International Journal of Sports Physiology and Performance, 2017, 12, 704-706
https://doi.org/10.1123/ijspp.2017-0137

Human Kinetics
INVITED COMMENTARY

Prediction: The Modern-Day Sport-Science and Sports-Medicine "Quest for the Holy Grail"

Alan McCall, Maurizio Fanchini, and Aaron J. Coutts

In high-performance sport, science and medicine practitioners employ a variety of physical and psychological tests, training and match monitoring, and injury-screening tools for a variety of reasons, mainly to predict performance, identify talented individuals, and flag when an injury will occur. The ability to "predict" outcomes such as performance, talent, or injury is arguably sport science and medicine's modern-day equivalent of the "Quest for the Holy Grail." The purpose of this invited commentary is to highlight the common misinterpretation of studies investigating association to those actually analyzing prediction and to provide practitioners with simple recommendations to quickly distinguish between methods pertaining to association and those of prediction.

Keywords: research and development, evidence, decision making, quality control, innovation

In high-performance sport science and medicine, practitioners employ a variety of physical and psychological tests, training and match monitoring, and injury-screening tools for a variety of reasons, mainly to predict performance, identify talented individuals, and flag when an injury may occur. The ability to "predict" outcomes such as performance, talent, or injury is arguably sports science and medicine's modern-day equivalent of the "Quest for the Holy Grail." The popularity of this topic is demonstrated through its prevalence in social-media discussions by both sport-science and sports-medicine academics and practitioners, with proponents both for and against. Much of the recent debate on prediction, particularly in the injury domain, stems from a review[1] by Professor Roald Bahr,[1] who posits that such prediction tests and tools "do not work—and probably never will." Perhaps the confusion and debate around prediction is linked to the terminology used by authors regarding a mismatch between their statistical modeling and subsequent interpretation of findings (ie, analyzing association and interpreting this as prediction). With such confusion, lay readers can therefore be misled into thinking that a test or tool is predictive, when in fact it is actually associative. To illustrate this point, we performed 3 separate literature searches using PubMed and the keywords *performance* AND *prediction, talent* AND *prediction,* and *injury* AND *prediction*. We then screened the abstracts of returned articles to check whether statistical outcomes actually reflected an analysis of predictive ability. After the *performance* AND *prediction* search, 61 articles were returned but only 14 (23%) reported using predictive analyses. Similarly, the *talent* AND *prediction* search returned 10 articles, of which only 1 (10%) included a mention indicative of predictive analysis. Finally, the search for *injury* AND *prediction* revealed that only 19 (35%) out of 55 articles (see Figure 1) had reported statistical modeling reflecting predictive analyses. The remaining studies in each domain had used statistical approaches pertaining to association.

Potential Consequences for Sport-Science and Sport-Medicine Practitioners

In high-performance sports, a major responsibility of science and medicine practitioners is to provide recommendations to key stakeholders such as players, coaches, and senior administrators that are subsequently used to inform important decisions. Indeed, these decisions can directly affect athlete performance, selections, or signings in addition to player availability for training and match play. Due to the potential impact of the advice provided, care must be taken to ensure that our own understanding of the information being provided is correct and the manner in which the information is delivered to these stakeholders is appropriate.

For example, if we examine the case of providing advice on injury prediction, some reports in the literature provide thresholds or cutoff values for various load metrics[2] and Functional Movement Screen scores[3] that indicate an increased risk of injury. Both these metrics have found popularity with practitioners; unfortunately, however, such indicators often show disappointing performance when used to identify individual players who will be injured. This may result in a misclassification of players (eg, indicating that a player will incur an injury and he or she does not) or, in the worst case, in a loss of "credibility" of the indicator itself. Take the common case from team sports where coaches are often alerted by the sport-science or medical staff to the "flagging" of players who are deemed to be at increased injury risk after objective load analysis that showed that the players' training loads had exceeded preset cutoff thresholds. In such cases, coaches often choose to continue to train intensively—despite the risk warning—most often without subsequent injury. In these cases, where advice is not followed and players avoid injury, the resultant effect is that coaches may be less likely to trust in future recommendations. It is due to examples like these that practitioners need to be aware of the importance of understanding the statistical underpinning behind the analysis

McCall is with the Medical Dept, Arsenal Football Club, London, UK. Fanchini is with the Sport Science and Performance Dept, US Sassuolo, Sassuolo, Italy. Coutts is with the Faculty of Health, University of Technology Sydney (UTS), Sydney, Australia. Address author correspondence to Alan McCall at amccall@arsenal.co.uk.

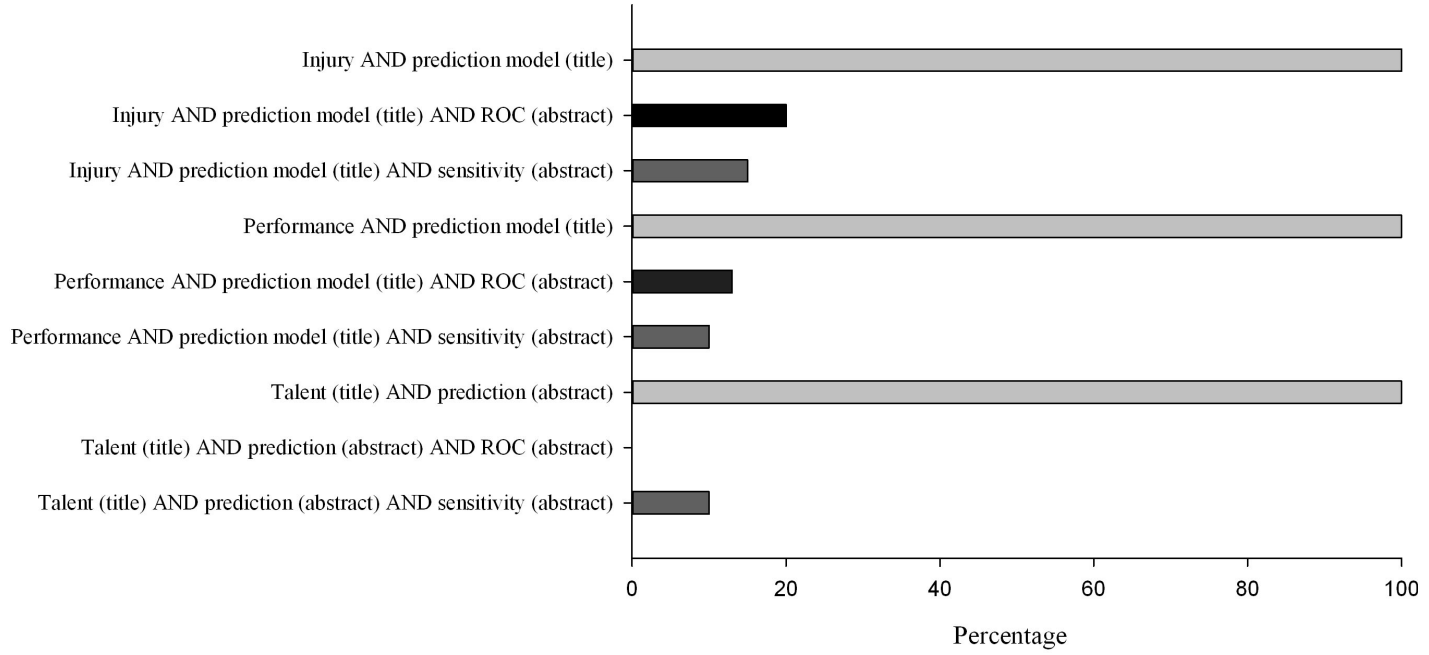

Figure 1 — Frequency of the words *ROC* (19%) and *sensitivity* (15%) in the abstracts of studies that included the words *performance* OR *talent* OR *injury* AND *prediction model* in the title (considered as reference 100%).

they are using and also the way in which they communicate their recommendations.

Association Versus Prediction

The absence of a clear definition of terms in the sport-science and -medicine domain and misuse of these terms have created a lack of understanding of the difference between association, or explanatory power, and predictive power.[4] Explanatory power (association) can be important in identifying a population who may perform well in a given competition, might turn into the next big star, or may perhaps be susceptible to injury. This provides us with evidence to implement a specific physical or psychological test or screening and monitoring tool. Predictive power allows us to classify at the individual level—this athlete will perform well, he or she will obtain a professional contract, or he or she will incur an injury—with a good level of accuracy. It is beyond the scope of this commentary to go into in-depth detail of each of the definitions and interpretations of statistical outcomes for association versus prediction. However, we would like to provide some specific and simple insights on the distinction between association and prediction and provide some key terms (Figure 2) that practitioners should look out for to quickly distinguish between explanatory and predictive abilities of physical, psychological, talent, or screening tests and monitoring tools.

Association

Studies investigating association can help us understand why a particular outcome occurs (explanatory studies) and may be due to direct or indirect causation.[5] Such studies are typically concerned with investigating whether there is a general relationship where 1 variable provides information about another,[5] for example, a higher fitness level is related to a better race performance. *Correlation* is

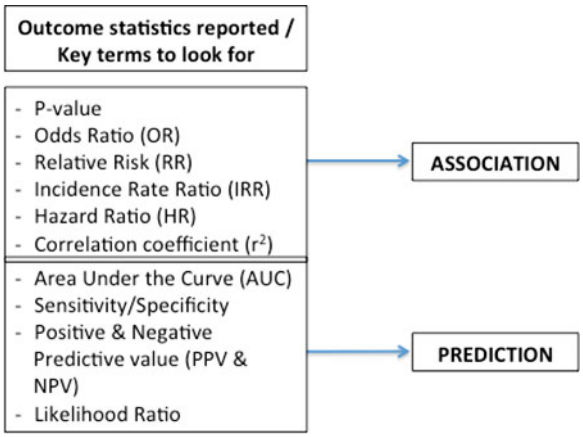

Figure 2 — Common statistical terms used to distinguish between associations and prediction. This list can be used to reduce confusion and help practitioners interpret methods and better understand how to apply them.

also often used interchangeably with *association* or *dependence;* however, while *dependence* is synonymous with *association, correlation* is different in that it implies specific types of association such as monotone trends or clustering but not causation.[5] If a significant association is identified between 1 or more factors and an outcome, it may be tempting to conclude that these can be used to predict.[1] However, even strong statistical associations do not necessarily imply that the factor can discriminate between individuals likely to incur a specific outcome and those who do not.[5]

Some of the most commonly reported outcome statistics for association include the *P* value, a correlation coefficient (r^2), and, in the case of injury studies, odds ratio (OR),[6] relative risk (RR),

risk difference, incidence-rate ratio (IRR), incidence-rate difference, and hazard ratio (HR).[7] If a research paper displays any of these statistical outcomes, it suggests that the authors have investigated association. As an example of the mismatch between association and prediction,[6] in the injury domain, an epidemiology article by Pepe et al[6] showed that data with a significant association and odds ratio of 3.0 (ie, implying a 3-times higher risk) would not be sufficient to conclude that the factor was accurate and predictive. Rather, in the authors' example, the OR would need to be ≥16.0 to have some acceptable level of predictive accuracy. It is important to note that this is not a rule of thumb but, rather, an example of the general concept that association does not necessarily imply prediction.

Prediction

In an ideal scenario, physical and psychological tests, and talent-identification and screening or monitoring tools would be able to detect all athletes who go on to, for example, perform well, obtain a professional contract, or incur an injury. While a test or tool with such 100% predictive accuracy is highly unlikely, those that can identify more athletes who go on to achieve or incur a specific outcome given a certain profile would provide confidence that implementing that specific test or tool would provide acceptable predictive ability on which to base confident recommendations. Statistical outcomes or terms that are used to explain the predictive ability include

- Sensitivity (the ability of a test to correctly identify an outcome) and specificity (the ability of a test to correctly identify no outcome).[8]
- Positive and negative predictive value. Positive predictive value—how likely it is that an individual will incur a specific outcome given a positive test result. Negative predictive value—how likely is it that the individual does not incur an outcome given that the test result is negative.[8]
- Likelihood ratio—how much more likely it is that an individual who tests positive will incur a specific outcome.[8]
- Receiver operating characteristics (ROC) curves—a plot of false positives against true positives for all cutoff values.[8]

Recommendations

In the sport-science and -medicine domain, there is clearly a mismatch between methods, analyses, and use of terms pertaining to association and prediction. Both have very different meanings, and neither should be confused with the other. In this commentary, we have attempted to provide some clarification between these 2 distinct entities and provide practitioners with a simple guide to distinguish between when a marker is associated with an outcome and when it can actually predict an outcome. Here is our key take-home message to practitioners and researchers:

- Explanatory power (association) and predictive power are different qualities, but both are important to interpret the efficacy of performance, talent-identification, and injury-prevention strategies—strong association measures provide evidence to support implementation, while higher predictive capacity will allow classification at the individual level.
- Practitioners: If a research paper indicates that the analysis provides "predictive abilities" of an important outcome variable, it should be checked carefully for methods that can infer predictive capabilities rather than those that examine association.
- Researchers should use correct terminology to appropriately differentiate between association and prediction.

References

1. Bahr R. Why screening tests to predict injury do not work—and probably never will. . . .: a critical review. *Br J Sports Med*. 2016;50:776–780. PubMed doi:10.1136/bjsports-2016-096256
2. Soligard T, Schwellnus M, Alonso JM, et al. How much is too much? (part 1)" International Olympic Committee consensus statement on load in sport and risk of injury. *Br J Sports Med*. 2016;50:1030–1041. PubMed doi:10.1136/bjsports-2016-096581
3. Chorba RS, Chorba DJ, Boullon LE, Overmyer CA, Landis JA. Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *N Am J Sports Phys Ther*. 2010;5(2):47–54. PubMed
4. Shmueli G. To explain or to predict? *Stat Sci*. 2010;25:289–310. doi:10.1214/10-STS330
5. Altman N, Krzywinski M. Association, correlation and causation. *Nat Methods*. 2015;12:899–900. PubMed doi:10.1038/nmeth.3587
6. Pepe MS, Janes M, Longton G, et al. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidem.* 2004;159:882–890. PubMed doi:10.1093/aje/kwh101
7. Nielsen RØ, Malisoux L, Møller M, Theisen D, Parner ET. Shedding light on the etiology of sports injuries: a look behind the scenes of time-to-event analyses. *J Orthop Sports Phys Ther*. 2016;46:300–311. PubMed doi:10.2519/jospt.2016.6510
8. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. *Contin Educ Anaesth Crit Care Pain*. 2008;8:221–223. doi:10.1093/bjaceaccp/mkn041