



REAL-WORLD EVIDENCE IN SPORT

Matthew S. Tenan PhD ATC FACSM¹, William Adams PhD ATC FACSM^{2,3,4,5}

¹Rockefeller Neuroscience Institute, West Virginia University, Morgantown, WV, USA; ²United States Olympic & Paralympic Committee, Colorado Springs, CO, USA; ³U.S. Coalition for the Prevention of Illness and Injury in Sport, Colorado Springs, CO, USA; ⁴Department of Kinesiology, University of North Carolina at Greensboro, Greensboro, NC, USA; ⁵School of Sport, Exercise and Health Sciences, Loughborough University, National Centre for Sport and Exercise Medicine (NCSEM), Loughborough, UK

KEY POINTS

- Real-world data are prevalent and growing in sport, but the quality, documentation and storage of this data is inconsistent.
- Much of sport is context-specific and data from only one domain (e.g. strength & conditioning or sports medicine) can lead practitioners towards incomplete or misleading conclusions.
- Current data practices in sport focus on descriptive analyses or prediction, neither of which should be used to change clinical practice, improve athlete care or increase performance.
- The generation of real-world evidence, designed to change clinical practice, is dependent both on valid analytical methods used to generate evidence, but also the reliability, validity and relevance of the underlying real-world data.
- Real-world evidence requires counterfactual thinking (e.g. how would this outcome change if I do “this” instead of “that”) and reflecting that thought process both in data analyses and ongoing clinical practice.

INTRODUCTION

Across all disciplines, the availability and use of data is widespread with the origins of this growth born from the creation of the internet. From a medicine perspective, the shift from paper records to electronic health records (EHR) was cemented when the Institute of Medicine (IOM) put forth their guidelines for the transition in 1997 (IOM, 1997). In many ways, modern Sport Science technology was made possible by the invention of low-power 9-degrees of freedom (9DOF) sensor chips, also called inertial measurement units, that were developed in 2013 (MPU-9150, 2024), which allowed practitioners to quantify athlete movement on a larger scale. Given the advent of these advancements in the sports medicine and sport science space, the use of athlete management system software and dashboards in sport has exploded at an exponential rate.

Despite the substantial increase in the recording of data from all aspects of sport, the inconsistent storage and documentation of this data make it challenging to use for performance, medical or research purposes over the long term. Furthermore, data are often kept in disciplinary ‘silos’ (e.g., sports medicine, nutrition, strength and conditioning, performance, etc.), meaning there is no integration of data between these teams. This lack of a cohesive framework for athletic support staff to share data in a structured way make it nearly impossible for any singular discipline to practice optimally because they lack the full information to make decisions and act on those decisions.

Metaphorically speaking, it is the equivalent of a poker or blackjack player trying to win without looking at their cards; they lack all the information they need to make the best decisions. By integrating data across all sources (e.g., sports medicine, sport science, etc.) and then using this information to inform athlete-centered decisions, we can make giant strides forward in serving our athletes.

Real-world data in sport is defined as data which is routinely collected from a variety of sources relating to the health or performance of an athlete/team for the delivery of healthcare or training. While all professions want to claim they are “data-driven” or “evidence-based”, simple missteps in the use, handling and/or analysis of real-world data make it more likely that our decisions are based on a flawed or incomplete picture of the athlete than a data-informed picture (Tenan, 2023b). This Sport Science Exchange (SSE) article addresses uses of real-world data, how to turn real-world data in to real-world evidence (RWE) and why no singular discipline in a sport organization can claim to be “data-driven” or “evidence-based” without the integration other disciplines’ data sources for decision-making.

USES OF REAL-WORLD DATA IN SPORT

The United States Food & Drug Administration (FDA) has published extensively on the use of real-world data (USFDA, 2023). Many of their guidelines are helpful in understanding how we should use real-world

data in sport; for example, we can re-word the FDA's purely medical uses of real-world data and place them in a sport or athletics context:

- Generation of hypotheses for future randomized-control trials in sport
- Informing prior probability distributions in a Bayesian statistical model
- Assessing the feasibility of doing research on a particular population (e.g. just linebackers in American football)
- Assembling cohorts of research participants for rare injuries requiring stratification (i.e. a multi-university study on anterior cruciate ligament rupture as a function of the menstrual cycle in female athletes)
- Identifying new biomarkers for performance or injury
- Developing prognostic/predictive indicators of injury or performance
- The creation of dashboards which show athlete or team data with simple analytics such as z-scores, means, maximum, minimum, correlations or points-over-replacement (added by the authors, not FDA)
- Generation of evidence to change clinical practice (i.e. RWE)

In the above list, the first four bullets are purely related to the planning of future *prospective* research. While those are important for researchers, we will not focus on them in this SSE article as most real-world data used in sport is retrospective. The following three bullets make up nearly all uses of real-world data in sport at the time of this writing. The identification of new biomarkers for performance or injury makes up a large proportion of the published research in the realm of sport science and medicine. These are *purely descriptive* analyses of a biomarker, including describing movement patterns of different athletes or positions on a team, or sleep patterns on a team across a season. The next bullet is the development of predictive models, which are sometimes published in the open literature but are more likely to be used internally by many teams or private companies. It is common in sport science and medicine to take descriptive analyses and treat them as predictive models, despite a lack of demonstrated validity in their out-of-sample predictive ability. Common examples of this issue are the minimal clinically important difference (MCID) (Boyer et al., 2022; Tenan & Boyer, 2023), “modifiable risk factors” from a descriptive multivariable model (Losciale et al., 2023; Tenan, 2023a) and the acute:chronic workload ratio (Impellizzeri et al., 2020a,b,c,d, 2021; Kalkhoven et al., 2021). The second-to-last bullet, the development of dashboards for practitioner use, will often have both descriptive purposes (e.g. a line of some metric across time for an athlete) as well as some predictive model (i.e. ‘readiness score’ or ‘injury score’). Descriptive, predictive and dashboard uses of real-world data are common, but so is the misunderstanding about how this data should be validly used and interpreted (Tenan, 2023b).

At the core, we have to question “why do we want to describe performance or injury?” and “why do we want to predict performance or injury?” The answer to both questions will be something akin to “because we can use this information to change how we optimally train or provide healthcare services to athletes.” However, there is a logical (but hidden) step being taken in this process; it is implied that we already *know* what needs to be done to change the future which has been predicted by our prediction model or we *know* that what we are describing definitively leads to better performance on the pitch/field/court for that athlete or team. In reality, we know that many medical prediction models do not lead to better outcomes once deployed to clinicians (Kappen et al., 2018; Schertz et al., 2023; Zhou et al., 2021). However, we often have no empirical evidence that training for a specific performance metric (e.g. squat strength, counter-movement jump, 40-yard dash, etc.) or surrogate medical endpoint (e.g. decreased heel-striking, hip flexor strength, functional movement screening, patient-reported outcomes) factually cause an increase in game performance or decrease injury, particularly in team sports. Chelsea Football Club does not go through their season with the goal of “describing what it is like to win the Premiership” (a passive process), they train throughout the season doing things to which they believe will cause them to win the Premiership (an active process). The Kansas City Chiefs do not want to predict if they will win the National Football League title, they want to train and play in such a way that it causes them to win the title. A key aspect is active causation and the counterfactual, “what causes me to win versus what causes me to lose.”

THE GENERATION OF REAL-WORLD EVIDENCE (RWE) IN SPORT

RWE provides the sport science and sports medicine practitioner with causal information from real-world data to improve their practice that descriptive or predictive analyses cannot. To do this, we need to realize that sport is not special; we don't need to invent novel methods, nor should we claim that something is impossible to know. Rather, sport can apply concepts and tools from philosophy, statistics, computer science, physics, economics and epidemiology to problems that exist in the specific domain of sport. This is not to say that we currently have all the tools or analytic methods we need to solve every sport-related question, even if the perfect data did exist! Imperfect data or imperfect methods may stop us from answering every question we have in sport, but they do not stop us from answering meaningful questions as long as our assumptions are clearly stated, and the limitations of the evidence are clear. The worst outcome of any research in sport is that it fools both the practitioner and their stakeholders into believing something with absolute certainty when the true evidence for causation is extremely weak.

Causality and Considering the Counterfactual

While we will leave hefty philosophical discussions on definitions of causality to philosophers and physicists, it is pertinent to consider the famous philosopher David Hume's definition of causality in 1748:“...

we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second. Or in other words where, if the first object had not been, the second never had existed” (Hume, 1993).

Expressed a different way: in order for causation to occur, one must follow another in time (i.e. “an object, followed by another”) and if the ‘causal object’ goes away, a counterfactual will arise (i.e. “if the first object had not been, the second never had existed”). This counterfactual reasoning is a key aspect of why we can have confidence that a randomized-control trial (RCT) tells us that the intervention *causes* a different outcome than what is seen in the ‘control’ group. Neither descriptive analyses nor predictive ones provide us with the counterfactual “if this-then that” statement that can be applied to our disciplinary practice.

Considering the Counterfactual: A Sport Example

It is, perhaps, easier to understand counterfactual reasoning in the context of a sport science experiment. In our hypothetical experiment, we have a 10-person basketball team playing a single game. We select 5 players to consume one ‘control beverage’ that is indistinguishable from our ‘super sauce beverage’ (which 5 players consume), and we want to know if this super sauce beverage causes players to score more points in the game. After the game ends, we have the data in Table 1. A naïve analysis would simply determine the difference in points scored between the groups and conclude that our super sauce beverage works! Let’s head to market, right? Not so fast. As you can also see in Table 1, we have a bunch of “unfilled” cells of data and unless we can fill in those missing data points and know exactly how many points that player would have scored in the exact same game with the other beverage, we cannot make causal conclusions. This is especially important in a team sport where the number of points scored by an individual player may be caused by a multitude of factors that are different for each game. For this reason, it is often said that counterfactuals to understand causation are a “missing data problem” (Ding & Li, 2018; Robins, 1986; Shpitser et al., 2015). This missing data problem can be solved either experimentally with varying designs of an RCT or through RWE studies.

If Not a Randomized-Control Trial, Then Real-World Evidence

In addition to counterfactual reasoning, the second aspect of an RCT that allows us to draw causal conclusions is that of randomization. If you are randomly selecting a large enough number of people to be in different groups for a study, this allows a researcher to ignore unmeasured (and often unmeasurable) variables that would affect the outcome, allowing us to focus on the causal effect of the intervention (Senn, 2013). While the RCT has many virtues, they are often either impractical or impossible to execute in sport. It is hard to imagine a coach or general manager allowing their staff to provide a supplement or medical treatment to one half of the team and a placebo supplement/treatment to the other half. However, that is more likely than the owners

Athlete	Drink	Points with Control	Points with Super Sauce
1	Super Sauce		22
2	Control	15	
3	Super Sauce		10
4	Super Sauce		12
5	Super Sauce		5
6	Control	4	
7	Control	5	
8	Super Sauce		6
9	Control	7	
10	Control	1	
Observed Average		6.4	11.0

Table 1: A Sport Example

of all teams in the National Basketball Association allowing researchers to randomly assign all of their current players to a new set of 30 teams and apply a set of interventions at the team level! This is where RWE analyses play a vital part in our ability to make definitive statements about what intervention or treatment causes a player/team to perform better, not get injured or return to play faster in the real-world.

Observational RWE used for decision-making has a lot in common with a high-quality RCT, despite the system for obtaining data being completely passive:

- The ‘study’ is designed prospectively, not after analyzing the data (a poor practice closely related to hypothesizing after the results are known, HARKing).
- The study outcome must be meaningful, ideally a primary outcome (injury or game/match/event performance) and not a surrogate measure of performance or injury.
- The intervention or interventions must be compared against a meaningful control condition.
- The data obtained must be fit to answer the question based on the study design, including relevant end points, appropriate levels of missing data, consistency in documentation and recording of meaningful confounders.

Every decision a clinician makes in sport has a counterfactual. If you are training an athlete one way, you are not training them another way. If you are applying a prophylactic brace or recovery treatment, you are not simultaneously also applying another treatment. Each of these interventions are made with the idea that they are changing the counterfactual world where the intervention (training, recovery treatment, brace, etc.) had not occurred.

The Casual Framework for Real-World Evidence in Sport

Observational causal inference is an entire area of research and expertise, extending beyond the scope of this SSE. What is important to know is that there are two primary frameworks: the Potential Outcomes Framework and the Structural Causal Model. While both approaches have seemingly easy to apply software packages (Blöbaum et al., 2022; Greifer, 2022; Mayer et al., 2023; Sant'Anna & Zhao, 2020; Sharma & Kiciman, 2020; Textor et al., 2016) it is essential to emphasize that applying observational causal inference analytics in a valid way is an extremely advanced skill set, even for individuals with a Ph.D. in statistics, epidemiology or computer science. As such, it is necessary to seek out experts in this field and to realize what the sport practitioner can bring to the table that allows for the creation of RWE. Since the sport practitioner is involved with the athletes and recording of data on a daily basis, they are also able to speak to its reliability, how often

data points are missing, why the data may be missing or why certain things are measured at different frequencies. Oftentimes, experts in observational causal inference will not have a deep understanding of the theoretical processes which underlie the way in which someone gets injured or how their athletic performance increases or decreases. In analytical terms, this is called the *data generating process*, and this is the core information necessary that a sport practitioner can provide to RWE experts (Ho et al., 2023). The most helpful thing a practitioner can do to facilitate RWE at their organization, beyond championing organizational data sharing and collaboration, is draw a flow chart which articulates this data generating process, also called a directed acyclic graph (DAG). In the process of drawing this DAG (see Figure 1 as an example), it will become clear that no practitioner in sport acts in isolation and it is vitally important that data from multiple disciplines are aggregated together in order to determine what causes athlete or team performance to improve or injury to decline.

ELIMINATE DATA TURF BATTLES WITHIN YOUR ORGANIZATION

All disciplines within a sport organization need to have a shared vision on how athlete health and winning are intertwined, balanced and optimized (Tenan & Alejo, 2024). The sports medicine staff cannot optimize player health without knowing the ongoing mechanical, physiological and psychological loads being placed on the athletes. The nutrition

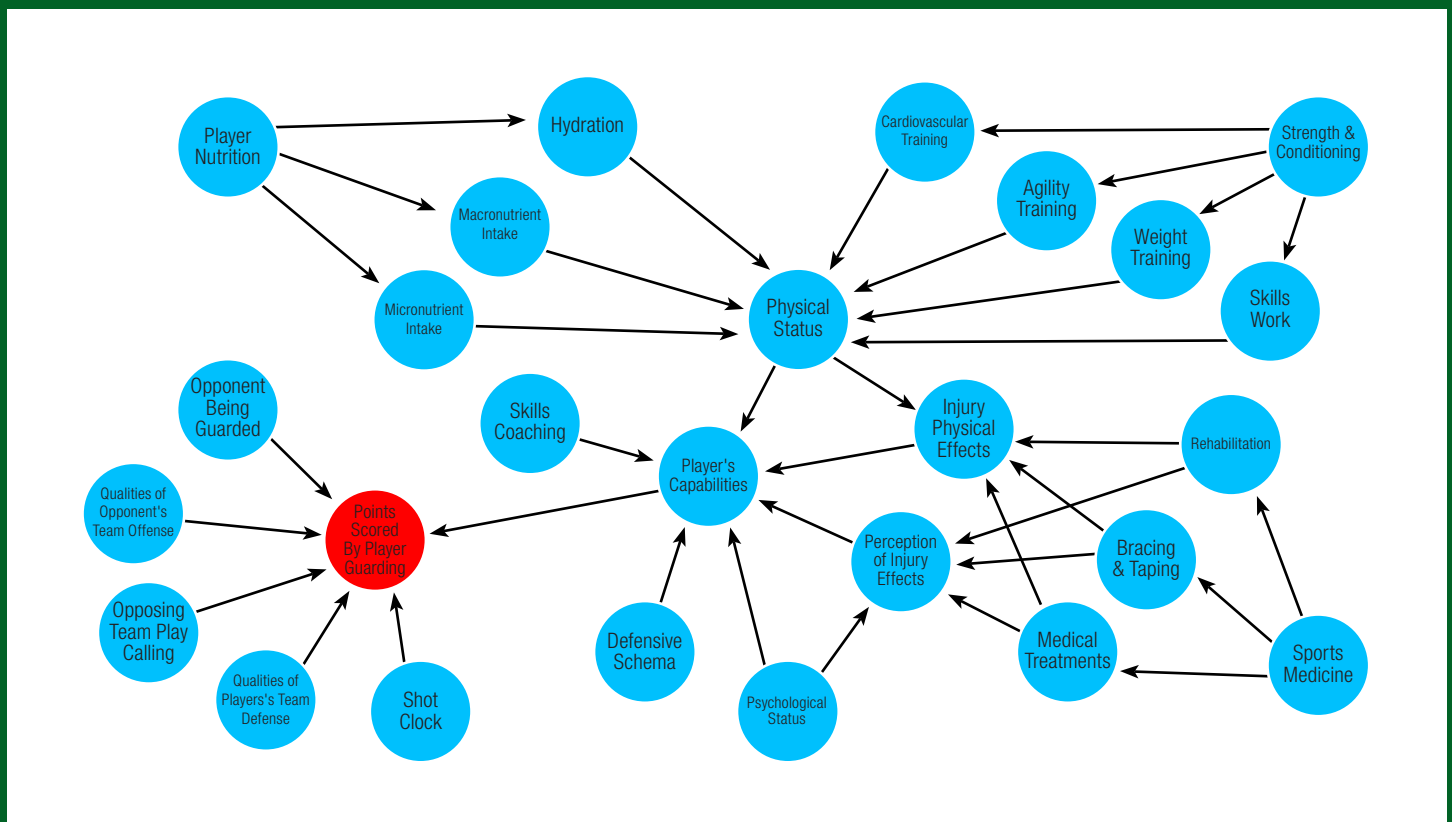


Figure 1. A conceptual Directed Acyclic Graph (DAG) depicting the data concepts (blue nodes) causing an outcome in sport (red node). This simplified example is what causes a basketball player to play effective defense (i.e. stop the opponent they are guarding from scoring).

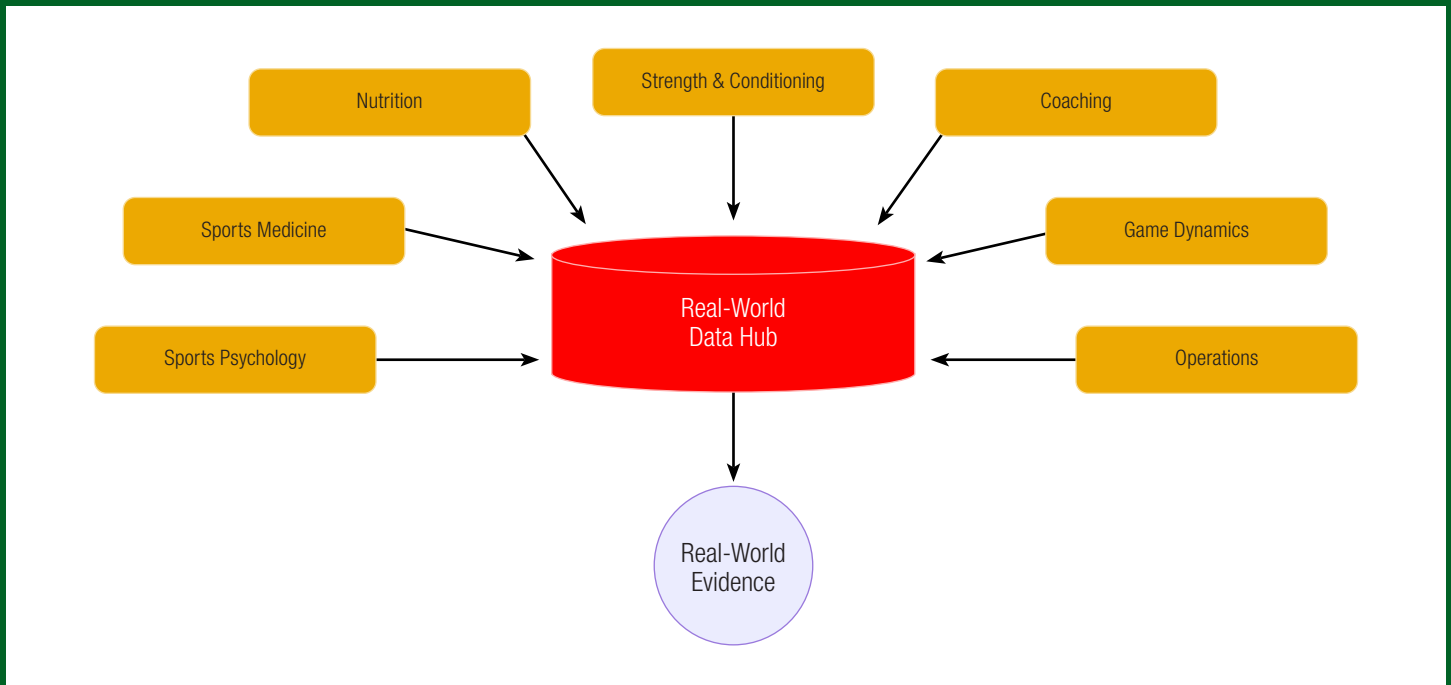


Figure 2. Flowchart of the necessary data organizational structure to create RWE in sport.

staff cannot determine ideal feeding patterns without knowing general caloric expenditure and that athlete's psychological relationship to food. The strength and conditioning staff cannot increase performance on the pitch/field/court without knowing the style of play preferred by the coaching staff, if a player has a lingering injury or if there is a periodized nutrition plan in place for weight management. Defining appropriate data governance, how disciplines within an organization share data in an automated way, is a key aspect of a high-functioning sports franchise that facilitates the creation of RWE that will guide the franchise towards greater success.

CONCLUSIONS

A well-run sport organization will have a real-world data "hub" which centralizes the data from each discipline (Figure 2). This hub is a centralized database that automatically extracts, cleans and organizes data from the athlete injury software, athlete management system, nutrition diaries/tracking, psychological testing, sport science technology, technology employed by the team/league during games and other pertinent sources. The centralization of data allows for RWE analyses that account for the data generating processes defined by the practitioners.

- The sport scientist wants to know if their implemented recovery technology has improved in-game performance after accounting for opponents, injuries, coaching decisions, etc. We can now do that.

- The sports medicine staff wants to know if a specific bracing/taping strategy decreases ankle sprains after accounting for surface types, play tactics, opponents, player fatigue, etc. We can now do that.
- The coaching staff wants to know if implementing a different defensive strategy is going to be effective against a specific opponent, knowing the physical and psychological status of their team. We can now do that.

These are all theoretical, but practical examples exist of what can be accomplished analytically with RWE studies when all disciplines within an organization are working together, sharing data and prioritizing organizational effectiveness over internal turf battles. In fact, Figure 2 is wrong. RWE in sport is not an end-product for a well-run organization, RWE is what provides a meaningful feedback mechanism for each discipline within the organization to continually improve their practice (Figure 3).

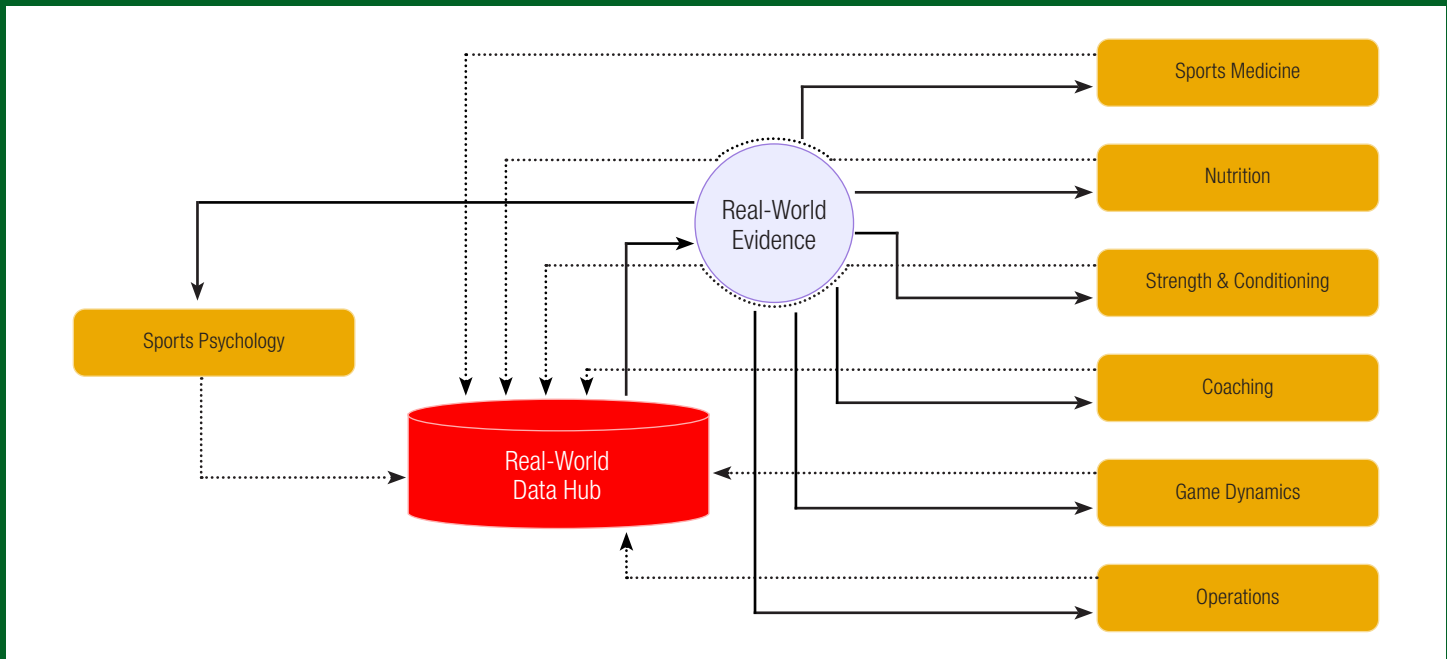


Figure 3. Real-World Evidence is the feedback mechanism that allows all disciplines to increase the quality of their practice.

The views expressed are those of the authors and do not necessarily reflect the position or policy of PepsiCo, Inc.

REFERENCES

- Blöbaum, P., P. Götz, K. Budhathoki, A.A. Mastakouri, and D. Janzing (2022). DoWhy-GCM: An extension of DoWhy for causal inference in graphical causal models. *ArXiv Preprint ArXiv:2206.06821*.
- Boyer, C.W., I.E. Lee, and M.S. Tenan (2022). All MCIDs are wrong, but some may be useful. *J. Orthop. Sports Phys. Ther.* 52:6.
- Ding, P., and F. Li (2018). Causal inference: A missing data perspective. *Stat. Sci.* 33:214–237.
- Greifer, N. (2022). WeightIt: Weighting for covariate balance in observational studies. <https://ngreifer.github.io/WeightIt/>
- Ho, M., M. van der Laan, H. Lee, J. Chen, K. Lee, Y. Fang, W. He, T. Irony, Q. Jiang, X. Lin, Z. Meng, P. Mishra-Kalyani, F. Rockhold, T. Song, H. Wang, and R. White (2023). The current landscape in biostatistics of real-world data and evidence: Causal inference frameworks for study design and analysis. *Stat. Biopharmaceut. Res.* 15:43–56.
- Hume, D. (1993). *An enquiry concerning human understanding: With Hume's abstract of a treatise of human nature and a letter from a gentleman to his friend in Edinburgh* (E. Steinberg, Ed.; Second Ed.). Hackett Publishing Company, Inc.
- Impellizzeri, F.M., P. Menaspà, A.J. Coutts, J. Kalkhoven, and M.J. Menaspà (2020a). Training load and its role in injury prevention, Part I: Back to the future. *J. Athl. Train.* 55:885–892.
- Impellizzeri, F.M., A. McCall, P. Ward, L. Bornn, and A.J. Coutts (2020b). Training load and its role in injury prevention, Part 2: Conceptual and methodologic pitfalls. *J. Athl. Train.* 55:893–901.
- Impellizzeri, F.M., M.S. Tenan, T. Kempton, A. Novak, and A.J. Coutts (2020c). Acute:Chronic workload ratio: Conceptual issues and fundamental pitfalls. *Int. J. Sports Physiol. Perform.* 15:907–913.
- Impellizzeri, F., S. Woodcock, A.J. Coutts, M. Fanchini, A. McCall, and A. Vigotsky (2020d). Acute to random workload ratio is 'as' associated with injury as acute to actual chronic workload ratio: Time to dismiss ACWR and its components. *ArXiv Preprint* <https://doi.org/10.31236/osf.io/e8kt4>
- Impellizzeri, F.M., S. Woodcock, A.J. Coutts, M. Fanchini, A. McCall, and A.D. Vigotsky (2021). What role do chronic workloads play in the acute to chronic workload ratio? Time to Dismiss ACWR and Its underlying theory. *Sports Med.* 51:581–592.
- Institute of Medicine (1997). *The computer-based patient record: An essential technology for health care*, Revised Edition. National Academies Press.
- Kalkhoven, J.T., M.L. Watsford, A.J. Coutts, W.B. Edwards, and F.M. Impellizzeri (2021). Training load and injury: Causal pathways and future directions. *Sports Med.* 51:1137–1150.
- Kappen, T.H., W.A. van Klei, L. van Wolfswinkel, C.J. Kalkman, Y. Vergouwe, and K.G.M. Moons (2018). Evaluating the impact of prediction models: Lessons learned, challenges, and recommendations. *Diagnost. Prognost. Res.* 2:11.
- Losciale, J.M., G.S. Bullock, G.S. Collins, A.J.H. Arundale, T. Hughes, N.K. Arden, and J.L. Whittaker (2023). Description, prediction, and causation in sport and exercise medicine research: Resolving the confusion to improve research quality and patient outcomes. *J. Orthop. Sports Phys. Ther.* 53:381–387.
- Mayer, I., P. Zhao, N. Greifer, N. Huntington-Klein, and J. Josse (2023). CRAN task view: Causal inference. *Comprehensive R archive network (CRAN)*. <https://CRAN.R-project.org/view=CausalInference>
- MPU-9150. (2024). TDK InvenSense. Retrieved January 16, 2024, from <https://invenSense.tdk.com/products/motion-tracking/9-axis/mpu-9150-2/>
- Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect. *Mathem. Model.* 7:1393–1512.
- Sant'Anna, P.H.C., and J. Zhao (2020). Doubly robust difference-in-differences estimators. *J. Econometr.* 219:101–122.
- Schertz, A.R., S.A. Smith, K.M. Lenoir, and K.W. Thomas (2023). Clinical impact of a sepsis alert system plus electronic sepsis navigator using the epic sepsis prediction model in the emergency department. *J. Emerg. Med.* 64:584–595.
- Senn, S. (2013). Seven myths of randomisation in clinical trials. *Stat. Med.* 32:1439–1450.
- Sharma, A., and E. Kiciman (2020). DoWhy: An end-to-end library for causal inference. *ArXiv Preprint ArXiv:2011.04216*.
- Shpitser, I., K. Mohan, and J. Pearl, J. (2015). Missing data as a causal and probabilistic problem. *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*, 802–811.

- Tenan, M.S. (2023a). Asking the right question is key to getting a valuable answer. *J. Orthop. Sports Phys. Ther.* 53:726.
- Tenan, M.S. (2023b). Missing data in sport science: A didactic example using wearables in American football. *Sports Med.* 53:1109-1116.
- Tenan, M.S., and C.W. Boyer (2023). The minimal clinically important difference: Letter to the editor. *Am. J. Sports Med.* 51:NP51–NP52.
- Tenan, M., and B. Alejo (2024). Athlete health and human performance will not improve without transdisciplinary collaboration and data sharing in elite sport. *SportRxiv*. <https://doi.org/10.51224/SRXIV.336>
- Textor, J., B. van der Zander, M.S. Gilthorpe, M. Liškiewicz, and G.T. Ellison (2016). Robust causal inference using directed acyclic graphs: The R package 'dagitty'. *Int. J. Epidemiol.* 45:1887–1894.
- U.S. Food and Drug Administration (2023). Framework for FDA's real-world evidence program. <https://www.fda.gov/media/120060/download>
- Zhou, Q., Z. Chen, Y. Cao, and S. Peng (2021). Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: A systematic review. *Npj Dig. Med.* 4:154.