

Single-Item Self-Report Measures of Team-Sport Athlete Wellbeing and Their Relationship With Training Load: A Systematic Review

Ciara Duignan, PhD*†; Cailbhe Doherty, PhD, BSc*†;
Brian Caulfield, PhD, MSci, BSci*†; Catherine Blake, PhD†

*Insight Centre for Data Analytics and †School of Public Health, Physiotherapy and Sports Science, University College Dublin, Ireland

Background: Single-item athlete self-report measures consist of a single question to assess a dimension of wellbeing. These methods are recommended and frequently used for athlete monitoring, yet their uniformity has not been well assessed, and we have a limited understanding of their relationship with measures of training load.

Objective: To investigate the applications and designs of single-item self-report measures used in monitoring team-sport athletes and present the relationship between these measures and measures of training load.

Data Sources: PubMed, Scopus, and SPORTDiscus were searched between inception and March 2019.

Study Selection: Articles were included if they concerned adult athletes from field- or court-sport domains, if athlete wellbeing was measured using a single-item self-report, and if the relationship with a measure of modifiable training load was investigated over at least 7 days.

Data Extraction: Data related to participant characteristics, self-report measures, training load measures, and statistical analysis and outcomes were extracted by 2 authors (C.D. and C.D.).

Data Synthesis: A total of 21 studies were included in the analysis. A narrative synthesis was conducted. The measures

used most frequently were muscle soreness, fatigue, sleep quality, stress, and mood. All measures presented various relationships with metrics of training load from no association to a very large association, and the associations were predominantly trivial to moderate in the studies with the largest numbers of observations. Relationships were largely negative associations.

Conclusions: The implications of this review should be considered by users in the application and clinical utility of single-item self-report measures in athlete monitoring. Great emphasis has been placed on examining the relationship between subjective and objective measures of training load. Although the relationship is still unclear, such an association may not be expected or useful. Researchers should consider the measurement properties of single-item self-report measures and seek to establish their relationship with clinically meaningful outcomes. As such, further study is required to inform practitioners on the appropriate objective application of data from single-item self-report measures.

Key Words: subjective measures, wellness, athlete monitoring

Key Points

- A variety of approaches are used to apply and analyze single-item self-report measures in team-sport athlete monitoring.
- Composite and single-item wellness measures presented various relationships with measures of training load, ranging from no association to a very large association.
- Although self-report measures have established value for users in communication facilitation and information disclosure, further research is required to establish their objective clinical utility.
- In future work, investigators should use evidence-based considerations to develop (ie, measurement properties) and analyze self-report measures in order to encourage robust and uniform research methods that can inform clinical practice.

In the era of data gathering in the athlete-monitoring continuum, researchers and practitioners in sports science and medicine find themselves in pursuit of meaningful signposts to facilitate athlete progress through optimal workload and recovery while reducing injuries and assembling minute components of competitive advantage. The recommendations for such signposts have focused on an array of objective and subjective measures of training load and recovery.¹ Surveying monitoring trends in high-

performance sport, Taylor et al² found that most practitioners (70%) placed equal focus on load quantification and the monitoring of fatigue and recovery; 84% used self-report questionnaires.

Although multi-item self-report measures with published validity and reliability have been considered responsive to training-induced changes in athletes' wellbeing,³ sports programs tend to favor brief, custom self-report measures in practice due to their ease of use, sport specificity, and

Table 1. Inclusion and Exclusion Criteria for Study Eligibility

Inclusion criteria
Study concerned team-sport athletes from field or court sports at any level.
Study measured athlete wellbeing using 1 or multiple single-item self-report variable(s).
Athletes were in training and a parallel measure of load was provided (<i>load</i> was defined as being modifiable by a coach [ie, an external load measure or session rating of perceived exertion, which can be modified by session duration]).
Analysis of the relationship between the self-report measures and training load was conducted; acceptable methods were correlation or regression analysis, including linear mixed models.
Measures were collected over a minimum period of 1 wk (ie, 7 d).
Athletes were aged > 18 y.
Exclusion criteria
Studies were of a nontraining experimental intervention (eg, recovery strategies, nutritional supplementation).
Studies were not reported in English.
Articles included injured participants or participants who did not train.

automation capacity.^{2,4,5} In a review, Saw et al³ identified subscales that may be of use in monitoring immediate responses to training load and concluded that previous recommendations^{6–8} for daily monitoring with self-administered measures remained appropriate owing to the need for immediate, daily adjustments to training. However, the relationship between these measures and athlete workload remains unclear.

Athlete self-report measures (ASRMs) in practice often comprise brief, single-item checklists derived from validated questionnaires, symptoms of overtraining, or sport-specific outcomes that are intended to be completed daily. *Single item* in this context refers to the single-question measurement of an aspect of wellbeing, such as rating general fatigue on a Likert scale, as opposed to a multi-item measure in which several questions may be used to quantify fatigue. Yet selectively combining scales or items from multiple empirical measures negates their established psychometric properties.⁹ In addition, many of the custom measures used in practice were not based on empirically derived, valid, or reliable scientific evidence.⁹ Further challenges with the utility of the data obtained using this method include questionnaire fatigue, data accuracy, and practitioner data burden,^{10,11} particularly concerning the establishment of meaningful change and actionable signposts or “red flags.”^{9,12}

A key challenge for practitioners designing and implementing these measures is the lack of understanding regarding the clinical utility of single-item self-report measures of athlete wellbeing. *Clinical utility* in sport science and medicine may be described as the relevance or usefulness of an intervention or process.¹³ Evidence has suggested that ASRMs in practice are used predominantly as status indicators of athlete readiness and facilitators of communication,^{10,14} but whether and how well these measures respond to training load or reflect recovery or readiness remains unclear. For instance, whether the relationship between athletes’ self-assessed wellbeing using single-item measures and workload is strong or weak, linear or nonlinear, negative or positive, or even exists at all is unknown. This may be due, in part, to the heterogeneity of

Table 2. Literature Search Strategy in PubMed

Title or abstract: (athlete* OR sportspeople OR sportsperson* OR sportsman OR sportsmen OR sportswoman OR sportswomen OR “sports people” OR “sports person” OR “sports man” OR “sports men” OR “sports woman” OR “sports women” OR player* OR baseballer* OR “baseball player*” OR basketball* OR “basketball player*” OR batsman OR batsmen OR “camogie player*” OR cricketer* OR footballer* OR “football player*” OR “Australian football player*” OR “American football player*” OR “Gaelic footballer*” OR “field hockey player*” OR handballer* OR “handball player*” OR hurler* OR “lacrosse player*” OR netballer* OR “netball player*” OR “rugby union player*” OR “rugby league player*” OR “soccer player*” OR softball* OR “softball player*” OR volleyballer* OR “volleyball player*”)

AND

(self-report* OR “self report*” OR diary OR diaries OR questionnaire* OR self-evaluation OR “self evaluation” OR self-appraisal OR “self appraisal” OR self-assessment OR “self-assessment” OR self-rating OR “self rating” OR subjective OR perceive* OR perceptual)

AND

(wellbeing OR well-being OR “well being” OR wellness OR “mental state*” OR “state of mind” OR mood* OR emotion* OR anxiety OR confidence OR self-esteem OR self-efficacy OR motivation OR depression OR stress OR health OR psychological OR tension OR feeling* OR “physical state” OR “physical functioning” OR “perceived recovery” OR “perceived strength” OR soreness OR “quality of life” OR readiness OR vitality OR vigor OR vigour OR sleepiness OR “sleep quality” OR fatigue OR tiredness OR alertness OR distress OR “social function” OR appetite OR overtrain* OR overreach*)

the measures themselves, how they are applied, and the fact that they are frequently developed commercially or designed in-house.

Custom single-item self-report measures are widely used in elite sport² to measure athlete readiness and the training-load response. As such, exploring their relationship with measures of training load is warranted to inform the design and implementation of future self-report measures in sport. Clarity is needed on the specifications of the measures in use and whether a relationship exists between single-item self-report wellness measures and training load. Because the types of training, applications of monitoring, and athletes’ motivations differ between team and individual sports, it is important to consider these studies in isolation.^{5,10} Therefore, the purposes of our review were to (1) investigate the application and design of single-item self-report measures used to monitor team-sport athletes and (2) examine the association between these measures and measures of training load.

METHODS

Search Strategy and Study Selection

The design and reporting of this review were conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).¹⁵ This review was preregistered with PROSPERO (CRD42019138105). Criteria for study eligibility can be found in Table 1. The literature search was conducted using the PubMed, Scopus, and SPORTDiscus databases between inception and March 1, 2019. The search strategy for this review is presented in Table 2 and had no applied restrictions. A total of 18 361 articles were found and

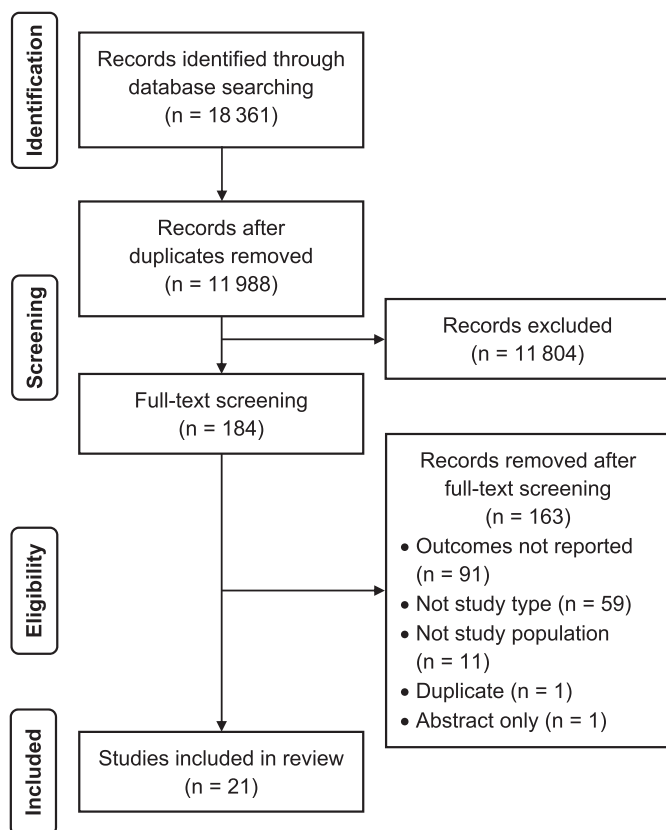


Figure 1. Study selection diagram.

imported into Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia) for screening. Abstracts were screened for eligibility by 2 authors (C.D. and C.D.), whereas full-text screening was completed by the first author (C. Duignan). A total of 21 studies were included in the review (Figure 1).

Data Extraction and Synthesis

Data were extracted by 2 authors (C.D. and C.D.) and entered in a custom-designed spreadsheet. Data related to participant characteristics (Table 3), self-report measures (Table 4), and statistical analysis and outcomes (see Supplemental Table 1, available online at <http://dx.doi.org/10.4085/1062-6050-2020-20.S1>) were recorded. A meta-analysis was deemed inappropriate because of the heterogeneity of both the self-report and training load measures used in the included studies; therefore, a narrative synthesis of the literature was conducted.

Assessment of Study Quality

The methodologic quality of the studies was assessed independently by 2 reviewers (C.D. and C.D.) using a modified Downs and Black checklist.³⁷ The original Downs and Black checklist comprises 27 questions that assess the quality of randomized and nonrandomized studies of health care interventions.³⁷ Modified versions of the checklist are commonly used to establish the quality of observational studies in sports science and medicine.³⁸ We deemed appropriate and retained 13 questions (items 1, 2, 3, 5, 6, 7, 9–12, 16, 18, and 25). Discrepancies in quality scoring were

discussed among 3 authors (C.D., C.D., and C.B.), and a consensus was reached. Full-quality scoring is presented in Supplemental Table 2.

RESULTS

The 21 studies in the review involved 6 sports and 500 participants (soccer = 181, American football = 159, Australian football = 87, rugby sevens = 48, field hockey = 12, volleyball = 13). Assessment durations varied from 8 days to 36 weeks and included undefined season-long monitoring (Table 3). Most researchers described the self-report measure as evaluating *wellness*; thus, we used this term to present the results collectively. Associations between wellness measures and training load measures have been summarized in the text and are presented in Supplemental Table 3.

Study Quality

The study quality assessment is presented in Supplemental Table 2. In 12* of the 21 studies, the athletes' sex(es) was (were) not stated. Where sex was known (n = 176), female athletes were underrepresented in the sample (n = 29, 16%). Principal confounders of the research in this review were factors such as player position, fitness, ambient temperature, and training type. Six investigations^{16,20,26,32,35,36} partially accounted for the distributions of these factors, whereas 7 studies^{16,17,21,26–28,35} accounted for confounding factors in their data analysis. None of the authors reported the required measure of random variability (interquartile range) for the ordinal data of their self-report measures; however, 1 study²⁵ was awarded a score for this question because the self-report measures were deemed not to be a main outcome. The statistical tests used were determined to be appropriate in all cases. A total of 15 groups^{16,20–25,27,30–36} described participants lost to follow-up; this score was not awarded when the reason for missing data was not reported.

Wellness Assessment

The number of questions ranged from 3 to 8, and 5-, 7-, or 10-point Likert scales were used. All single-item variables were related to physical and psychological health except in 1 investigation²³ in which the desire to train was a variable; individual results for this variable were not provided. The most commonly used measures were muscle soreness (n = 20), fatigue (n = 20), sleep quality (n = 19), stress (n = 14), and mood (n = 6). A total of 14 studies provided or used a composite measure, 10 of which^{16–20,24,27,28,30,31} were summed wellness scores and 4 of which^{21–23,32} were averaged scores. Six other variables were assessed (motivation to train, sleep quantity, time of sleep, mental fatigue, energy, and perceived recovery). When the direction of the scale was evident (n = 18), a higher score reflected better wellness in 11 studies,^{16–18,23,24,28,32–36} and a lower score represented better wellness in 7 studies.^{19,20,25,27,29–31} The authors of 12 investigations[†] cited a 1995 review by Hooper and Mackinnon⁶ in describing the design of their self-report

*References 16, 17, 19, 24, 26, 28, 31–36.

†References 16–18, 20–23, 26, 30, 31, 35, 36.

Table 3. Study Characteristics

Study	Sport	Level	One Team?	Sample, N	Sex, n		Duration, wk
					Females	Males	
Buchheit et al ¹⁶ (2013)	Australian football	Professional	Y	18	NS	NS	2
Clemente et al ¹⁷ (2017)	Soccer	Professional	Y	35	NS	NS	Season (unspecified)
Clemente et al ¹⁸ (2019)	Volleyball	Elite	Y	13	0	13	36
Elloumi et al ¹⁹ (2012)	Rugby sevens	Elite	Y	16	NS	NS	8
Fessi et al ²⁰ (2016)	Soccer	Professional	Y	17	0	17	2
Gallo et al ²¹ (2016)	Australian football	Professional	Y	36	0	36	10
Gallo et al ²² (2017)	Australian football	Professional	Y	33	0	33	23
Gathercole et al ²³ (2015)	Rugby sevens	Elite	Y	12	12	0	6
Govus et al ²⁴ (2018)	American football	Elite	Y	58	NS	NS	8
Haller et al ²⁵ (2019)	Soccer	Professional	Y	22	0	22	16 (estimated)
Henderson et al ²⁶ (2019)	Rugby sevens	Professional	Y	20	NS	NS	Season (unspecified)
Ihsan et al ²⁷ (2017)	Field hockey	Elite	Y	12	0	12	1 (9 d)
Malone et al ²⁸ (2018)	Soccer	Elite	N	48	NS	NS	Season (unspecified)
Mara et al ²⁹ (2015)	Soccer	Elite	Y	17	17	0	18
Moalla et al ³⁰ (2016)	Soccer	Professional	Y	14	0	14	16
Rabbani et al ³¹ (2018)	Soccer	Professional or semiprofessional	Y	8	NS	NS	1 (8 d)
Sampson et al ³² (2019)	American football	Collegiate	Y	42	NS	NS	17
Thorpe et al ³³ (2015)	Soccer	Professional	Y	10	NS	NS	2 (17 d)
Thorpe et al ³⁴ (2017)	Soccer	Elite	Y	10	NS	NS	2 (17 d)
Wellman et al ³⁵ (2017)	American football	Collegiate	Y	29	NS	NS	3 (20 d)
Wellman et al ³⁶ (2019)	American football	Collegiate	Y	30	NS	NS	In-season (unspecified)

Abbreviations: N, no; NS, not stated; Y, yes.

Table 4. Self-Report Measure Characteristics

Study	No. of Questions		Method	Muscle Soreness	Fatigue	Sleep Quality	Stress	Mood	Other	Sum or Average
	Scale	Range								
Buchheit et al ¹⁶ (2013)	5	5, with 0.5 increments	Not stated	Y	Y	Y	Y	Y	NA	Sum
Clemente et al ¹⁷ (2017)	4	7	Tablet, custom application	Y	Y	Y	Y	NA	NA	Sum (Hooper Index)
Clemente et al ¹⁸ (2019)	4	7	Tablet, custom application	Y	Y	Y	Y	NA	NA	Sum (Hooper Index)
Elloumi et al ¹⁹ (2012)	8	7	Not stated	Y	NA	NA	Y	NA	Y	Sum ^e
Fessi et al ²⁰ (2016)	4	7	Not stated	Y	Y	Y	Y	NA	Y	Sum (Hooper Index)
Gallo et al ²¹ (2016)	5	7	Personal or squad device	Y	Y	Y	Y	Y	NA	Average ^e
Gallo et al ²² (2017)	5	7	Personal or squad device	Y	Y	Y	Y	Y	NA	Average ^e
Gathercole et al ²³ (2015)	7	7	Not stated	Y	Y	Y	Y	NA	Y	Average ^e
Govus et al ²⁴ (2018)	3	5	Squad device	Y	Y	Y	NA	NA	Y	z score ^e
Haller et al ²⁵ (2019)	4	10 cm	Not stated	NA	Y	Y	NA	NA	Y	NA
Henderson et al ²⁶ (2019)	5	7	Personal or squad device (Smartabase ^a)	Y	Y	Y	Y	NA	Y	NA
Ihsan et al ²⁷ (2017)	4	10	Personal device	Y	Y	Y	NA	Y	NA	Sum
Malone et al ²⁸ (2018)	5	7	Not stated	Y	Y	Y	Y	NA	Y	Sum
Mara et al ²⁹ (2015)	2	10	Cloud-based spreadsheet (Google Drive ^b)	Y	Y	NA	NA	NA	Y	NA
Moalla et al ³⁰ (2016)	4	7	Not stated	Y	Y	Y	Y	NA	NA	Sum (Hooper score)
Rabbani et al ³¹ (2018)	4	7	WhatsApp ^c	Y	Y	Y	Y	NA	NA	Sum (Hooper Index)
Sampson et al ³² (2019)	3	5	Not stated	Y	Y	Y	NA	NA	NA	Average
Thorpe et al ³³ (2015)	3	7	Not stated	Y	Y	Y	NA	NA	NA	NA
Thorpe et al ³⁴ (2017)	3	7	Not stated	Y	Y	Y	NA	NA	NA	NA
Wellman et al ³⁵ (2017)	6	5	Pen and paper	Y	Y	Y	Y	Y	Y	NA
Wellman et al ³⁶ (2019)	6	5	Smartphone application (CoachMePlus ^d)	Y	Y	Y	Y	Y	Y	NA

Abbreviations: NA, not applicable; Y, yes.

^a Fusion Sport, Boulder, CO.

^b Google LLC, Mountain View, CA.

^c WhatsApp Inc, Menlo Park, CA.

^d Buffalo, NY.

^e Denotes where results for individual variables were not given.

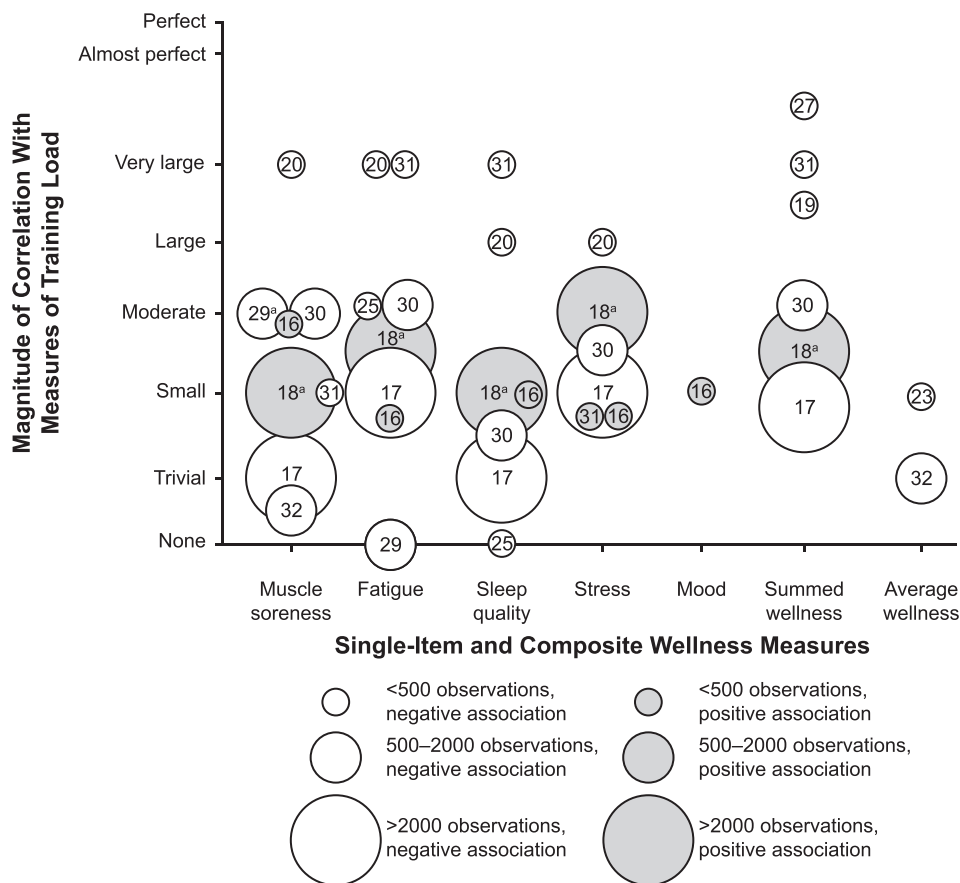


Figure 2. Results of studies that used Pearson and Spearman correlation analyses. Reference numbers for the studies are shown in the circles. For illustrative purposes only, the magnitudes are as stated in the original studies (see Supplemental Table 3 for descriptors). ^a Indicates a correlation was present only in some parts of the season.

measure: 4 sets of authors^{17,18,20,31} labeled the summed wellness score the *Hooper Index*, and 1 set of authors³⁰ labeled it the *Hooper Score* (Table 4).

Thirteen studies explored (1) the relationship between pretraining wellness and subsequent load output (ie, same day)^{16,18,21,24,26–28} and (2) the relationship between training load and subsequent wellness (ie, next day)^{18,22,32–36}, whereas in 8 studies,^{17,19,20,23,25,29–31} the combination of data used for the correlation analysis (ie, whether they used same-day versus next-day data) was not clear. Researchers used a variety of derived variables, true values, changes in values and z scores, means and group means, normalization, and fixed and random effects in their analyses. Results were presented as a mix of significance testing, effect sizes and correlation values, and interpretations of magnitude and subsequent inferences that were mostly based on the work of Hopkins et al,^{39,40} Batterham and Hopkins,⁴¹ and Hopkins⁴² (see Supplemental Table 3). Figure 2 illustrates the magnitude of the correlations as interpreted in the original studies (Pearson and Spearman analyses only) between single-item and composite wellness scores and training load on the basis of the number of observations used in the original studies. If the number of observations was not stated in the original study, the maximum number of potential observations was estimated by multiplying the sample size by either the number of sessions or duration of the study. Investigations are grouped by analysis type for

the results summaries, whereas Supplemental Table 3 provides further details.

Composite Wellness Scores: Summed and Averaged

Associations between summed wellness and training-load measures were reported in 8 studies (see Supplemental Table 3). Pearson and Spearman correlations revealed small to moderate,^{17,18,30} large,^{19,20} and very large²⁷ positive and negative associations between summed wellness scores and measures of training load. Two groups used a summed wellness score analyzed with linear mixed models. Govus et al²⁴ found that a 1-unit increase in wellness z score was associated with a trivial increase in player load, and Malone et al²⁸ found trivial to very likely negative effects of the wellness z score on training output measures.

An average wellness score was computed in 4 studies. Gallo et al^{21,22} calculated linear mixed models and demonstrated no effect of days postmatch or match load on the weekly wellness profile²² and predominantly trivial effects of the wellness z score on training output variables.²¹ Using the Spearman correlation, Sampson et al³² identified a trivial negative association between the wellness z score and the acute:chronic workload ratio on the previous day. Using the Pearson correlation, Gathercole et al²³ reported a small negative correlation between wellness and training impulse for a 7-day average.

Individual Wellness Variables

Muscle Soreness. Associations between muscle soreness and training load measures were reported in 14 investigations. Pearson and Spearman correlations varied from trivial^{17,32} to small,^{18,31} moderate,^{16,29,30} and large²⁰ negative and positive associations between measures of muscle soreness and measures of training load.

Applying linear mixed models, Thorpe et al^{33,34} found no correlation between training load and muscle soreness. Wellman et al^{35,36} used categorical scores to determine that players who rated muscle soreness higher had greater training loads than those who scored lower. Govus et al²⁴ reported that muscle soreness was not related to player load or session rating of perceived exertion (sRPE); when the variables were modelled individually, a 1-unit increase in muscle soreness corresponded with a trivial decrease in sRPE. Conversely, Henderson et al²⁶ described increased muscle soreness as associated with a trivial increase in the physical performance factor value.

Fatigue. Associations between measures of fatigue and training load were reported in 13 studies. Pearson and Spearman correlations yielded results that varied from no correlation²⁹ to small,^{16–19} moderate,^{25,30} large,²⁰ and very large³¹ negative and positive associations between measures of fatigue and measures of training load.

Using linear mixed models, researchers found a large negative³³ association between total high-intensity running and fatigue and small to moderate associations between fatigue and cumulative days of high-speed running distance.³⁴ Based on categorical values, Wellman et al^{35,36} determined that players who rated fatigue higher had greater training loads than those who scored lower, whereas fatigue was not included in the best-fit model of Henderson et al.²⁶

Sleep Quality. The association between sleep quality and training load was provided by 13 groups. Pearson and Spearman correlations indicated no associations²⁵ to trivial,¹⁷ small,^{16,18,30} large,²⁰ and very large³¹ negative and positive associations between measures of sleep quality and measures of training load.

In 3 studies using linear mixed models, researchers found no associations^{24,33,34} between measures of training load and sleep quality. Wellman et al³⁵ demonstrated no differences in training load related to sleep quality, but Wellman et al³⁶ noted differences between maximal-intensity deceleration distance and sprint distance for players who rated sleep quality as different on specific days of the week. Sleep quality was not included in the best-fit model of Henderson et al.²⁶

Stress. The association between stress and training load was reported in 9 studies. Pearson and Spearman correlations displayed results that ranged from no correlation¹⁸ to small,^{16–18} moderate,³⁰ and large²⁰ positive and negative associations between measures of stress and training load.

Authors²⁶ using linear mixed models reported that increased stress was associated with a trivial increase in the physical performance factor value, whereas Wellman et al^{35,36} used categorical values to show that those who perceived less stress had lower training loads than those who perceived more stress.

Mood. Three sets of investigators depicted the association between mood and training load. Buchheit et al¹⁶ used the Pearson correlation and observed a small positive

association. Wellman et al³⁵ used linear mixed models and found that those with more favorable mood scores had lower training loads than those with less favorable responses. However, Wellman et al³⁶ noted no differences in movement variables for measures of mood.

Other. Associations between training load and other individual variables were reported in 5 studies. Using Spearman correlations, researchers detected no association between training load and sleep quantity across the season²⁹ and a trivial negative association between energy and the acute:chronic workload ratio.³² In addition, no correlation was present between training data and the measures of mental fatigue or sleep duration.²⁵ Govus et al²⁴ used linear mixed models and reported that pretraining energy was trivially positively related to player load and not related to sRPE. Henderson et al²⁶ indicated that a higher level of perceived recovery was associated with a trivial decrease in the physical performance factor value.

DISCUSSION

Our aims in this review were to identify the applications and designs of single-item self-report measures used to monitor team-sport athletes and present their relationship with measures of training load. Predominant findings were a paucity of evidence-based practice in the adoption and design of single-item self-report measures beyond general recommendations, a variety of data-collection and -analysis techniques, and a spectrum of no correlation to very large correlations with measures of training load. We provided an in-depth critical analysis of the use of single-item self-report measures in team sports and highlighted the lack of quality and rigor in how these measures were approached and used.

Readers should be vigilant in interpreting the results regarding the direction of the scale (ie, higher or lower scores reflecting better wellness) with respect to positive or negative relationships with workload, the direction of the proposed effect, the type of load measure (ie, internal sRPE or external load measures), and the analytic approaches used. Heterogeneity of the approaches made the results challenging to synthesize with clarity (Figure 3).

Self-Report Measure Overview

A concern regarding the use of single-item self-report variables was the lenient interpretation and citation of previous recommendations. This was frequently related to the work of Hooper and Mackinnon⁶ and Hooper et al.⁴³ Whereas the authors of 12 studies[‡] referenced a 1995 review by Hooper and Mackinnon⁶ for their chosen self-report measures, most of these researchers used a design closer to the questionnaire used by Hooper et al⁴³ in an original paper from the same year rather than the recommendations in the cited review (ie, the original study⁴³ involved a 7-point scale of 4 items, whereas in the review, the authors⁶ recommended a 5-point scale of 7 items). This included, in some cases, the construction of a Hooper Index or Hooper Score (also referred to as the *Hooper Scale* in other studies); however, in the referenced studies, the authors suggested neither the terms nor the summing method. This “naming” alludes to the idea of a

[‡]References 16–18, 20–23, 26, 30, 31, 35, 36.

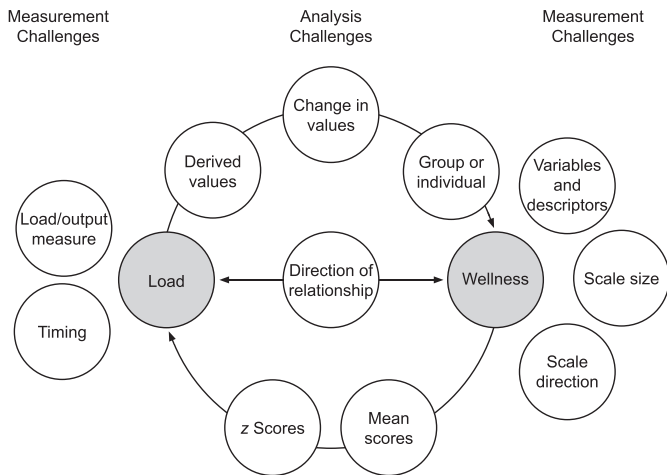


Figure 3. Challenges in measurement and analysis.

validated, formative construct and was subsequently cited by others.

Relationship Between Measures of Wellness and Measures of Training Load

The existence, size, and direction of a relationship between single-item self-report measures and measures of training load varied for both composite and individual measures, creating a challenge in drawing inferences or conclusions from the results. The existence of a relationship appeared more prominent in studies that used correlation measures than in those that used linear mixed models, potentially because the mixed models accounted for the correlation within repeated measures for each athlete. This was most notable in Ihsan et al,²⁷ who normalized measures to time and RPE, but this investigation had one of the smallest numbers of observations (maximum of 72) of the included works.

Whereas Saw et al³ found the ASRM to be responsive to acute and chronic training loads, they used a bespoke method of measuring sensitivity and consistency that was not replicable. In addition, the self-report measures used in the review of Saw et al³ were typically retrospective; they often referred to a time period in the previous week (eg, profile of mood states) or the previous 3 days (eg, recovery stress questionnaire for athletes), in contrast to the predominantly daily measured single-items evaluated here. They identified the vigor/motivation, physical symptoms/injury, nontraining stress, fatigue, physical recovery, general health/wellbeing, and “being in shape” subscales as useful in their review; only 1 subscale (ie, fatigue) directly overlaps with the measures in our review.

Review Context

One of the most challenging aspects of researching the use of self-report measures of athlete wellbeing and, by extension, conducting a narrative synthesis of this literature, was the tendency for observations to be made in an uncontrolled “real-world” environment. Authors of 4 studies^{17,18,28,30} explicitly stated that the measures used were part of the normal team monitoring routine, whereas authors of 17 studies^{16,19–27,29,30–36} did not specify whether the methods existed previously or had been implemented

for research purposes. Heterogeneity among the included studies was further demonstrated by the variety of inclusion and exclusion strategies, including the requirement or “happening” of full participation ($n = 6$),^{16,20,30,31,33,34} minimal training or competition participation requirements ($n = 5$),^{17,18,22,25,36} the exclusion of goalkeepers ($n = 2$),^{20,29} and the requirement for ordinal self-report data to be normally distributed ($n = 2$).^{21,28} Investigators in a further 7 studies^{19,23,24,26,27,32,35} did not reference any requirements for the inclusion or exclusion of participants or data. In only 1 study did the authors²⁵ comment that interim results from data collection were not provided to the coaching staff. It was unclear in all other cases whether the self-report data were or may have been used by practitioners to manipulate the training design or participation throughout the study, thereby creating interdependency between consecutive observations. Caution must be applied in making recommendations on the basis of such works.⁴⁴ In light of these factors, we echo the recommendation of Fullagar et al,⁴⁵ who suggested that studies such as these that lack experimental control should be presented as case reports and authors should resist making inferences beyond general observations.

Whereas research requires controlled experiments to establish confidence in a result, one can also argue that controlled, laboratory-based tests do not transfer to the real-world environment. If the isolated relationship between workload and ASRM data is clinically relevant, then we require controlled data collection to identify whether a relationship exists and, if so, its underlying nature. However, in a real-world scenario, these measures are unlikely to be influenced by training load in isolation, given that they will also be determined by factors such as recovery and psychological and social influences. This indicates that these measures may be more reflective of complex “readiness” than a linear training “response.” Yet defining an optimal state of readiness continues to be elusive; authors of the studies included in this review who explicitly evaluated the effect of pretraining wellness on subsequent training output found predominantly trivial results.

Considerations for Practice

The existence and nature of any relationship between training load and wellbeing create an interesting debate for practitioners seeking to use these outcomes to inform their clinical decision making. For example, the clinical importance of a specific summed wellness z score of -1 corresponding to a difference of 4-m sprint distance in training²⁸ is unclear. Whether these small correlations and magnitude-based inferences of heterogeneous self-report data can adequately inform the design of future measures and whether they justify the appropriateness of a single-item ASRM in its currently recommended form must be considered. Readers should also be aware of the variety of scoring, summing, and averaging techniques used to analyze the self-report measures in this review when they consider the adoption or are planning the design of single-item measures in practice. Saw et al⁹ outlined recommendations for developing custom ASRMs to ensure acceptable psychometric properties; however, readers may also take into account design requirements relative to their needs. For

instance, if the primary role of the ASRM is facilitating communication and prompting information disclosure, a simple approach may suffice.^{10,14}

With regard to current practice, investigators have suggested that data from self-report measures are indeed used predominantly as status indicators and facilitators of communication rather than as decision-making tools¹⁴ and that their value is predicated on athletes' honesty and practitioners' interpretations.¹⁰ Although many practitioners seek clarity and simplicity in athlete monitoring, the complexity of the relationship between wellness and training load and the difficulties associated with their measurement necessitate substantial investment of human resources to enable the best use of an ASRM. Indeed, Hooper and Mackinnon⁶ proposed that the use of these self-analysis tools to identify trends toward overtraining syndrome depended largely on interpretation by a coach. Furthermore, many guidelines that inform the design and implementation of single-item ASRMs have been based on studies that used these measures for monitoring overtraining and, as such, narrow Likert scales may not be sensitive to smaller day-to-day variations in the athlete's state. Further research is required to apply clinically meaningful and individual "traffic-light" approaches with ordinal scales, such as those used in ASRMs.¹² These factors are also important to recognize when considering the potential value of introducing a self-report system in a sports program, given the challenges associated with implementation, adherence, dishonesty, resource investment, and system factors.^{10,46}

Review Limitations and Future Work

Whereas our review provides insight into this complex area of sports medicine research, a number of limitations should be acknowledged. First, the possibility that studies identifying correlations between self-report measures and workload were more likely to be published than those not identifying such correlations was high but also difficult to detect. A funnel plot and its associated statistics are typically used in intervention research to elucidate such bias; nonetheless, this was not possible due to the heterogeneity of the studies (and their associated primary and secondary outcomes). Publication bias, if it exists, may result in overestimation of the strength of the relationship between the measures evaluated. Next, study heterogeneity, including the variations in the self-report and training load measures used, the different frequencies and durations of administration, and the alternative statistical analysis approaches made synthesizing this literature particularly challenging and drawing definitive conclusions impossible. These limitations highlight the need for further research and more critical consideration in the design, use, and analysis of self-report measures in research and practice beyond interpreting general recommendations for athlete self-appraisal.

In addition, we were unable to identify long-term relationships between single-item ASRMs and training load measures because the included studies focused primarily on daily analyses. In future evaluations, investigators may use single-item self-report data to represent different stages of the season or binary measures, such as changes in the days posttraining or competition. Researchers may also investi-

gate the ability of single-item self-report measures to reflect athlete readiness or recovery, although defining such a state of readiness is a subject for debate.

CONCLUSIONS

Team-sport programs are interpreting general recommendations in their use of single-item self-report measures, which predominantly feature muscle soreness, fatigue, sleep quality, stress, mood, and often a composite score. Studies with the largest numbers of observations showed predominantly trivial to moderate associations between single-item self-report measures and measures of workload. Where associations were found, the directions of the relationship were predominantly negative (ie, when training load increased, wellness decreased). Although the nature of the relationship is still unclear, such an association may not be expected or useful. Self-report measures have established value for users in sport programs through certain aspects, such as communication,^{10,14} yet further assessment is required to establish their clinical utility beyond the role of complementary tool. Future authors should consider the measurement properties of single-item self-report measures and establish their relationship with clinically meaningful outcomes. The potential may exist for such measures to be incorporated in predictive risk tools in conjunction with established risk factors; however, further research is required to inform practitioners on the appropriate objective application of data from single-item self-report measures in team sport.

ACKNOWLEDGMENTS

This work was supported by funding from Science Foundation Ireland (Drs Duignan and Caulfield) under the grant for the Insight Centre for Data Analytics (SFI/12/RC/2289_P2) and the Health Research Board (Dr Doherty) under grant No. ARPP-A-2018-002.

REFERENCES

1. Soligard T, Schweltnus 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(17):1030–1041. doi:10.1136/bjsports-2016-096581
2. Taylor K-L, Chapman D, Cronin J, Newton M, Gill N. Fatigue monitoring in high performance sport: a survey of current trends. *J Aust Strength Cond.* 2012;20(1):12–23.
3. 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.* 2016;50(5):281–291. doi:10.1136/bjsports-2015-094758
4. Saw AE, Main LC, Gastin PB. Monitoring athletes through self-report: factors influencing implementation. *J Sports Sci Med.* 2015;14(1):137–146. doi:10.1519/JSC.0000000000000499
5. 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(9):2518–2526. doi:10.1519/JSC.0b013e31827fd600
6. Hooper SL, Mackinnon LT. Monitoring overtraining in athletes: recommendations. *Sports Med.* 1995;20(5):321–327. doi:10.2165/00007256-199520050-00003
7. Fry RW, Morton AR, Keast D. Overtraining in athletes: an update. *Sports Med.* 1991;12(1):32–65. doi:10.2165/00007256-199112010-00004

8. Kenttä G, Hassmén P. Overtraining and recovery: a conceptual model. *Sports Med.* 1998;26(1):1–16. doi:10.2165/00007256-199826010-00001
9. Saw AE, Kellmann M, Main LC, Gastin PB. Athlete self-report measures in research and practice: considerations for the discerning reader and fastidious practitioner. *Int J Sports Physiol Perform.* 2017;12(suppl 2):S127–S2135. doi:10.1123/ijssp.2016-0395
10. Duignan CM, Slevin PJ, Caulfield BM, Blake C. Exploring the use of mobile athlete self-report measures in elite Gaelic games: a qualitative approach. *J Strength Cond Res.* Published online August 8, 2019. doi:10.1519/JSC.0000000000003334
11. Halson SL. Monitoring training load to understand fatigue in athletes. *Sports Med.* 2014;44(suppl 2):139–147. doi:10.1007/s40279-014-0253-z
12. Robertson S, Bartlett JD, Gastin PB. Red, amber, or green? Athlete monitoring in team sport: the need for decision-support systems. *Int J Sports Physiol Perform.* 2017;12(suppl 2):73–79. doi:10.1123/ijssp.2016-0541
13. Lesko LJ, Zineh I, Huang SM. What is clinical utility and why should we care? *Clin Pharmacol Ther.* 2010;88(6):729–733. doi:10.1038/clpt.2010.229
14. Saw AE, Main LC, Gastin PB. Role of a self-report measure in athlete preparation. *J Strength Cond Res.* 2015;29(3):685–691. doi:10.1519/JSC.0000000000000698
15. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097
16. Buchheit M, Racinais S, Bilsborough JC, et al. Monitoring fitness, fatigue and running performance during a pre-season training camp in elite football players. *J Sci Med Sport.* 2013;16(6):550–555. doi:10.1016/j.jsams.2012.12.003
17. Clemente FM, Mendes B, Nikolaidis PT, Calvete F, Carriço S, Owen AL. Internal training load and its longitudinal relationship with seasonal player wellness in elite professional soccer. *Physiol Behav.* 2017;179:262–267. doi:10.1016/j.physbeh.2017.06.021
18. Clemente FM, Mendes B, Palao JM, et al. Seasonal player wellness and its longitudinal association with internal training load: study in elite volleyball. *J Sports Med Phys Fitness.* 2019;59(3):345–351. doi:10.23736/S0022-4707.18.08312-3
19. Elloumi M, Makni E, Moalla W, et al. Monitoring training load and fatigue in rugby sevens players. *Asian J Sports Med.* 2012;3(3):175–184.
20. Fessi MS, Nouira S, Dellal A, Owen A, Elloumi M, Moalla W. Changes of the psychophysical state and feeling of wellness of professional soccer players during pre-season and in-season periods. *Res Sport Med.* 2016;24(4):375–386. doi:10.1080/15438627.2016.1222278
21. Gallo TF, Cormack SJ, Gabbett TJ, Lorenzen CH. Pre-training perceived wellness impacts training output in Australian football players. *J Sports Sci.* 2016;34(15):1445–1451. doi:10.1080/02640414.2015.1119295
22. Gallo TF, Cormack SJ, Gabbett TJ, Lorenzen CH. Self-reported wellness profiles of professional Australian football players during the competition phase of the season. *J Strength Cond Res.* 2017;31(2):495–502. doi:10.1519/JSC.0000000000001515
23. Gathercole R, Sporer B, Stellingwerff T. Countermovement jump performance with increased training loads in elite female rugby athletes. *Int J Sports Med.* 2015;36(9):722–728. doi:10.1055/s-0035-1547262
24. Govus AD, Coutts A, Duffield R, Murray A, Fullagar H. Relationship between pretraining subjective wellness measures, player load, and rating-of-perceived-exertion training load in American college football. *Int J Sports Physiol Perform.* 2018;13(1):95–101. doi:10.1123/ijssp.2016-0714
25. Haller N, Ehlert T, Schmidt S, et al. Circulating, cell-free DNA for monitoring player load in professional football. *Int J Sports Physiol Perform.* 2019;14(6):718–726. doi:10.1123/ijssp.2018-0756
26. Henderson MJ, Fransen J, McGrath JJ, Harries SK, Poulos N, Coutts AJ. Individual factors affecting rugby sevens match performance. *Int J Sports Physiol Perform.* 2019;14(5):620–626. doi:10.1123/ijssp.2018-0133
27. Ihsan M, Tan F, Sahrom S, Choo HC, Chia M, Aziz AR. Pre-game perceived wellness highly associates with match running performances during an international field hockey tournament. *Eur J Sport Sci.* 2017;17(5):593–602. doi:10.1080/17461391.2017.1301559
28. Malone S, Owen A, Newton M, et al. Wellbeing perception and the impact on external training output among elite soccer players. *J Sci Med Sport.* 2018;21(1):29–34. doi:10.1016/j.jsams.2017.03.019
29. Mara JK, Thompson KG, Pumpa KL, Ball NB. Periodization and physical performance in elite female soccer players. *Int J Sports Physiol Perform.* 2015;10(5):664–669. doi:10.1123/ijssp.2014-0345
30. Moalla W, Fessi MS, Farhat F, Nouira S, Wong DP, Dupont G. Relationship between daily training load and psychometric status of professional soccer players. *Res Sport Med.* 2016;24(4):387–394. doi:10.1080/15438627.2016.1239579
31. Rabbani A, Baseri MK, Reisi J, Clemente FM, Kargarfard M. Monitoring collegiate soccer players during a congested match schedule: heart rate variability versus subjective wellness measures. *Physiol Behav.* 2018;194:527–531. doi:10.1016/j.physbeh.2018.07.001
32. Sampson JA, Murray A, Williams S, Sullivan A, Fullagar HKF. Subjective wellness, acute: chronic workloads, and injury risk in college football. *J Strength Cond Res.* 2019;33(12):3367–3373. doi:10.1519/JSC.0000000000003000
33. Thorpe RT, Strudwick AJ, Buchheit M, Atkinson G, Drust B, Gregson W. Monitoring fatigue during the in-season competitive phase in elite soccer players. *Int J Sports Physiol Perform.* 2015;10(8):958–964. doi:10.1123/ijssp.2015-0004
34. Thorpe RT, Strudwick AJ, Buchheit M, Atkinson G, Drust B, Gregson W. The influence of changes in acute training load on daily sensitivity of morning-measured fatigue variables in elite soccer players. *Int J Sports Physiol Perform.* 2017;12(suppl 2):S2107–S2113. doi:10.1123/ijssp.2016-0433
35. Wellman AD, Coad SC, Flynn PJ, Climstein M, McLellan CP. Movement demands and perceived wellness associated with pre-season training camp in NCAA Division I college football players. *J Strength Cond Res.* 2017;31(10):2704–2718. doi:10.1519/JSC.0000000000002106
36. Wellman AD, Coad SC, Flynn PJ, Siam TK, McLellan CP. Perceived wellness associated with practice and competition in National Collegiate Athletic Association Division I football players. *J Strength Cond Res.* 2019;33(1):112–124. doi:10.1519/JSC.0000000000002169
37. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377–384. doi:10.1136/jech.52.6.377
38. Johnston W, O'Reilly M, Argent R, Caulfield B. Reliability, validity and utility of inertial sensor systems for postural control assessment in sport science and medicine applications: a systematic review. *Sports Med.* 2019;49(5):783–818. doi:10.1007/s40279-019-01095-9
39. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc.* 2009;41(1):3–13. doi:10.1249/MSS.0b013e31818cb278
40. Hopkins KD, Hopkins BR, Glass GV. *Basic Statistics for the Behavioral Sciences.* 3rd ed. Boston, MA: Allyn & Bacon; 1996.

41. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. *Int J Sports Physiol Perform.* 2006;1(1):50–57.
42. Hopkins WG. Measures of reliability in sports medicine and science. *Sport Med.* 2000;30(1):1–15.
43. Hooper SL, Mackinnon LT, Howard A, Gordon RD, Bachmann AW. Markers for monitoring overtraining and recovery. *Med Sci Sports Exerc.* 1995;27(1):106–112.
44. Horton R. Offline: the gravy train of systematic reviews. *Lancet.* 2019;394(10211):1790. doi:10.1016/S0140-6736(19)32766-7
45. Fullagar HHK, McCall A, Impellizzeri FM, Favero T, Coutts AJ. The translation of sport science research to the field: a current opinion and overview on the perceptions of practitioners, researchers and coaches. *Sports Med.* 2019;49(12):1817–1824. doi:10.1007/s40279-019-01139-0
46. Duignan CM, Slevin PJ, Caulfield BM, Blake C. Mobile athlete self-report measures and the complexities of implementation. *J Sports Sci Med.* 2019;18(3):405–412.

Address correspondence to Ciara Duignan, PhD, Insight Centre for Data Analytics, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland. Address e-mail to ciara.duignan@insight-centre.org.