The Establishment and Refinement of the National Basketball Association Player Injury and Illness Database

**Christina D. Mack, PhD, MSPH*†; Peter Meisel, MSPH‡;
Mackenzie M. Herzog, PhD, MPH*†; Lisa Callahan, MD§;
Eva E. Oakkar, PhD, MS*; Taylor Walden, BA‡; Joseph Sharpe, ATC∥;
Nancy A. Dreyer, PhD, MPH*†; John DiFiori, MD‡§**

*IQVIA Real-World & Analytics Solutions, Research Triangle Park, Durham, NC; †University of North Carolina at Chapel Hill; ‡National Basketball Association, New York, NY; §Hospital for Special Surgery, New York, NY; ∥National Basketball Athletic Trainers Association, Charlotte, NC

The National Basketball Association (NBA; also referred to as "the league") has established a centralized, audited electronic medical record system that has been linked with external sources to provide a platform for robust research and to allow the NBA to conduct player health and safety reviews. The system is customized and maintained by the NBA and individual teams as part of the employment records for each player and is deployed uniformly across all 30 teams in the league, thereby allowing for standardized data on injuries, illnesses, and player participation in NBA games and practices. The electronic medical record data are enriched by linkage with other league external data sources that provide additional information about injuries, players, game and practice participation, and movement. These data linkages allow for the assessment of potential injury trends, development of injury-prevention programs, and rule changes, with the ultimate goal of improving player health and wellness. The purpose of this article is to describe this NBA injury database, including the details of data collection, data linkages with external data sources, and activities related to reporter training and data quality improvement.

Key Words: electronic medical records, electronic health records, sports injuries, injury prevention

The US National Basketball Association (NBA; also referred to as "the league") health and safety program has established the systematic collection of injury and illness data across all teams, which is linked with external sources provided by the NBA and its teams to create a platform for robust research and to allow the NBA to conduct player health and safety reviews. These data are important for assessing potential injury trends in the league and shaping the development of injury-prevention programs and rule changes, with the ultimate goal of improving player health and wellness.

In 2012, to advance its efforts in this area, the NBA began establishing a centralized injury database on the backbone of an electronic medical record (EMR) system in accordance with the collective bargaining agreement (CBA) between the NBA and the National Basketball Players Association. The EMR's primary purpose is to facilitate teams' clinical documentation of player injuries. A secondary aim of this database is to enable league health and safety reviews and other injury-research and -prevention efforts. To create a high-quality database of player injuries, the injury data in the EMR are systematically inspected using audit and data-quality processes and then linked with many other contextual and exposure-related data sources.

The purpose of this paper is to describe the NBA injury database, including the details of data collection, linkages with external data sources, and activities related to reporter training and data quality improvement.

INJURY REPORTING IN THE NBA

Within the team structure, player health is typically managed by team physicians and certified athletic trainers (ATs), often with support from others, including physical therapists, sports scientists, and strength and conditioning coaches. The record for each player is maintained within a standardized, audited system that has been customized for the NBA and is deployed uniformly across all 30 teams in the league. Injury and illness data for all NBA players on a roster are entered into the EMR system by team medical staff (primarily ATs) as events occur. If a player changes teams, the new team is granted access to the applicable health record, allowing for continuity of a team's management of an injury record and the player's overall history. Records of "2-way players" (certain players contracted to play primarily in the G League—the NBA's official minor league—with a limited number of NBA service days) are included only for the time they participate in NBA team activities.
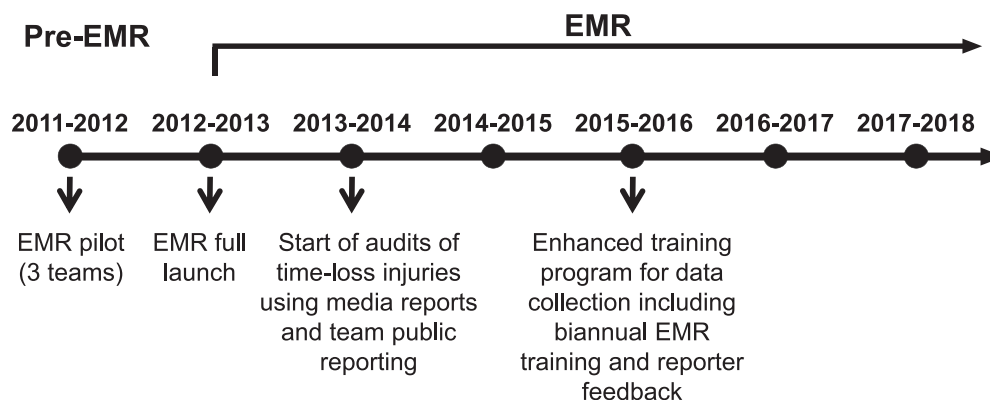
Figure 1. Evolution of injury reporting in the National Basketball Association. Abbreviation: EMR, electronic medical record.

In accordance with the CBA, data collection was transitioned into the league-wide, centralized EMR system in 2012 and adopted by all teams at the start of the 2012–2013 NBA season (Figure 1). Before 2012, injury data were collected by teams centrally through a surveillance system,[1] but data entry was not audited by the NBA. Although the injury database has evolved over time, the injury reporting definition has been stable since 2007–2008. Before the 2013–2014 NBA season, as reporting within the EMR was established, mandated, and standardized, the league initiated structured reporter guidance, training programs, feedback mechanisms, and auditing and data quality-control processes. These audit and data-quality activities, coupled with continuous improvements in the technical EMR system, facilitate efficient and complete data collection.

DATA LINKAGES

The EMR injury data are enriched by linkage with external data sources from the NBA and its teams that provide additional information about injuries, players, participation, and movement, allowing for a better understanding of player exposure and potential confounders (Figure 2). Daily linkage to the NBA game statistics database provides up-to-date roster information and game participation to the second. Additionally, the linkage with game statistics allows for the determination of game schedules (eg, back to backs) and travel information on both a team and player level. Information on travel for each team includes distance and direction traveled, amount of travel time, and change in time zones. In-game optical-camera data recording movement of players and the ball have been collected league wide since the 2013–2014 season and are used to calculate certain metrics (eg, speed, distance run, mechanical load, and accelerations, as well as game movement coordinates).[2] Other individual characteristics collected by the league or teams and related to players, such as height, weight, age, body composition, and the number of years they played basketball in college, have also been integrated.

Data from the EMR and the league contain key variables, including *injury ID*, *athlete ID*, and *team ID*, that allow for linkage of the various datasets from both sources in the NBA database. Although much of this information (eg, team rosters and game statistics) can be found in public sources, only data from official NBA sources are integrated

into the database. These data are joined using profile data; information such as athlete names, dates of birth, and other identifiers are securely stored and not provided to external researchers or included within datasets used for analyses unless required. Processes are in place to govern secure data transfers and data storage between the league, the EMR provider, and research and technical entities.

DEFINITIONS

Detailed descriptive information is recorded for injuries, including specific type (eg, sprain, strain, fracture), basketball activity (eg, game, practice, shootaround, conditioning), specific event (eg, shooting, rebounding), onset (acute versus nonacute), and mechanism (contact or noncontact). Clinical treatment information such as physician visits, medication use, and rehabilitation data (eg, modalities and protocols) is also collected in the EMR.

Injury and Illness

Team medical staff are required to report injuries and illnesses in accordance with a uniform injury definition. In this system, an *injury* is defined as resulting in 1 or more of the following: a missed game or practice, the issuing of a prescription medication, or the need for "extraordinary care," which refers to sutures or topical skin adhesive; injection or aspiration; an x-ray, ultrasound, magnetic resonance imaging, computed tomography scan, bone scan, or other special test; intravenous administration of fluids; or a hospital visit. Team staff are instructed to close an injury record when the player has been cleared for full participation without restrictions and is no longer undergoing treatment for the injury or illness.

Specific Injury Definitions

Team medical staff enter injury diagnoses in the EMR based on prespecified drop-down options for body region (eg, head, lower extremity), body part (eg, nose, ankle, knee), body site (eg, nasal bone, deltoid ligament, anterior cruciate ligament), laterality (eg, left, bilateral), and injury type (eg, fracture, sprain, strain). To facilitate league health and safety reviews and research, clinical definitions have been developed to combine injury diagnoses into 108 specific injuries (eg, lateral ankle sprain, metatarsal bone stress injury) and 64 broader injury categories (eg, ankle sprain, bone stress injury). The specific injuries and
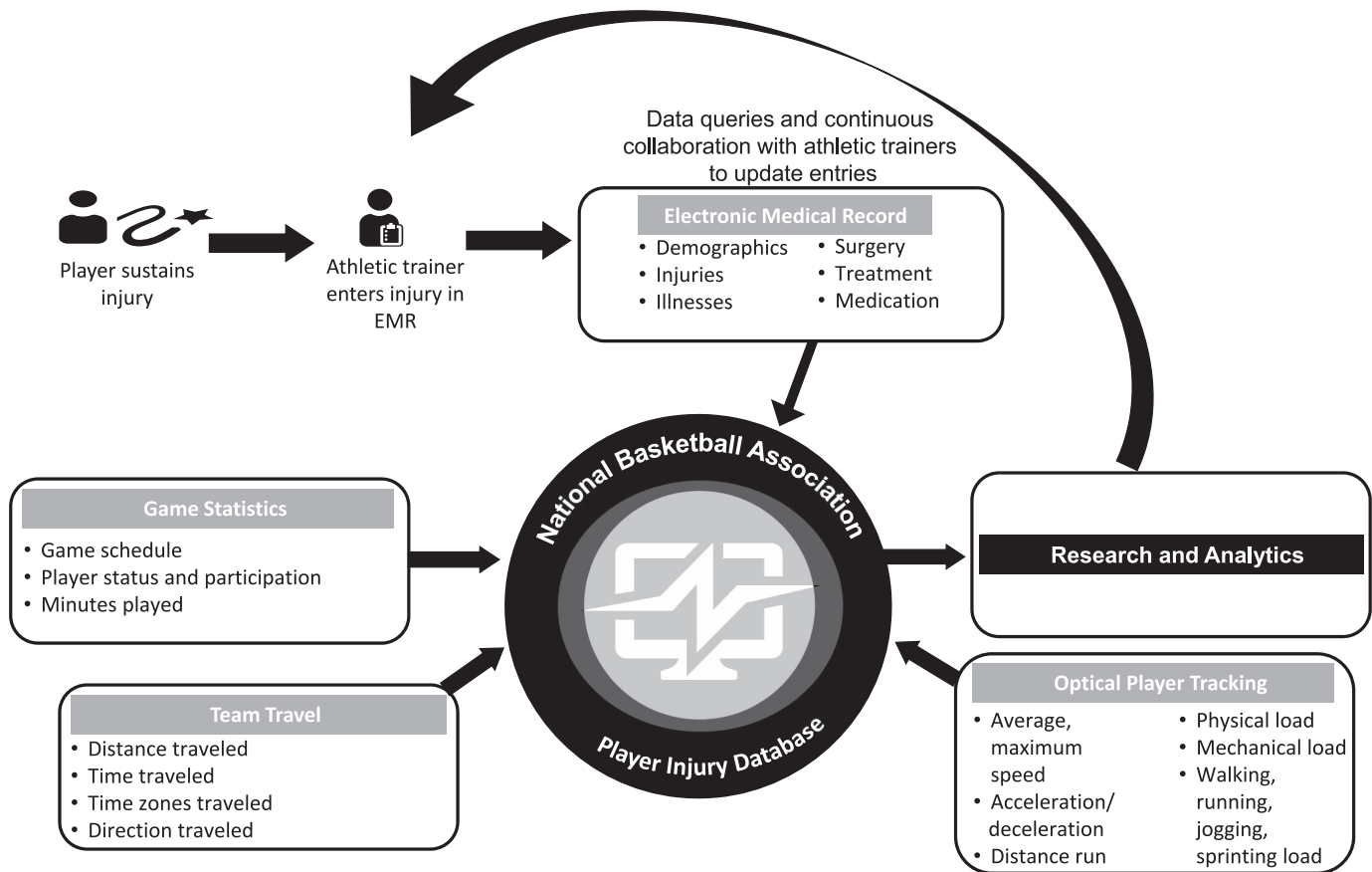
Figure 2. National Basketball Association player injury database. Abbreviation: EMR, electronic medical record.

categories are mutually exclusive and focus on injuries of high frequency, severity, or clinical interest. If an injury occurs that is not part of the current nomenclature, the injury is added as needed, both as an option in the EMR and to the corresponding clinical definitions.

Basketball Participation and Exposure

The database contains numerous metrics to quantify a denominator of exposure to injury during basketball participation when calculating rate metrics.[3] Granular information is available for player game participation, quantifying both whether the player played in the game and the duration of participation in minutes (player-minutes of participation per game; Table). This allows for the calculation of metrics such as the injury rate per game played or the injury rate per minute on the court. Teams are required to enter practice participation data into the EMR, including the date of the team's practice, whether each player participated, and, if the player did not participate, the specific reason why the player was unable to participate (eg, linking the practice to a specific injury or illness). However, although game participation is audited against league game data, practice participation is not audited and is subject to team-specific schedules and reporting.

In addition to the player-level injury rate per game or per minute played, methods for quantifying basketball participation include *team-game* (1 team participating in 1 game to calculate the injury rate per game at the team level), *team-season* (1 team participating for 1 season), and *player-season* (1 player with any participation in an NBA season to calculate the injury rate per season). Through the integration of optical-camera data, additional participation metrics (such as distance run, physical and mechanical loads, and physical and mechanical intensity) for players are available, which can be used to quantify exposure.

Time Loss and Non–Time Loss

As noted earlier, team medical staff are required to record player participation in the EMR and report any games in which a player was unable to participate because of rest or a specific injury or illness. As a result, injuries can be examined in total or stratified by whether games were missed as a result. To assess the effect of injury and time to return to play, the total number of games missed because of injuries can be calculated overall, by specific injury types, and by teams or players.

Acute Versus Nonacute Onset

Injuries can be reported as having an acute or nonacute onset. *Acute injuries* are defined as having occurred because of a specific, identifiable event. *Nonacute injuries* are defined as being unrelated to a specific event and often correlate with chronic, insidious, preexisting, or overuse injury subclassifications; examples of nonacute injuries are bone stress injuries and tendinopathies.

Table. National Basketball Association Players and Games, 2013–2014 Through 2017–2018

Metric/Season	Year				
	2013–2014	2014–2015	2015–2016	2016–2017	2017–2018
Players, No.[a]	552	557	582	572	606
Preseason					
Games	115	118	109	102	78
Player-games	2796	2831	2730	2640	2042
Player-minutes	53 770.27	54 310.25	50 317.73	48 185.37	35 491.27
Regular season					
Games	1230	1230	1230	1230	1230
Player-games	25 618	25 981	26 078	26 140	26 107
Player-minutes	595 207.20	595 207.18	594 856.75	594 806.25	593 856.07
Postseason					
Games	89	81	86	79	82
Player-games	1906	1653	1900	1737	1729
Player-minutes	43 170.78	39 330.33	41 530.47	38 070.52	39 560.53

[a] Includes players who played at least 1 preseason, regular season, or postseason game.

Reinjury

The definition used for a *reinjury* is an injury of the same type and at the same site as an index injury that occurs after a player has returned to full participation and is no longer receiving treatment for the index injury.[4] Accordingly, ATs are instructed to enter reinjuries into the EMR as new injury records, which allows for collection of the same comprehensive information about the reinjury as that collected for index injuries.[5] This definition has been provided to team medical staff and is reinforced through reporter training. If injury exacerbations occur while the player is still undergoing treatment, medical staff can enter the date of the exacerbation within the existing injury record.

IMPLEMENTATION AND DATA QUALITY

Reporter Guidance and Training

Athletic trainers play a central role in entering information into the injury database, serving as both primary reporters and consistent health care providers for the athletes and teams. Team physicians and other members of the team medical staff also contribute to injury and return-to-play documentation. As such, a key component of the injury database is communication about reporting practices, standardization, and guidance to ATs and team physicians. Communicating the results of league health and safety reviews, research, and prevention efforts originating from research in the database is central when emphasizing to team medical staff the importance and utility of the data.

To enhance consistent data entry and gather user feedback, in-person and remote meetings are conducted on a recurring basis. These meetings highlight new EMR features, focus on data-entry topics, and present select findings from health and safety reviews. In addition, the training explains how EMR changes may affect data-entry processes and interpretations. For example, training may introduce a new field to capture the exact time of injury during a game, describe the correct EMR field and the way to log treatment, or provide guidance on how to uniformly define the injury onset. Summarized information on injury incidence and characteristics is disseminated annually along with data-quality reports and reminders. Additional contacts with team medical staff occur throughout the season when needed, as well as through individual queries regarding specific data entries during the data-curation process.

Auditing and Data Quality

Systematic auditing and field cross-checks are performed to assess data accuracy. Systematic auditing involves a rigorous line-by-line review by the league office and epidemiology partners (IQVIA). Audits cover all injuries reported in the EMR and include comparisons with team media injury reports. If a discrepancy is identified, the NBA league office works with team medical staff to resolve the concern and update the EMR appropriately.

In addition to the audits, field cross-checks are regularly conducted to scan the data for missing or contradictory entries. When problems are found, team medical staff are contacted to correct the EMR so that it accurately reflects the circumstances of the injury event. For example, injuries reported in the EMR as occurring during a game must be associated with a game date; if participation data show that the player did not participate in a game on the date of injury, the team is contacted to review the entry. Logic checks also cover onset discrepancies: for example, an injury reported as chronic but with a corresponding activity indicating an acute event. Auditing and data-quality processes are continuously adapted and improved based on these findings.

AVAILABILITY FOR RESEARCH

In accordance with the CBA and subject to appropriate approvals, the injury database can be used to allow authorized academicians and researchers to access the data and conduct studies designed to improve player health and broaden medical knowledge. This may include researchers affiliated with the NBA, its teams, or the National Basketball Players Association or unaffiliated researchers who are authorized to use these data pursuant to a process overseen by the NBA's Research Committee. The database supports research questions critical to player

health and injury, NBA game schedules, and measures of player load.

DISCUSSION

Consistent, accurate, and timely injury reporting is an important aspect of enhancing player health league wide; however, collecting such data, particularly in the context of clinical care in the sports environment, can be challenging. Sports injury surveillance has historically been a voluntary system of reporting,[6–9] at times focused on a somewhat narrow definition of a *reportable injury*. In certain cases, the data in these systems have been less reliable for understanding true patterns of injury occurrence, as it is difficult to quantify differences in teams that choose to report versus those that do not. It is also difficult to identify which data are missing and the subsequent effect on interpretation. Newer systems for data collection have been difficult to establish in many contexts, largely because of decentralized reporting structures, the added time commitment on staff for systematic reporting, and sometimes unwieldy technology. As in any clinical setting, establishing a database that promotes consistent, accurate, and timely recording of injuries and illnesses is not an easy process. Progress across the field has been made in recent years, particularly in the supporting technologies, allowing for the crude estimation of the incidence and burden of injuries, but these programs are still often challenged by limitations in the completeness of data entry, a lack of roster-based denominators, and definitions that underestimate the true effects of injury.

The introduction of a league-wide EMR system alongside development of a systematic program for data monitoring and linkages has resulted in a more complete system that can be used to provide timely and meaningful epidemiologic insights about injuries. Through linkage with exposure data, movement metrics, and sophisticated game statistics, the database enables a more accurate measure of athlete game exposure, and standardized, mandated reporting coupled with auditing processes allows for relatively complete descriptions of player injuries. The connection with per-player basketball participation metrics specifically is one of the more scientifically important aspects of this system, as it allows for robust injury analytics that account for the amount of injury exposure in an NBA game setting.[3] These efforts are particularly valuable because they enhance the ability to address injury-related questions as they arise: for example, examinations of potential increases or decreases in injuries during a season, the effect of schedules,[10] or evaluations of changes to the game relevant to injury-prevention efforts.

Improvements in data capture and quality over the years influence how data can be used and interpreted. Changes in incidence across years can be attributable to true differences in injury occurrence but may also be due to changes induced by teams' improved reporting. For instance, the first year of the EMR was a transition time with challenges similar to those of EMR adoption in any other employment or clinical setting. Consequently, data on injuries before 2013–2014 are used for research purposes only under limited circumstances, such as an analysis of consistently defined and reported injuries (eg, anterior cruciate ligament ruptures). Data from the 2013–2014 season forward are the strongest and most consistent, having been subjected to more consistent audit and data-quality processes. These data-quality and -analytic approaches could potentially be applied, perhaps with some modifications, to other sports and levels of competition.

CHALLENGES AND FUTURE DIRECTIONS

No surveillance system is perfect, and the context of sport injury presents a set of methodologic challenges that are still being resolved. Differences in accuracy and completeness of reporting among teams may exist; however, audit processes are designed to reduce this variability, and reporting by team is monitored each season to ensure that there are no significant, unexpected differences. Further, a record may not be updated as consistently if a player moves between the G League and the regular league. Also, as in any health or employment records system, follow-up in the NBA database is restricted to active players and current employees, and no information is available for players after they leave the league, which affects the ability to understand the full effects of injuries. Similarly, although medical examinations are completed at the time of entry into the league, information on injury and health history before the NBA is limited. As players move between NBA teams, continuity of care with primary providers is lost, which introduces the potential for variability. Non–team-related offseason training may also be incomplete, which limits the ability to capture and describe all basketball activity.

Although the consistent entry of injury recurrences and nonacute injuries has been prioritized in recent years, it can be challenging to distinguish new injuries from exacerbations of prior injuries, which may result in underreporting or misclassification of reinjuries. Also, nonacute injuries, by nature, have an onset that occurs over a prolonged period of time, and it can be difficult for providers to enter an accurate onset date. For these reasons, definitions are refined through discussions with team medical staff on a continuing basis.

Another key challenge to understanding the complexities of playing a professional sport relates to developing metrics that can be used to characterize performance, particularly in the context of return to play after an injury. For example, metrics such as minutes played and distance run are affected by both player fitness and external factors such as a coach's decision on how much playing time to give a particular player, the injury status of other players on the team, and the performance of teammates. As with all data collection for injury surveillance, a goal of these efforts is to ensure that injuries and the accompanying clinical and contextual information are reported consistently and completely across all teams.

CONCLUSIONS

A database using a sport-specific EMR coupled with quality checks and ongoing communication with team medical staff can support robust, timely, and responsive health and safety reviews. The NBA EMR system also may more broadly illustrate how an EMR can be used for research. The power inherent in the EMR linkage with sources of player-exposure data is a step forward

compared with traditional injury-reporting systems that were not centralized or linked in the past or that were based entirely within the construct of clinical care. Monitoring player injuries and providing data to make evidence-based clinical and policy-level decisions, especially those aimed at injury prevention, are high priorities for the NBA and its teams. Regular quality assessment of EMR data entry, accompanied by periodic audits, data queries, and use of the data for various types of reporting, coupled with continuous refinement over time, has resulted in a powerful program that is being used to enhance player health in the NBA and may serve as a research model for other sport programs.

REFERENCES

1. Starkey C. Injuries and illnesses in the National Basketball Association: a 10-year perspective. *J Athl Train*. 2000;35(2):161–167.
2. NBA announces multiyear partnership with Sportradar and Second Spectrum. National Basketball Association Web site. https://pr.nba.com/nba-announces-multiyear-partnership-sportradar-second-spectrum/. Accessed March 12, 2019.
3. Knowles SB, Marshall SW, Guskiewicz KM. Issues in estimating risks and rates in sports injury research. *J Athl Train*. 2006;41(2):207–215.
4. Shrier I, Steele RJ, Zhao M, et al. A multistate framework for the analysis of subsequent injury in sport (M-FASIS). *Scand J Med Sci Sports*. 2016;26(2):128–139.
5. Fuller CW, Bahr R, Dick RW, Meeuwisse WH. A framework for recording recurrences, reinjuries, and exacerbations in injury surveillance. *Clin J Sport Med*. 2007;17(3):197–200.
6. Dick R, Agel J, Marshall SW. National Collegiate Athletic Association Injury Surveillance System commentaries: introduction and methods. *J Athl Train*. 2007;42(2):173–182.
7. Centers for Disease Control and Prevention (CDC). Sports-related injuries among high school athletes—United States, 2005–06 school year. *MMWR Morb Mortal Wkly Rep*. 2006;55(38):1037–1040.
8. Powell JW, Barber-Foss KD. Injury patterns in selected high school sports: a review of the 1995–1997 seasons. *J Athl Train*. 1999;34(3):277–284.
9. Drakos MC, Domb B, Starkey C, Callahan L, Allen AA. Injury in the National Basketball Association: a 17-year overview. *Sports Health*. 2010;2(4):284–290.
10. Mack CD, Herzog MM, DiFiori JP, Meisel PL, Dreyer NA. A second look at NBA game schedules: response to Teramoto et al. *J Sci Med Sport*. 2018;21(3):228–229.

Address correspondence to Christina D. Mack, PhD, MSPH, IQVIA Real-World & Analytics Solutions, Research Triangle Park, 4820 Emperor Boulevard, Durham, NC 27703. Address e-mail to Christina.Mack@IQVIA.com.