



# The Case for Adopting a Multivariate Approach to Optimize Training Load Quantification in Team Sports

Dan Weaving<sup>1,2,3\*</sup>, Ben Jones<sup>1,2,4,5</sup>, Kevin Till<sup>1,2,4</sup>, Grant Abt<sup>3</sup> and Clive Beggs<sup>1</sup>

<sup>1</sup> Institute for Sport, Physical Activity and Leisure, Leeds Beckett University, Leeds, United Kingdom, <sup>2</sup> Leeds Rhinos Rugby League Club, Leeds, United Kingdom, <sup>3</sup> School of Life Sciences, University of Hull, Kingston upon Hull, United Kingdom, <sup>4</sup> Yorkshire Carnegie, Leeds, United Kingdom, <sup>5</sup> The Rugby Football League, Leeds, United Kingdom

**Keywords:** multivariate analysis, global positioning systems (GPS), training load, reductionism, external training load, orthogonal analysis

## OPEN ACCESS

### Edited by:

Billy Sperlich,  
Integrative and Experimentelle  
Trainingswissenschaft, Universität  
Würzburg, Germany

### Reviewed by:

Shaun Jamss McLaren,  
Teesside University, United Kingdom  
Lars Donath,  
University of Basel, Switzerland

### \*Correspondence:

Dan Weaving  
d.a.weaving@leedsbeckett.ac.uk

### Specialty section:

This article was submitted to  
Exercise Physiology,  
a section of the journal  
Frontiers in Physiology

**Received:** 22 June 2017

**Accepted:** 27 November 2017

**Published:** 12 December 2017

### Citation:

Weaving D, Jones B, Till K, Abt G and  
Beggs C (2017) The Case for  
Adopting a Multivariate Approach to  
Optimize Training Load Quantification  
in Team Sports.  
*Front. Physiol.* 8:1024.  
doi: 10.3389/fphys.2017.01024

Professional sports teams are investing substantial resources in monitoring the training load (TL) in their players in an attempt to achieve favorable training outcomes such as increases in performance and a reduction in negative outcomes such as injury. This investment is likely to increase as organizations explore the most recent developments in wearable technology that allow a wide variety of objective physiological and other measures to be collected concurrently and over long periods of time. The question of how all of this data can be used is one that many in our field are now asking (Foster et al., 2017). To answer this, we have to start with a definition of TL. Soligard et al. (2016) recently defined TL as:

“the sport and non-sport burden (single or multiple physiological, psychological, or mechanical stressors) as a stimulus that is applied to a human biological system (including subcellular elements, a single cell, tissues, one or multiple organ systems, or the individual)”

To quantify this construct, a common approach is to determine the ratio of a single measure across two moving-average time periods (e.g., acute- and chronic-training-load-ratio [A:C]). Suboptimal (either too high or low) TL is associated with an increased risk of injury (Hulin et al., 2016). However, while many TL methods (e.g., total distance, high-speed distance, session rating of perceived exertion [sRPE]) are collected, they are used individually as “predictor” variables in these analyses. Therefore, the initial consideration should be to determine the variable(s) that provide the most valid representation of the actual load imposed on each athlete.

Establishing the validity of a TL measure is typically examined through its agreement with a criterion which represents the true value. For example, the speed derived from a global positioning system (GPS) device is compared to that derived from a radar gun (Roe et al., 2016). In this instance, the confidence that the criterion measure represents the true value is high. In contrast, determining the validity of internal TL methods is problematic due to the limited physiological markers that are available in the field, and that there is no criterion method of measuring the internal TL. In addition, the definition highlighted previously (Soligard et al., 2016) demonstrates the complexity of the internal TL construct. Therefore, despite sRPE having been reported to correlate highly with Banister’s training impulse (TRIMP) ( $r = 0.75$ ) (Lovell et al., 2013) and Edward’s TRIMP ( $r = 0.70$ ) (Kelly et al., 2016) models, in these examples the shared variance is only 56 and 49%, respectively. This means that about half of the variance is unexplained. Are we therefore adopting a reductionist approach by assuming that by association, a single measure can capture the whole (true) internal TL imposed?

Physiological systems are complex, with many disparate factors affecting the outcomes of training. In essence, every bout of exercise/training imposes specific physiological, biomechanical,

and psychological demands which vary not only with the prescribed “dose” (i.e., sets, repetitions, duration etc) but also with the mode (e.g., strength training vs. sport-specific training) of exercise (Soligard et al., 2016; Cardinale and Varley, 2017). Therefore, it is unlikely that a single independent variable will be able to capture this complexity and provide a valid measure of TL (either internal or external) and consequently, a holistic representation of TL has been suggested (Cardinale and Varley, 2017). By taking a univariate approach, we are in danger of omitting valuable information that could contribute to explaining the relationships between the imposed TL, and changes in fitness/performance/injury. For example, it is common practice to collect multiple TL variables concurrently. Recent investigations have shown that a single TL variable is unable to capture a meaningful proportion of the variance provided by multiple internal and external TL variables, which is exacerbated by the mode of training (e.g., technical-tactical, high-intensity-interval-training, sprint-training) (Weaving et al., 2014, 2017). Therefore, as the internal TL is governed largely by the external TL, external TL measures are likely to contribute “surrogate” information about the internal TL imposed and provide information that can also relate to training outcomes (Oxendale et al., 2016). In data science terms, the information contained collectively in, and between, these variables, has great potential to inform and optimize our understanding of training dose-response relationships. However, appropriately unlocking this information (without statistical/mathematical violation) can be difficult. As the variables associated with TL are often strongly correlated, multicollinearity (i.e., the degree to which variables are similar to one another) is frequently a problem. In addition, because player cohorts are small, it is often the case that the number of measured variables can exceed the number of players. As such, TL datasets can pose a considerable challenge when using traditional techniques such as logistic and multiple linear regression, thereby limiting their applicability when adopting multivariate (rather than univariate) TL analyses. However, through the use of dimension reduction techniques such as principal component analysis (PCA) (Weaving et al., 2014, 2017) and single value decomposition (SVD) (Till et al., 2016), which are immune to multicollinearity issues, it is possible to capture the complexity of a system in just a few orthogonal composite variables (i.e., variables that provide

unique information). Because most of the variance in the system is captured in these orthogonal composite variables, it means that complex higher-dimensional systems can be represented on 2D and 3D scatter plots with minimal loss of information (Till et al., 2016). Furthermore, because the new variables are orthogonal it means that they are not correlated in any way, thus ensuring that they capture different attributes of the TL “system.” Single value decomposition and eigen-decomposition are at the heart of other useful data science techniques, such as partial least squares correlation analysis (PLSCA) (Beggs et al., 2016), which have great potential with respect to TL quantification. Rather than taking a conventional statistical approach, PLSCA utilizes the concept of shared information to gain new insights into the relationships between groups of variables (i.e., both predictor and response variables) in complex datasets. For example, using PLSCA, the relationship between multiple TL variables (e.g., total-distance, high-speed-distance, and s-RPE) and multiple “fatigue” variables can be investigated in a single analysis, allowing stronger inferences to be made of the “dose-response” nature of these broad theoretical constructs that we wish to represent.

Despite the perceived increases in computational demands placed on practitioners, the authors feel that this multivariate approach warrants further investigation, at least initially in research, given the importance of TL measures in optimizing the preparation of team-sport players. It is then envisaged that this approach could be integrated into athlete monitoring software platforms to “combine” unique aspects of information provided by multiple TL variables. Although developing our understanding of what individual TL measures represent is important (i.e., validity), it is hoped that multivariate approaches will further develop our knowledge of the dose-response nature of TL monitoring with important training outcomes such as the changes in fitness, performance, and injury risk.

## AUTHOR CONTRIBUTIONS

DW, GA, and BJ: conceptualized the idea, wrote the introduction, and rationale to the commentary. DW and CB: wrote the discussion of the analysis approach to multivariate data. GA, KT, BJ, DW, and CB: drafted the manuscript, revised critically for important intellectual content.

## REFERENCES

- Beggs, C. B., Magnano, C., Belov, P., Krawiecki, J., Ramasamy, D. P., Hagemeyer, J., et al. (2016). Internal jugular vein cross-sectional area and cerebrospinal fluid pulsatility in the aqueduct of sylvius: a comparative study between healthy subjects and multiple sclerosis patients. *PLoS ONE* 11:e0153960. doi: 10.1371/journal.pone.0153960
- Cardinale, M., and Varley, M. C. (2017). Wearable training-monitoring technology: applications, challenges and opportunities. *Int. J. Sports Physiol. Perform.* 12, S255–S262. doi: 10.1123/ijsp.2016-0423
- Foster, C., Rodriguez-Marroyo, J. A., and de Koning, J. J. (2017). Monitoring training loads: the past, the present, and the future. *Int. J. Sports Physiol. Perform.* 12(Suppl. 2), S22–S28. doi: 10.1123/ijsp.2016-0388
- Hulin, B. T., Gabbett, T. J., Lawson, D. W., Caputi, P., and Sampson, J. A. (2016). The acute:chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *Br. J. Sports Med.* 50, 231–236. doi: 10.1136/bjsports-2015-094817
- Kelly, D. M., Strudwick, A. J., Atkinson, G., Drust, B., and Gregson, W. (2016). The within-participant correlation between perception of effort and heart-rate-based estimations of training load in elite soccer players. *J. Sports Sci.* 34, 1328–1332. doi: 10.1080/02640414.2016.1142669
- Lovell, T. W., Sirotic, A. C., Impellizzeri, F. M., and Coutts, A. J. (2013). Factors affecting perception of effort (session rating of perceived exertion) during rugby league training. *Int. J. Sports Physiol. Perform.* 8, 62–69. doi: 10.1123/ijsp.8.1.62

- Oxendale, C. L., Twist, C., Daniels, M., and Highton, J. (2016). The relationship between match-play characteristics of elite rugby league and indirect markers of muscle damage. *Int. J. Sports Physiol. Perform.* 11, 515–521. doi: 10.1123/ijsp.2015-0406
- Roe, G., Darrall-Jones, J., Black, C., Shaw, W., Till, K., and Jones, B. (2016). Validity of 10 Hz GPS and timing gates for assessing maximum velocity in professional rugby union players. *Int. J. Sports Physiol. Perform.* 13, 1–14. doi: 10.1123/ijsp.2016-0256
- Soligard, T., Schweltnus, M., Alonso, J. M., Bahr, R., Clarsen, B., Dijkstra, H. P., et al. (2016). How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *Br. J. Sports Med.* 50, 1030–1041. doi: 10.1136/bjsports-2016-096581
- Till, K., Jones, B. L., Cogley, S., Morley, D., O'Hara, J., Chapman, C., et al. (2016). Identifying talent in youth sport: a novel methodology using higher-dimensional analysis. *PLoS ONE* 11:e0155047. doi: 10.1371/journal.pone.0155047
- Weaving, D., Jones, B., Till, K., Marshall, P., and Abt, G. (2017). Multiple measures are needed to quantify training loads in professional rugby league. *Int. J. Sports Med.* 38, 735–740. doi: 10.1055/s-0043-114007
- Weaving, D., Marshall, P., Earle, K., Nevill, A., and Abt, G. (2014). Combining internal- and external-training-load measures in professional rugby league. *Int. J. Sports Physiol. Perform.* 9, 905–912. doi: 10.1123/ijsp.2013-0444

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Weaving, Jones, Till, Abt and Beggs. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.