Red, Amber, or Green? Athlete Monitoring in Team Sport: The Need for Decision-Support Systems

Robertson, Samuel, Bartlett, Jonathan D and Gastin, Paul B (2017) Red, Amber, or Green? Athlete Monitoring in Team Sport: The Need for Decision-Support Systems. International Journal of Sports Physiology and Performance, 12 (Suppl 2). 73 - 79. ISSN 1555-0265

Red, amber or green? Athlete monitoring in team sport: the need for decision support systems

Robertson S[1], Bartlett, J.D[1], Gastin, P.B[2]

[1]Institute of Sport, Exercise and Active Living (ISEAL),
Victoria University,
Melbourne,
Australia.

[2]Centre for Sport Research,
School of Exercise and Nutrition Sciences,
Deakin University,
Burwood, Victoria, Australia

Running head: Decision support systems in team sports

Address for correspondence:
Dr Sam Robertson
Institute of Sport, Exercise and Active Living (ISEAL)
Victoria University
Victoria
Australia
Tel: +61 396 806 151
Email: sam.robertson@vu.edu.au

Word Count: 3396
Abstract Word Count: 156
Number of Tables: 1
Number of Figures: 4

Abstract

Decision support systems are used in team sport for a variety of purposes including evaluating individual performance and informing athlete selection. A particularly common form of decision support is the traffic light system, where colour coding is used to indicate a given status of an athlete with respect to performance or training availability. However despite relatively widespread use, there remains a lack of standardisation with respect to how traffic light systems are operationalised. This paper addresses a range of pertinent issues for practitioners relating to the practice of traffic light monitoring in team sports. Specifically, the types and formats of data incorporated in such systems are discussed, along with the various analysis approaches available. Considerations relating to the visualisation and communication of results to key stakeholders in the team sport environment are also presented. In order for the efficacy of traffic light systems to be improved, future iterations should look to incorporate the recommendations made here.

Key words: load, training, physical performance, injury

Introduction

Decision support systems are computer-based information systems that provide objective evidence relating to the decision-making of organisations.[1] Such systems utilise historical data to generate a recommendation or assessment to a user, with the decision often provided based on output generated by a software-based algorithm. [2.3] In sport, decision support systems have been used for purposes such as tournament scheduling, [4] evaluating athlete performance [5] and informing team selection. [6] A number of challenges are required to be overcome in order for decision support systems to provide ongoing value to organisations. These include a willingness of users to accept and act on findings/recommendations, appropriate integration of the system into the organisation's workflow as well as ensuring consistent use by practitioners.[7] Although evidence supporting their use is to date equivocal (see [8-9] for examples of unsuccessful implementations), relative success in fields such as medicine [2,3] make decision support systems an attractive proposition for sporting organisations in managing recent increases in data generation.

In team sports, one form of decision support, 'traffic light systems' are becoming more popular in their use to inform and support the decisions of practitioners. Although the nature of these decisions may vary, they often relate to the type and level of training an athlete is to undertake, or their availability to participate in competition. Also used to monitor student progress in education, [10] traffic light systems function by flagging red, amber or green, thereby providing a rapid insight into how different from the norm a daily score is for a given measurement. For instance, green may be interpreted as things should continue as per normal, amber suggests caution that if left unattended could pose a risk, whilst red raises an alarm and indicates action is required in order to bring the response back closer to the norm. Considering the constraints, time-pressures and challenges that practitioners face in the day-to-day fast paced environment of high performance sport, the ease of application, visual appeal and translational ability of the traffic light approach make them an attractive option in applied sporting environments. Nevertheless, evidence of their basis as an objective decision support system is scarce.

In performance sport contexts, measurements used in the traffic light systems are often derived directly from the athlete (both subjective and objective data), with the evidence base built using historical data. Types of data considered by practitioners using traffic light systems include self-reported athlete wellness, [11] musculoskeletal screening scores, [12] training load, [13] fitness and fatigue [14] and physiological testing/benchmarking. [15] Typically, users will use this information to adjust training programs and/or treatment in an effort to avoid undertraining/overtraining, reduce the likelihood of injury/illness incidence and determine the effectiveness of training programmes to ensure maintenance of performance. [16]

Validating the decision not to train

One of the most common outputs of traffic light systems used in a decision support context relates to a determination on the volume and intensity of training an athlete will undertake for a given session (or period of time). A common issue with traffic light systems is that it is often not clear what is used to validate the decision to restrict an athlete's training. A number of problems arise when these systems are attempted to be validated, especially when using either injury prevention and/or performance-based metrics.

For instance, individual player performance as a construct has proved difficult to define in team sports; due largely in part to the multifaceted nature of game play. [17] Further, considerable individual performance variation is likely to be observed depending on what is occurring at the team level. [18] Using injury incidence as a measure is no less problematic. As injury occurs at such a low incidence relative to the total number of sessions/matches players participate in, any decisionsupport system for training availability is almost certainly destined to be conservative in its approach. The implications of this is that athletes may be missing sessions that they may participate in without adverse effect, thereby exerting a flow on effect to performance.

Another fundamental problem with both forms of data is that access to injury or individual performance information is not available prior to the training session or match of interest. As such, traffic light systems in their current format are limited as a *predictive* tool. All they can do is (partially?) *explain* why an injury did or did not occur, or why a player did or did not perform to their usual standard (see Shmueli, 2010 [19] for a description of the differences between explanation and prediction). Of course that is of limited use to a practitioner making decisions on the athlete's availability. Further, in order to make an accurate prediction based on historical data, a large number of data points are required, which necessitates a long lead in time and therefore limits those in the early stages of implementing a monitoring program. Consequently, proxies for under-recovery or susceptibility to injury, derived from the literature and or practitioner experience, are used as early warning signs for decision making with the intent to mitigate the risk of an undesirable outcome. So how can the evidence behind traffic light methods be improved, without losing the practical qualities that make them so popular in the first place?

Despite the abovementioned methodological issues pertaining to injury prediction, nonetheless there has been a range of research investigating the relationships between the incidence of injury with player wellness, [20] musculoskeletal screening test scores, [21] fitness levels [22] and training load. [23] As many elite team sport athletes are being assessed in some capacity on an almost daily basis, the ability to analyse these athletes at the individual level has never been more feasible. The rise in popularity of data mining in sport [24] has also allowed for non-linear relationships between load metrics [25] and injury/performance [26] at the inter- and intra-individual level to be better elucidated and visualised. Consequently, it is evident that in order to obtain better answers to these questions, both large data sets and complex analyses are required. The benefits of improving objective decision support systems such as traffic lights are important to both the performance and financial health of sports organisations. Below we provide some guiding principles for practitioners that can help to improve the efficacy of the approach.

Step 1 – What type of data should be considered in the traffic light system?

Collecting, maintaining and analysing the types of data mentioned earlier has in many sports become a full-time job in itself. For the sports practitioner, reducing the volume of data to consider in making a decision on an athlete's training availability or determining their injury risk can greatly increase work efficiency. A pertinent example relating to data reduction can be drawn from Bartlett et al., 2016 [25] where the authors investigated the relationships between commonly collected training metrics and the session RPE response of athletes at a professional Australian Rules Football club. It was observed that the relationship between the distance covered by an individual in a session and the training time was almost a perfect one. Consequently, as is standard practice in relationship modelling [27] one of these metrics (training time) was removed from the model; in this case without any meaningful adverse effect on accuracy. Of course the data reduction could have instead been applied to the second metric. The duration of a training session is easier to measure than the distance an athlete has covered, which is of practical use to those without access to GPS or other player tracking systems. Amongst other benefits, the practice of data reduction helps to improve model parsimony, which in the event of multiple solutions existing to a single problem the simplest should be chosen (see Coutts' 2014 editorial on the relevance of Occam's Razor to sport science [28]).

So which considerations, in addition to the above, can help the practitioner to arrive at a decision on what to look at and what to leave out when designing their traffic light system? Figure 1 shows five main factors that should be considered by those working in high-level team sport environments, with an outline of each provided below.

**** INSERT FIGURE 1 ABOUT HERE ****

Validity of measurements and data entry

The validity of a measure or the strength of relationships between variables of interest should primarily inform the decision support system. For instance, concurrent validity refers to the extent to which a metric relates to an alternate, previously validated measure of the same construct administered at the same time (e.g., assessing training time and distance covered as per the example above). [29] Convergent validity relates to the extent to which two metrics that theoretically should be related to each other are, indeed, related to each other (e.g., an increase in heart rate as the intensity of a training session also increases). [30] As an example, if certain information relating to activities the player undertook the night before a training session showed limited relationship with the athlete's risk of injury or performance in competition, it would not make sense to measure it for that purpose. In the context of designing a traffic light system, an assessment of these forms of validity is essentially another form of the data reduction process. Whilst these and other forms of validity are not always measurable or relevant for all metrics included in the traffic light system, they should be assessed wherever possible. Alternately, a review of the literature can inform the approach, via evaluation of the suitability of both objective and subjective measures [31] and consideration of issues related to sport context and implementation. [32]

Of equal importance is consideration of the reliability of a traffic light system. Some level of random error is inherent and to be expected in any measurement. From a systematic perspective if technology shows meaningful differences between different devices, units or software versions or the methods of obtaining self-report data change, [33] then reliability will in turn also be affected. Therefore, this within- and between-athlete variability should be accounted for. With sufficient data, the latter consideration can be overcome through the development of separate models for each individual athlete.

Data interpretation and decision-making consequences

In professional team sports, where decisions relating to training availability need to be made within 1-2 hrs of training commencing, the traffic light system needs to be easily and quickly interpretable. Whilst coaches are expected to be learned and experienced in their content area, they are typically not statistically trained. Consequently, more sophisticated data formats may require conversion before being communicated to coaches and other practitioners. For instance, raw data may need to be converted to a normalised score (e.g., a *z*-score) to allow historical intra-individual or within-team, sport or gender comparisons. [34] Often this will also entail some form of visualisation, which may also vary in nature depending on the preferences or learning styles of the intended audience. Delivery flexibility and the ability to generate visualisations rapidly are crucial in ensuring all stakeholders can interpret results for their given use. *Cost effectiveness*

The cost effectiveness of a system includes features such as burden, time and cost/benefit. In its most simple form, burden relates to the number of staff and the resources required in order to collect, clean, interpret and report the data used in the traffic light system. This includes both the start-up cost (e.g., hardware and software, data storage solutions) and daily operation of the measurement system. Many companies working in elite sport have aimed to provide user-friendly software in order to expedite this process. However, if metrics of interest are not reported by the accompanying software, then further post-hoc analysis of raw data may be required. Burden can also exist in the form of staff being required to undertake further training in order to complete the collection and analysis of data. This may also extend to their ability to understand and interpret any results derived from these analyses. In addition, the burden on the athlete should also be considered and minimised as much as possible. [32]

Closely linked with interpretability and burden, the time required to collect, interpret and report is paramount to a successful, useful and meaningful decision support system. How much time it takes to manage data and implement a decision support system (especially in the context where thousands of observations can be obtained in one week for a single team) dictates the success of a given system. For example, analysing a continuous trace of 10 Hz GPS data for each player for each training session can allow for interesting insights into the movement of athletes, however, it can be time consuming. Consequently, the extent to which gaining this insight provides added benefit to informing a decision comparative to the time spent on the analysis needs to be examined.

In high-level sport, the decision support system should be considered in relation to its cost and benefit so as to determine its efficacy and value to an overall programme. Beyond the more tangible benefits such as possible improvements in performance and reductions in injuries and illnesses, other benefits such as communication between staff and athletes, building knowledge within the programme and supporting athlete self-management are all possible benefits of developing monitoring and decision support systems. [35]

Step 2 – In which format should traffic light system data be analysed?

The format in which data are analysed can alter the nature of the inferences made, irrespective of the analysis approach implemented. Whilst ideal where possible, the analysis of unconverted, raw data can result in substantially varied baseline values across different athletes, making between-individual comparisons challenging. As a result, in team sport settings *z*-scores continue to experience popularity based on their ability to allow for the standardised position of an individual within a group or with reference to their own baseline data to be articulated. Expressing data as a percentage change from baseline addresses this by allowing for the within-individual differences to be interpreted within context of others in the group. However where large within-individual variation exists in data, or where values are close to zero, artificially high values may result. Furthermore, the conversion of the data to a relative format may be less interpretable to some stakeholders. So which format should be used in traffic light systems? One of the key considerations in making this selection relates to the decision to focus on the individual or the group.

Individual vs group

The importance of considering the individual within a team has received increased attention of late. [25,36] However, it is well established that analysing larger numbers of athletes together can allow for greater inferences to be made relating to the sample population of interest, thereby increasing the confidence in such findings. [37] The approach taken is likely to depend on the question at hand. For instance, when considering a team sport training scenario, a typical approach for practitioners would be to use within-group comparisons and literature to determine the typical responses for a given training period. Figure 2 provides an example of a practical problem with this approach. The figure shows the average weekly training load for 39 players from an Australian Football League club over the course of a month during the season. Both weekly mean values and the variance differs substantially between players, thus the need for an individual approach is self-evident.

**** INSERT FIGURE 2 ABOUT HERE ****

Step 3 – How is traffic light data analysed and interpreted?

The consideration on whether to assess at the individual or group level will also have direct implications for the types of analysis undertaken. A range of commentaries and resources exist relating to the various approaches available to sport scientists. [38-39] However, perhaps the two most pervasive topics relate to determining what constitutes a meaningful change and the accounting for repeated measures in analyses.

Accounting for repeated measures

Most traffic light systems will incorporate repeated measures data. Many of these measurements occur on a daily basis; aggregated weekly or monthly values along with rolling averages are often then calculated to describe trends in the data as well as facilitate analysis. However, when group data is pooled without account for the dependency of repeated observations on the same individual/s, relationships between variables of interest can be overstated. [40] Generalised linear models such as linear mixed models and generalised estimating equations can account for this issue in the modelling process, however whilst relatively common in research, their use may require upskilling of practitioners. Although machine learning algorithms can allow for any potential non-linearity both between and within individuals to be uncovered, most approaches assume independence between observations. The development of models for each individual has been used as another method of avoiding the repeated measures issue. [25] However this will be more time consuming when large player numbers are involved. Further, in instances where limited data exists obtaining a well-fit model also may become a challenge.

Identifying a meaningful change

In sporting terms, it is important to identify what longitudinal changes in responses (i.e., to training) are meaningful, above and beyond 'normal' or random variability. Given the historical records of data now available to many professional teams, a number of approaches have been proposed in the literature to determine what constitutes a 'meaningful' change (often referred to as responsiveness). [41] The standard deviation (SD), effect size, smallest worthwhile change (SWC), coefficient of variation (CV) and risk ratio are just some examples of metrics used to determine this meaningful change. However, unsurprisingly each measure will provide different outputs.

**** INSERT TABLE 1 & FIGURE 3 ABOUT HERE ****

In the Figure 2 example, despite similar weekly mean training loads, the distinct levels of variance between each player results in substantially different thresholds for each player, therefore also resulting in different flags (Table 1). Table 1 shows an example of a traffic light system in operation. In this illustration a < 0.3 x CV of the weekly load is considered 'green' (a 'small' effect/difference), a 0.3 to 0.9 x CV equates to 'amber' (a 'moderate' difference) and a 0.9 to 1.6 x CV is 'red' (a large difference). It should be noted that this approach represents only one example and a variety of others experience use in the field. Such systems have clear implications for decision-making between individuals within a group. Clearly if one traffic light system was calibrated using a CV approach, and another using the SD, then the measurement and observation would be different, therefore, triggering a different course of follow up action. In complement to Figure 2, Figure 3 displays the weekly training load for the two athletes (#5 and #13) shown in Figure 2 and Table 1. An example traffic light system is shown for the month (incorporating the weekly load data) using the same traffic light thresholds discussed above The differences between the two outputs are clearly visible.

Whether considering the data from a training prescription or injury prevention perspective, given the noted differences for each player, it is apparent that differentiated loading approaches should be prescribed for each. For example, Player #13 shows large variation in their monthly load – due in part to the high load obtained in week 1. In the example, this has resulted in a decision to reduce the exposure to load in week 2; therefore, the system provides a red flag. Together, these 2 weeks demonstrate inconsistency in loading, possibly increasing the risk of injury/illness. [42] In rectifying this, closer attention (in the context of this example) should be placed on the absolute and relative changes in load so as to prescribe more consistent loading. In contrast, Player #5 demonstrates relative consistency in their loading (range 1250 AU – 1850 AU). As such, a red flag (a change of ~300 load units) may not pose any meaningful risk to injury/illness. Collectively, this shows a number of complexities and factors to consider when individualising training prescription in team sports. The system employed will thus require careful consideration of the relationships between each metric and those validity measures mentioned earlier in the article.

Step 4 – Communicating the findings

Increasing the transparency in which data is displayed in scientific research has received considerable attention of late. [43] Figures which are able to display the response of the individual within the group have become more sophisticated as more advanced visualisation packages are available in commercial software. Figure 4 shows an example of how the same group means and standard deviations can be replicated using individual data, as well as how different tests of statistical significance change as a result of this differently distributed data. [43] This provides further support for visualisation of both the individual and group in order to understand the nature of the dataset. The great appeal of the traffic light approach is its ability to convey information visually in an intuitive and easily interpretable manner. The use of integrated plots, automated colour coding and conditional formatting, and visual flagging of outliers, anomalies and trends (both desirable and undesirable) provides regular feedback to the coach and support staff to guide daily decision making.

*** INSERT FIGURE 4 ABOUT HERE ****

The future

Given the considerable human and financial investment in the pursuit of success, and the ethical importance of looking after individuals in our care, high performance sport will continue to evolve in search of better ways to train and monitor athletes and to make decisions about how best to manage them to ensure both safety and success. The future will likely involve a mix of existing and new measurement approaches and technologies. However, to be most effective, and to provide a sound basis for decision support, all of the following will need to be developed and enhanced:

- Robust selection of athlete monitoring measures, with due consideration to issues related to validity, reliability, data reduction and athlete burden.
- Establishment of evidence-based guidelines related to the determination of benchmarks and baselines and the subsequent boundaries used for categories (e.g., red, amber, green) within a decision support system.
- Development of database and dashboard software to enhance data management and visualisation.
- Application and exploration of analytic approaches to large datasets that account for longitudinal repeated measures data. Evaluation of multiple analysis approaches (i.e., machine learning vs linear models) to the same datasets.
- Improved integration within multidisciplinary teams and the upskilling of staff and coaches in sport science and data analysis.
- The strategic implementation of research and innovation within high performance programmes, including rigorous data collection and question driven projects.
- The pursuit of research that encourages practitioners and researchers to answer questions through analysis of larger scale datasets facilitated through greater collaboration across clubs, leagues and sports.

References

1. Sprague Jr RH. A framework for the development of decision support systems. *MIS Q*. 1980;1:1-26.

2. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, Sam J, Haynes RB. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA*. 2005;293:1223-1238.

3. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ*. 2005;330:765-772.

4. Kostuk KJ, Willoughby KA. A decision support system for scheduling the Canadian Football League. *Interfaces*. 2012;42:286-295.

5. Calder JM, Durbach IN. Decision support for evaluating player performance in rugby union. *Int J Sports Sci Coach*. 2015;10:21-37.

6. Ofoghi B, Zeleznikow J, MacMahon C, Dwyer D. Supporting athlete selection and strategic planning in track cycling omnium: a statistical and machine learning approach. *Inform Sciences*. 2013;233:200-213.

7. Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *JAMA*. 1998;280:1339-1346

8. Lewis G, Sharp D, Bartholomew J, Pelosi AJ. Computerized assessment of common mental disorders in primary care: effect on clinical outcome. *Fam Pract*. 1996;13:120-126.

9. Bogusevicius A, Maleckas A, Pundzius J, Skaudickas D. Prospective randomised trial of computer aided diagnosis and contrast radiography in acute small bowel obstruction. *Eur J Surg*. 2002;168:78-83.

10. Given K, Hannigan A, McGrath D. Red, yellow and green: what does it mean? How the progress test informs and supports student progress. *Med Teach*. 2016;14:1-8.

11. Gastin PB, Meyer D, Robinson D. Perceptions of wellness to monitor adaptive responses to training and competition in elite Australian football. *J Strength Cond Res*. 2013;27:2518-2526.

12. Halson SL. Monitoring training load to understand fatigue in athletes. *Sports Med*. 2014;44:139-147.

13. Cross MJ, Williams S, Trewartha G, Kemp S, Stokes KA. The influence of in-season training loads on injury risk in professional rugby union. *Int J Sports Physiol Perform*. 2016;11:350-355.

14. Gabbett TJ, Domrow N. Relationships between training load, injury, and fitness in sub-elite collision sport athletes. *J Sports Sci.* 2007;25:1507-1519.

15. Gastin PB, Meyer D, Huntsman E, Cook J. Increase in injury risk with low body mass and aerobic-running fitness in elite Australian football. *Int J Sports Physiol Perform.* 2015;10:458-463.

16. Taylor K, Chapman D, Cronin J, Newton MJ, Gill N. Fatigue monitoring in high performance sport: a survey of current trends. *J Aust Strength Cond.* 2012;20:12-23.

17. Duch J, Waitzman JS, Amaral LA. Quantifying the performance of individual players in a team activity. *PloS One.* 2010;5:e10937.

18. Woods CT, Robertson S, Collier NF. Evolution of game-play in the Australian Football League from 2001 to 2015. *J Sports Sci.* 2016;1-9.

19. Shmueli G. To explain or to predict?. *Stat Sci*. 2010:289-310.

20. Krumrei K, Flanagan M, Bruner J, Durall C. The accuracy of the functional movement screen to identify individuals with an elevated risk of musculoskeletal injury. *J Sport Rehabil*. 2014;23:360-364.

21. Dorrel BS, Long T, Shaffer S, Myer GD. Evaluation of the Functional Movement Screen as an Injury Prediction Tool Among Active Adult Populations A Systematic Review and Meta-analysis. *Sports Health.* 2015:1941738115607445.

22. Gabbett TJ, Domrow N. Relationships between training load, injury, and fitness in sub-elite collision sport athletes. *J Sports Sci.* 2007;25:1507-1519.

23. Fowler P, Duffield R, Waterson A, Vaile J. Effects of Regular Away Travel on Training Loads, Recovery, and Injury Rates in Professional Australian Soccer Players. *Int J Sports Physiol Perform*. 2015;10:546:552.

24. Ofoghi B, Zeleznikow J, MacMahon C, Raab M. Data mining in elite sports: a review and a framework. *Meas Phys Educ Exerc Sci.* 2013;17:171-186.

25. Bartlett JD, O'Connor F, Pitchford N, Torres-Ronda L, Robertson SJ. Relationships Between Internal and External Training Load in Team Sport Athletes: Evidence for an Individualised Approach. *Int J Sports Physiol Perform*. 2016; DOI: 10.1123/ijspp.2015-0791

26. Bittencourt NF, Meeuwisse WH, Mendonça LD, Nettel-Aguirre A, Ocarino JM, Fonseca ST. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept. *Br J Sports Med.* 2016: doi:10.1136/bjsports-2015-095850

27. Johnson JB, Omland KS. Model selection in ecology and evolution. *Trends Ecol Evol.* 2004;19:101-108.

28. Coutts AJ, Zavorsky GS, Galy O, Pyne DB, Guy JH, Edwards AM. In the age of technology, Occam's razor still applies. *Int J Sports Physiol Perform*. 2014;9:741.

29. Currell K, Jeukendrup AE. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med.* 2008;38:297-316.

30. Robertson SJ, Burnett AF, Cochrane J. Tests examining skill outcomes in sport: A systematic review of measurement properties and feasibility. *Sports Med.* 2014;44:501-518.

31. Saw AE, Main LC, Gastin PB. Monitoring the athlete training response: subjective self-reported measures trump commonly used objective measures: a systematic review. *Br J Sports Med.* 2015: doi:10.1136/bjsports-2015-094758

32. Saw AE, Main LC, Gastin PB. Impact of Sport Context and Support on the Use of a Self-Report Measure for Athlete Monitoring. *J Sports Sci Med.* 2015;14:732-739.

33. Saw AE, Main LC, Gastin PB. Monitoring athletes through self-report: Factors influencing implementation. *J Sports Sci Med*. 2015;14:137-146.

34. Schmidt JD, Register-Mihalik JK, Mihalik JP, Kerr ZY, Guskiewicz KM. Identifying impairments after concussion: normative data versus individualized baselines. *Med Sci Sports Exerc*. 2012;44:1621-1628.

35. Saw AE, Main LC, Gastin PB. Role of a self-report measure in athlete preparation. *J Strength Cond Res*. 2015;29:685-691.

36. Gastin PB, Fahrner B, Meyer D, Robinson D, Cook JL. Influence of physical fitness, age, experience, and weekly training load on match performance in elite Australian football. *J Strength Cond Res*. 2013;27:1272-1279.

37. Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. *Multivariate Behav Res.* 2001;36:249-77.

38. Atkinson G, Nevill AM. Selected issues in the design and analysis of sport performance research. *J Sports Sci.* 2001;19:811-827.

39. Hopkins W, Marshall S, Batterham A, Hanin J. Progressive statistics for studies in sports *Med Sci Sports Exerc.*. 2009;41:3.

40. Williamson DS, Bangdiwala SI, Marshall SW, Waller AE. Repeated measures analysis of binary outcomes: applications to injury research. *Accid Anal Prev.* 1996;28:571-579.

41. Liang MH. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. *Med Care.* 2000;38:84-90

42. Gabbett TJ. The training—injury prevention paradox: should athletes be training smarter and harder?. *Br J Sports Med.* 2016;50:273-280.

43. Weissgerber TL, Milic NM, Winham SJ, Garovic VD. Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol*. 2015;13:e1002128.