren-czerniak/home>.

MARIEL S. LAVIERI is an Associate Professor in the Department of Industrial and Operations Engineering at the Univ of Michigan. She received her BS in both Industrial and Systems Engineering and Statistics from the Univ of Florida. She received her MS and PhD in Management Science from the Univ of British Columbia. Her research applies operations research to healthcare topics with methods such as dynamic programming, stochastic control, and continuous, partially observable state space models. Her email is lavieri@umich.edu. Her website is <http://www-personal.umich.edu/~lavieri/>.

SPENCER W. LIEBEL is an Assistant Professor and Clinical Neuropsychologist in the Department of Neurology at the Univ of Utah School of Medicine. He received his PhD and MS in Clinical Psychology from the Univ of Georgia and BS in Psychology from Brigham Young Univ. His research and clinical interests center on sport-related concussion to clarify their behavioral and neural mechanisms to improve clinical detection, outcomes prediction, and recovery monitoring. His email is spencer.liebel@hsc.utah.edu. His website is <https://healthcare.utah.edu/fad/mddetail.php?physicianID=u0863033&name=spencer-w-liebel>.

MICHAEL A. MCCREA is the Shekar N. Kurpad, MD, PhD Chair and Professor of Neurosurgery and Co-Director of the Center for Neurotrauma Research at the Medical College of Wisconsin. He also serves as a research neuropsychologist at the Clement Zablocki VA Medical Center in Milwaukee, Wisconsin. He is currently co-PI of the NCAA-DoD Concussion Assessment, Research and Education (CARE) Consortium and co-I on the Transforming Research and Clinical Knowledge in TBI (TRACK-TBI) study. His email is mmccrea@mcw.edu. His website is <https://www.mcw.edu/departments/neurosurgery/people/michael-mccrea-phd-abpp>.

THOMAS W. MCALLISTER is the Albert Eugene Sterne Professor of Psychiatry at the Indiana Univ School of Medicine. Dr. McAllister has been working in the field of brain injury recovery for over 30 years. He has written widely on the neuropsychiatric sequelae of traumatic brain injury (TBI), and is one of the four Principal Investigators of the NCAA and U.S. Department of Defense funded Concussion Assessment, Research, and Education (CARE) Consortium. His email is twmcalli@iupui.edu. His website is <https://medicine.iu.edu/faculty/22290/mcallister-thomas>.

PAUL F. PASQUINA is a Professor and Chair of the Department of Physical Medicine & Rehabilitation at the Uniformed Services Univ of the Health Sciences (USUHS), where he is also the Director of the Center for Rehabilitation Sciences Research. He received his BS degree from West Point and M.D degree from USUHS. He is a retired U.S. Army Colonel, who currently also serves as the Chief of Rehabilitation at Walter Reed National Military Medical Center. His email is paul.pasquina@usuhs.edu. His website is <https://www.usuhs.edu/node/10525>.

STEVEN P. BROGLIO is a Professor of Kinesiology, Neurology, and Physical Medicine and Rehabilitation at the Univ of Michigan. He is the Associate Dean of Graduate Affairs for the School of Kinesiology and serves as the Director of the Michigan Concussion Center and the NeuroTrauma Research Laboratory. In these roles, he oversees clinical care, educational outreach, and multi-disciplinary research aimed at fundamental questions on concussion prevention, identification, diagnosis, management, and outcomes. His email is broglio@umich.edu. His website is <https://concussion.umich.edu/leadership/steven-broglio/>.

CARE CONSORTIUM INVESTIGATORS are the site investigators and teams from the 36 affiliations that support the work and research of the NCAA-DoD CARE Consortium. The NCAA-DoD CARE Consortium is a large multi-site study, consisting of 30 NCAA institutions and military service academies. The NCAA-DoD CARE Consortium email is careac@iupui.edu. The NCAA-DoD CARE Consortium website is <https://careconsortium.net>.