The Role of Veracity on the Load Monitoring of Professional Soccer Players: A Systematic Review in the Face of the Big Data Era

João Gustavo Claudino 1,2,*, Carlos Alberto Cardoso Filho 1, Daniel Boullosa 3,4,5,*, Adriano Lima-Alves 6,*, Gustavo Rejano Carrion 1, Rodrigo Luiz da Silva Gianoni 2,*, Rodrigo dos Santos Guimarães 7,8,*, Fúlvio Martins Ventura 9, André Luiz Costa Araújo 10, Sebastián Del Rosso 11, José Afonso 12, and Julio Cerca Serrão 1

1 Laboratory of Biomechanics, School of Physical Education and Sport, Campus São Paulo, Universidade de São Paulo, São Paulo 05508-030, State of São Paulo, Brazil; carlos.filho@alumni.usp.br (C.A.C.F.); gustavo.carrion@usp.br (G.R.C.); jcserrao@usp.br (J.C.S.)
2 Research and Development Department, LOAD CONTROL, Contagem 32280-440, Minas Gerais, Brazil; rodrigogianonii@yahoo.com.br
3 College of Healthcare Sciences, Campus Townsville, James Cook University, Townsville 4811, Australia; daniel.boullosa@gmail.com
4 INISA, Graduate Program of Movement Sciences, Campus Campo Grande, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Mato Grosso do Sul, Brazil
5 Research and Development Department, iLOAD Solutions, Campo Grande 79118-280, Mato Grosso do Sul, Brazil
6 Department of Physiology, Fortaleza Esporte Clube, Fortaleza 60510-290, Ceará, Brazil; adrianolimaat@gmail.com
7 Department of Physiology, América Futebol Clube, Contagem 32180-060, Minas Gerais, Brazil; rodrigo.sportcoach@gmail.com
8 Program de Doctorado en Ciencias del Deporte, Facultad de Ciencias del Deporte, Universidad de Extremadura, 10071 Cáceres, Spain
9 Department of Performance, Guarani Futebol Clube, Campinas 13100-200, State of São Paulo, Brazil; fulvio.ventura@gmail.com
10 Health and Performance Department, Goiás Esporte Clube, Goiânia 74343-010, Goiás, Brazil; andrearaaujo118@hotmail.com
11 Centro de Investigaciones en Nutrición Humana, Escuela de Nutrición, Facultad de Ciencias Médicas, Campus Córdoba, Universidad Nacional de Córdoba, Córdoba 5000, Argentina; delrossosebastian@gmail.com
12 Centre for Research, Education, Innovation and Intervention in Sport, Faculty of Sport, Campus Porto, University of Porto, 4099-002 Porto, Portugal; jneves@fade.up.pt
* Correspondence: claudinojgo@usp.br

Abstract: Big Data has real value when the veracity of the collected data has been previously identified. However, data veracity for load monitoring in professional soccer players has not been analyzed yet. This systematic review aims to evaluate the current evidence from the scientific literature related to data veracity for load monitoring in professional soccer. Systematic searches through the PubMed, Scopus, and Web of Science databases were conducted for reports on the data veracity of diverse load monitoring tools and the associated parameters used in professional soccer. Ninety-four studies were included in the review, with 39 different tools used and 578 associated parameters identified. The pooled sample consisted of 2066 footballers (95% male: 24 ± 3 years and 5% female: 24 ± 1 years). Seventy-three percent of these studies did not report veracity metrics for any of the parameters from these tools. Thus, data veracity was found for 54% of tools and 23% of parameters. The current information will assist in the selection of the most appropriate tools and parameters to be used for load monitoring with traditional and Big Data approaches while identifying those still requiring the analysis of their veracity metrics or their improvement to acceptable veracity levels.

Keywords: data analytics; fitness; illness; injury; performance; recovery
1. Introduction

Load monitoring is used to ensure the best workload that optimizes physical fitness and sports performance while preventing athletes from injury or illness in soccer [1]. In this context, a true individualization of all training factors in team sports is mandatory for a better fitness–fatigue equilibrium through an individualized monitoring approach [2]. Meanwhile, the number of injuries [3] and the physical performance of soccer players continue to increase over the years [4]. For instance, in a 13-year longitudinal analysis, it was found that hamstring injuries have annually increased by 4% in professional soccer players [3]. Factors associated with load monitoring, such as player load, match frequency, playing style, team management, and the continuity of technical staff, have also influenced these trends [3]. On the other hand, it is well accepted that physical performance in soccer matches has continuously improved over the past years. For instance, total distance covered, high-intensity running and sprinting distances, and number of sprints increased by 2% (effect size (ES): 0.22), 30% (ES: 0.82), 35% (ES: 0.93), and 85% (ES: 1.46), respectively, across 7 seasons of the English Premier League (i.e., 2006–2007 compared to 2012–2013) [4]. However, the VO_{2max} of elite female [5] and male [6] soccer players has not changed between 1989 and 2007 for females and between 1989 and 2012 for males. This paradox of a similar aerobic fitness between decades, despite objectively increased match demands, highlights the need to select appropriate monitoring tools. Furthermore, considering that soccer is a complex sport in which players need to develop several physical capacities (e.g., acceleration, agility, endurance), the selection of the best tools to monitor the evolution of players’ physical fitness over the season is required for better managing the complex balance between training, competition, recovery, and evaluation [7–10].

We are living in a time where technology has experienced a fast evolution in the sports arena [11]. This technological growth has enabled the 24 h monitoring of athletes on an individual basis [12], thus allowing large data sets, i.e., Big Data, to be gathered in sport settings. Careful analyses of these large datasets can enhance our knowledge in sport science and medicine, thus supporting the making-decision process for designing appropriate training and competitive strategies [13,14]. However, some premises, which are known as the four Vs of Big Data, should be attained from these datasets for these purposes: (1) volume, (2) variety, (3) velocity, and (4) veracity [15]. Volume refers to the size of the datasets; variety refers to different data formats and data sources; velocity describes the speed at which data is generated and processed for analyses; and veracity refers to the accuracy, quality, relevance, uncertainty, reliability, and predictive values of the collected data [14–17]. Focusing on veracity to identify the best monitoring tools and parameters in soccer is mandatory to identify the “errors” [18]. Furthermore, veracity is a preliminary step to effectively meet all these premises and, thus, optimize the applicability of Big Data to soccer with the best cost-to-benefit ratio. For instance, according to IBM [15], poor data quality generates an economic cost of around USD 3.1 trillion per year in the USA. If this value were the gross domestic product (GDP) of any country, it would be placed in the top 6 GDPs in the world.

Therefore, this systematic review aims to evaluate data veracity for load monitoring in professional soccer. This information is important for a deep understanding of the tools and parameters used in load monitoring as well as the effective implementation of Big Data analyses in professional clubs.

2. Materials and Methods

We adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [19].

2.1. Sources and Study Selection Process

Three electronic databases (PubMed, Scopus, and Web of Science) were systematically searched from inception to August 2020. The command line (“monitoring” OR “monitor”) AND (“training load” OR “load”) AND (“soccer” OR “football”) was used in the
title, abstract, and keywords during the electronic searches. The selection process was completed by two authors (J.G.C. and C.A.C.F.). If some doubt was found during this step, a third author (D.B.) assisted in the decision. The quality appraisal was completed by one author (J.G.C.).

2.2. Eligibility Criteria

Initially, the articles reviewed were identified after reading the titles and abstracts based on the following inclusion criteria:

1. The study was written in English;
2. The study was published as original research in a peer-reviewed journal, and a full-text article was available;
3. Data were reported for soccer players;
4. The participants were professional soccer players;
5. Load monitoring parameters were included.

Then, after this first screening, the eligibility criteria according to PECO were applied in the remaining full manuscripts. (P)articipants: healthy, professional soccer players of any age and sex. (E)xposure: exposure to load monitoring in the training session and/or on match day. (C)omparators: control groups without load monitoring are accepted but not mandatory. (O)utcomes: load monitoring tools and parameters.

2.3. Data Extraction

Five authors (A.L.M., G.R.C., R.G., F.M.V., and S.D.R.) extracted the following information from the included full-text articles: authors, year, sample information (mean, age, sex, sample size, competitive level), study design, veracity data, and information on load monitoring tools and parameters. The parameters were classified according to the applied method (e.g., laboratory or portable methods; venous or capillary blood samples). Discrepancies were resolved through discussion until consensus was reached. A narrative synthesis of data was performed.

The median was used when a range of a veracity metric was reported in a manuscript. All veracity metrics of the reported parameters were used to determine the median and range (minimum–maximum) for each tool.

2.4. Quality Assessment

The quality of all studies was evaluated by one author (J.G.C.) using objective evaluation criteria (see supplementary material, Table S1) based on a previous study by Saw et al. [20]. Scores were assigned based on how well each criterion was met, assuming a maximum possible score of 8 (low risk of bias). Studies with a score of ≤4 were considered poor and were, therefore, subsequently excluded.

2.5. Veracity Analysis

Veracity analysis was performed based on the use of measures related to accuracy, uncertainty, reliability, and quality of data [18,21–23]. All these measures were only obtained when explicitly reported by the authors’ own data. For the used instruments, accuracy was checked in the manufacturer’s official documentation or website. Accuracy is understood as the closeness of the agreement between the result of a measurement and the true value of the measurand [24]. Uncertainty and reliability were obtained by absolute and relative consistencies using the standard error of measurement (SEM) [25] (also known as the typical error of measurement (TEM) or typical error (TE) [26], beyond the measures generated from the SEM itself as the minimum difference (MD) or minimum detectable change (MDC)) and the intraclass correlation coefficient (ICC) [25–27]. The SEM is the measure of the amount of error variance in a set of obtained scores, different from the standard error of estimate (SEE), which is the standard deviation of true scores if the observed score is held constant [28]. Standard uncertainty is the uncertainty of the result of a measurement expressed as a standard deviation [29]. Precision is understood as the closeness of agreement between the
results of successive measurements of the same measurand, carried out under the same conditions of measurement. Precision is also called repeatability [24]. Limits of agreement (LOAs) refer to the reference interval for the test–retest differences expected for 95% of the population [30]. The coefficient of variation (CV) refers to the variation around the average, expressed as a percentage; data quality by means of the CV is interpreted by the level of instability [31].

Some reference values are assumed to better understand the results. SEM results lower than 5% have been suggested [32,33] and classified as good (SEM = < 5.0%), moderate (SEM = 5.0–9.9%), or poor (SEM ≥ 10.0%) [34–36]. However, it is recommended to present the corresponding value with confidence intervals (CIs) [37,38] because this is a measure of how much the measured test scores are spread around a “true” score [37,38]. Although the SEM may be better than the ICC [25] to evaluate reliability, both are reported in the present manuscript. According to Koo and Li [39], ICC values < 0.50 = poor, between 0.50 and 0.75 = moderate, between 0.75 and 0.90 = good, and >0.90 = indicative of excellent reliability. For the CV, it has been suggested that CV > 30.0% = large, CV < 30.0% and >10.0% = medium, and CV < 10.0% = small [40]. These reference values were used in the figure to determine the traffic-light systems (i.e., red, amber or green). For accuracy, MDC, SEE, precision, uncertainty and LOAs, no reference values were found in the scientific literature.

3. Results

The initial search returned 1035 articles (see Figure 1). After the removal of duplicates (n = 461), a total of 574 studies were retained for full-text screening. Following the eligibility assessment, 479 studies were excluded as they did not meet the inclusion criteria, while one full-text manuscript was not found. Finally, 94 studies were included in this systematic review [41–134].

3.1. Characteristics of the Studies and Risk of Bias

The pooled sample size and age included 2066 participants aged 24 ± 3 years, being composed mostly of male athletes (95%). The average duration of the load monitoring interventions was 168 days (range: 1–1034). The athletes were all professional soccer players from first divisions (81%), second divisions (12%), or third divisions (1%), and national teams (6%). A single study (1%) did not report the athletes’ level. Regarding the nationality of the soccer players, they were from England (16%), Spain (15%), Brazil (9%), Australia (7%), Portugal (7%), Netherlands (6%) Italy (6%), France (4%), Denmark (2%), Qatar (2%), Tunisia (2%), Russia (2%), Iran (2%), Austria (1%), Belgium (1%), Chile (1%), Colombia (1%), Czech Republic (1%), Germany (1%), Latvia (1%), Norway (1%), Singapore (1%), Switzerland (1%), and Wales (1%). Seven studies did not report the nationality of the athletes but reported the continent (Europe = 7%).

The average bias score for the studies was 6.6 (range: 4–8). Only one study had a high risk of bias, with a score ≤4 (see Table S1). There were 88 observational studies (i.e., 94%), with 82 prospective cohort studies [41,43–49,51,52,54–56,59–61,63–84,88,90,92–109,111–114,116–123,125–134], 3 retrospective cohort studies [85,110,115], 2 cross-sectional studies [50,62], and one diagnostic accuracy study [124]. There were six experimental studies (i.e., 6%), with three pre–post interventional studies [42,58,89], two cross-over randomized controlled trials [53,91], and one non-randomized trial [57] (see Table S2).

3.2. Main Findings

Most studies, i.e., 73%, did not report any veracity metric for all the identified parameters (n = 69). At least one veracity metric was reported in 18% of studies (n = 17), with a veracity metric for all parameters reported in 9% of studies (n = 8; see Figure ??). Thirty-nine different tools were used for load monitoring: electronic performance tracking systems (EPTS) = 64% (n = 60), rating of perceived exertion (RPE) = 48% (n = 45), heart rate monitoring (HR) = 29% (n = 27), countermovement jump test (CMJ) = 12% (n = 11), psycho-
metric questionnaire with Likert scale = 12% (n = 11), blood samples = 6% (n = 6), maximal running test = 6% (n = 6), salivary imunoglobulina A (s-IgA) = 6% (n = 6), blood lactate concentration = 5% (n = 5), video-computerized systems = 5% (n = 5), watch = 5% (n = 5), body composition = 3% (n = 3), integrative tool of training load assessment = 3% (n = 3), repeated sprint ability = 3% (n = 3), salivary cortisol = 3% (n = 3), salivary testostereone = 3% (n = 3), submaximal running test = 3% (n = 3), sprint test = 3% (n = 3), anaerobic speed reserve = 2% (n = 2), creatine kinase (CK) = 2% (n = 2), maximal oxygen uptake = 2% (n = 2), Visual Analog Scale Questionnaire = 2% (n = 2), total quality recovery (TQR) = 2% (n = 2), training planning = 2% (n = 2), actigraphy = 1% (n = 1), blood ammonia concentration = 1% (n = 1), coach rating of performance = 1% (n = 1), Edinburgh Mental Well-being Scale = 1% (n = 1), infrared thermography = 1% (n = 1), isometric force testing = 1% (n = 1), Neuro-muscular Efficiency Index = 1% (n = 1), oxidative stress = 1% (n = 1), Physical Activity Enjoyment Scale = 1% (n = 1), Profile of Mood State Questionnaire (POMS) = 1% (n = 1), RESTQ-Sport Scale = 1% (n = 1), sit-and-reach test = 1% (n = 1), squat jump test (SJ) = 1% (n = 1), tensiomyography = 1% (n = 1), and urine metabolomic = 1% (n = 1).

Figure 1. Study selection in a PRISMA flow diagram.

3.1. Characteristics of the Studies and Risk of Bias

The pooled sample size and age included 2066 participants aged 24 ± 3 years, being composed mostly of male athletes (95%). The average duration of the load monitoring interventions was 168 days (range: 1–1034). The athletes were all professional soccer players (81%), second divisions (12%), or third divisions (1%), and national teams (6%). A single study (1%) did not report the athletes’ level. Regarding the nationality of the soccer players, they were from England (16%), Spain (15%), Brazil (9%), Australia (7%), Portugal (7%), Netherlands (6%), Italy (6%), France (4%), Denmark (2%), Qatar (2%), Tunisia (2%), Russia (2%), Iran (2%), Austria (1%), Belgium (1%), Chile (1%), Colombia (1%), Czech Republic (1%), Germany (1%), Latvia (1%), Norway (1%), Singapore (1%), Switzerland (1%), and Wales (1%). Seven studies did not report the nationality of the athletes but reported the continent (Europe = 7%).

The average bias score for the studies was 6.6 (range: 4–8). Only one study had a high risk of bias, with a score ≤4 (see Table S1). There were 88 observational studies (i.e., 94%),
A total of 578 parameters were used to monitor the load from these 39 tools. Data veracity was reported for 54% of the tools in at least one parameter, which resulted in 23% of these parameters. Thus, most studies did not report metrics of veracity for their tools and parameters (see details in Figure ??, Table 1 and Table S2). Specifically, Table S2 presents a summary of all the selected studies, with further information on study design (and duration), sample level (n; sex; age; country of sample), tools (brand and model or reference), accuracy reported by the company, and parameters as well as veracity analyses (metrics).

| Parameter | CV | ICC | SEM |
|-----------|----|-----|-----|
| EPTS      | 6.5% (0.3%–83.5%) | 0.78 (0.61–0.94) | 7.5% (7.5%–30.0%) |
| RPE       | 7.6% (2.0%–29.0%) | 0.86 (0.77–0.89) | 17.0% (14.0%–20.0%) |
| HR Monitoring | 0.98 | 19.6% |
| CMJ       | 47.4% (3.7%–91.0%) | 1.7% (1.1%–1.8%) | 3.1% |
| Psychom Quest | 4.1% | S-IgA |
| Blood Samples | S-Cortisol | S-Testosterone |
| EPTS      | 6.5% (0.3%–83.5%) | 0.78 (0.61–0.94) | 7.5% (7.5%–30.0%) |
| RPE       | 7.6% (2.0%–29.0%) | 0.86 (0.77–0.89) | 17.0% (14.0%–20.0%) |
| HR Monitoring | 0.98 | 19.6% |
| CMJ       | 47.4% (3.7%–91.0%) | 1.7% (1.1%–1.8%) | 3.1% |
| Psychom Quest | 4.1% | S-IgA |
| Blood Samples | S-Cortisol | S-Testosterone |
| EPTS      | 6.5% (0.3%–83.5%) | 0.78 (0.61–0.94) | 7.5% (7.5%–30.0%) |
| RPE       | 7.6% (2.0%–29.0%) | 0.86 (0.77–0.89) | 17.0% (14.0%–20.0%) |
| HR Monitoring | 0.98 | 19.6% |
| CMJ       | 47.4% (3.7%–91.0%) | 1.7% (1.1%–1.8%) | 3.1% |
| Psychom Quest | 4.1% | S-IgA |
| Blood Samples | S-Cortisol | S-Testosterone |
| EPTS      | 6.5% (0.3%–83.5%) | 0.78 (0.61–0.94) | 7.5% (7.5%–30.0%) |
| RPE       | 7.6% (2.0%–29.0%) | 0.86 (0.77–0.89) | 17.0% (14.0%–20.0%) |
| HR Monitoring | 0.98 | 19.6% |
| CMJ       | 47.4% (3.7%–91.0%) | 1.7% (1.1%–1.8%) | 3.1% |
| Psychom Quest | 4.1% | S-IgA |
| Blood Samples | S-Cortisol | S-Testosterone |
| EPTS      | 6.5% (0.3%–83.5%) | 0.78 (0.61–0.94) | 7.5% (7.5%–30.0%) |
| RPE       | 7.6% (2.0%–29.0%) | 0.86 (0.77–0.89) | 17.0% (14.0%–20.0%) |
| HR Monitoring | 0.98 | 19.6% |
| CMJ       | 47.4% (3.7%–91.0%) | 1.7% (1.1%–1.8%) | 3.1% |
| Psychom Quest | 4.1% | S-IgA |
| Blood Samples | S-Cortisol | S-Testosterone |

Figure 2. Veracity metrics classifications. Median (minimum–maximum) and result (i.e., median) without range (i.e., minimum and maximum) because it was found just for one value (n = 1). Based on the current method, the respective reference values were used in this figure to determine the traffic-light systems (i.e., red, amber, or green). The median was used to
determine the color, and blank cells are due to the lack of veracity metrics for that case. $CV$ = coefficient of variation (standard deviation divided by the mean, expressed as a percentage, %); $ICC$ = intraclass correlation coefficient; $SEM$ = standard error of measurement; EPTS = electronic performance tracking system; RPE = rating of perceived exertion; HR = heart rate; CMJ = countermovement jump; Psychom Quest = Psychometric Questionnaire with Likert scale; S-IgA = salivary immunoglobulina A; S-Cortisol = salivary cortisol; S-Testosterone = salivary testosterone; CK = creatine kinase; MRT = maximal running test; RSA = repeated sprint ability; TQR = total quality recovery; IF = isometric force; POMS = Profile of Mood State Questionnaire; SJ = squat jump; Watch = wristwatch.

| Tools (n = 39): 54% (21/39) with Veracity Metrics | Parameters (n = 578): 23% (131/578) with Veracity Metrics | Veracity Metrics (Classification) |
|-------------------------------------------------|------------------------------------------------------------|----------------------------------|
| Total distance covered (m) (n = 8)              | SEM = 0.3–5.3%                                             | (Good–Moderate)                  |
|                                                  | CV = 8.3–43.0%                                             | (Small–Large)                    |
|                                                  | Bias = 0.8 ± 5.4 and 2–6% (Good–Moderate)                  |
| Average speed (m/min) (n = 3)                    | SEE = 4–5%                                                 | (Excellent)                      |
|                                                  | CV = 4.2%                                                  | (Good)                           |
| PlayerLoad (AU) (n = 3)                         | SEM = 0.6%                                                 | (Medium–Large)                   |
| Top speed (km/h) (n = 3)                        | CV = 20.0–37.0%                                            | (Good)                           |
| Distance covered at>21 km/h (m) (n = 2)         | SEM = 2.6–7.7%                                             | (Good–Moderate)                  |
| High metabolic power distance ($\geq$20 W/kg; m) (n = 2) | CV = 10%                                                  | (Small)                          |
| High-speed running (19.8–25.2 km/h; m) (n = 2)  | SEM = 4.5%                                                 | (Good)                           |
| Maximum acceleration ($>3$ m/s$^2$; m) (n = 2)   | CV = 32.0–36.6%                                            | (Large)                          |
| Deceleration effort ($>3$ m/s$^2$; number) (n = 1) | SEM = 5.6%                                                | (Large)                          |
| Decelerations ($\geq$3.3 m/s$^2$; number) (n = 1) | CV = 0.3 ± 2.5                                            | (Small)                          |
| Distance at acceleration zone count 2–3 (2 m/s$^2$–3 m/s$^2$; m) (n = 1) | SEM = 30.0%                                              | (Large)                          |
| Distance at deceleration zone count 2–3 ($−2$–$−3$ m/s$^2$; m) (n = 1) | CV = 31.0%                                               | (Large)                          |
| Distance covered between 13–18 km/h (m) (n = 1) | SEM = 13.3–17.4%                                          | (Poor)                           |
| Distance covered between 17–21 km/h (m) (n = 1) | SEM = 141.0–146.0%                                        | (Poor)                           |
| Distance covered between 18–21 km/h (m) (n = 1) | SEM = 21.0–25.5%                                          | (Poor)                           |
| Distance in Power Zone 20–25 w/kg (km) (n = 1)  | CV = 32.0%                                                | (Large)                          |
| Distance in Speed Zone 2 (no descriptions; km)  | CV = 35.0%                                                | (Large)                          |
| Duration (min) (n = 1)                          | SEM = 2.5%                                                | (Excellent)                      |
| Dynamic stress load (AU) (n = 1)                | ICC = 0.94                                                | (Excellent)                      |
| Efforts/sprints >21 km/h (number) (n = 1)        | SEM = 44.5–49.0%                                          | (Poor)                           |
| High-intensity events (5.4–9.0 km/h; m) (n = 1) | CV = 19.0%                                                | (Medium)                         |
| High-intensity events (9.0–12.6 km/h; m) (n = 1) | CV = 17.0%                                                | (Medium)                         |
| High-intensity events (12.6–14.4 km/h; m) (n = 1) | CV = 16.0%                                                | (Medium)                         |
| High-intensity running distance/min ($>15.1$ km/h; m/min) (n = 1) | SEM = 30.6%                                              | (Poor)                           |
| High-intensity running distance (16–22 km/h; m) (n = 1) | CV = 25.2%                                                | (Medium)                         |
| High-speed distance (19.8–25.5 km/h; m) (n = 1) | SEM = 8.1%                                                | (Moderate)                       |

1. **EPTS** in 64% (60/94) of Studies and 20% (59/294) of Parameters with Veracity Metrics.
| Parameter Description                                      | n  | CV  | Bias   | MDC  |
|------------------------------------------------------------|----|-----|--------|------|
| High-speed running (19.8–25.1 km/h; %) (n = 1)             |    |     |        |      |
| High-speed running (individualized; %) (n = 1)             |    |     |        |      |
| High-speed running distance (14.4–19.8 km/h; m) (n = 1)    |    |     |        |      |
| High-speed running time (17–21 km/h; min) (n = 1)          |    |     |        |      |
| Low-speed activities (<14.4 km/h; %) (n = 1)               |    |     |        |      |
| Low-speed activities (individualized; %) (n = 1)           |    |     |        |      |
| Moderate-speed running (14.4–19.8 km/h; %) (n = 1)        |    |     |        |      |
| PlayerLoad2D (AU) (n = 1)                                 |    |     |        |      |
| PlayerLoadV (%) (n = 1)                                   |    |     |        |      |
| Power score (w/kg) (n = 1)                                 |    |     |        |      |
| Sprinting (>25.2 km/h; %) (n = 1)                         |    |     |        |      |
| Sprinting (individualized; %) (n = 1)                     |    |     |        |      |
| Sprint distance (>25.5 km/h; m) (n = 1)                   |    |     |        |      |
| Sprint distance (no descriptions; km) (n = 1)             |    |     |        |      |
| Total acceleration distance (m) (n = 1)                   |    |     |        |      |
| Total deceleration distance (m) (n = 1)                   |    |     |        |      |
| Very high-speed distance (>19.8 km/h; m) (n = 1)          |    |     |        |      |
| Very high speed running distance (>22 km/h; m) (n = 1)    |    |     |        |      |
| Very high speed running time (>21 km/h; min) (n = 1)      |    |     |        |      |
| Foster session-RPE (AU) (n = 4)                            |    |     |        |      |
| Borg RPE scores (Scale 6–20 by Borg, 1982) (n = 1)        |    |     |        |      |
| Carrie session-RPE (Scale 0–10 by Carrie, 2012) (n = 1)  |    |     |        |      |
| Monotony (Scale 1–10 by Foster et al., 2001) (n = 1)      |    |     |        |      |
| Morandi session-RPE (Scale 0–10 by Morandi et al., 2020) |    |     |        |      |
| Morandi RPE score (Scale 0–10 by Morandi et al., 2020)   |    |     |        |      |
| Muscular session-RPE (Scale 1–10 by Foster et al., 2001)  |    |     |        |      |
| Respiratory session-RPE (Scale 1–10 by Foster et al., 2001) |    |     |        |      |
| Strain (Scale 1–10 by Foster et al., 2001)                |    |     |        |      |
| Sum of all muscular RPE scores (Scale 1–10 by Foster et al., 2001) |    |     |        |      |
| Sum of all respiratory RPE scores (Scale 1–10 by Foster et al., 2001) |    |     |        |      |
| Mean percentage of maximum HR (%) (n = 3)                 |    |     |        |      |
| HR mean (bpm) (n = 2)                                     |    |     |        |      |
| Ln RMSSD (ms) (n = 2)                                     |    |     |        |      |
| Banister’s TRIMP (Banister, 1991) (n = 1)                 |    |     |        |      |
| Maximal HR (bpm) (n = 1)                                  |    |     |        |      |
| Time above 85% HRmax (%) (n = 1)                          |    |     |        |      |
| TRIMPmod (Stagno et al., 2007) (n = 1)                    |    |     |        |      |
### Table 1. Cont.

|   |   |   |   |
|---|---|---|---|
|   |   |   |   |
| **CMJ Test** |   |   |   |
| 4 | in 12% (11/94) of Studies and 8% (1/2) of Parameters with Veracity Metrics | Jump height (cm) (n = 1) | ICC = 0.98 |
|   |   |   |   |
| **Psychometric Questionnaire with Likert scale** |   |   |   |
| 5 | in 12% (11/94) of Studies and 56% (5/9) of Parameters with Veracity Metrics | Hooper Index (n = 2) | SEM = 19.6% (Poor) |
|   |   | Energy Levels (n = 1) | CV = 15.0% (Medium) |
|   |   | Lower body soreness (n = 1) | CV = 14.0% (Medium) |
|   |   | Readiness to Train (n = 1) | CV = 20.0% (Medium) |
|   |   | Sleep quality (n = 1) | CV = 20.0% (Medium) |
| **Blood Samples** |   |   |   |
| 6 | in 6% (6/94) of Studies and 13% (6/48) of Parameters with Veracity Metrics | Cortisol (n = 1) | ICC = 0.74 (Moderate) |
|   |   | Plasma CK (n = 1) | CV = 1.8% (Small) |
|   |   | Plasma CRP (n = 1) | CV = 1.7% (Small) |
|   |   | Plasma LDH (n = 1) | CV = 1.1% (Small) |
|   |   | Testosterone (n = 1) | ICC = 0.83 (Good) |
|   |   | T/C ratio (n = 1) | ICC = 0.82 (Good) |
| **Maximal Running Test** |   |   |   |
| 7 | in 6% (6/94) of Studies and 40% (2/5) of Parameters with Veracity Metrics | Final velocity of the 30-15 Intermittent Fitness Test (Buchheit, 2008) (n = 1) | SEM = 7.7–8.0% (Moderate) |
|   |   | Total distance covered (m) Yo-Yo Intermittent Recovery Test Level 1 (Bangsbo et al., 2008) (n = 1) | ICC = 0.48 (Poor) |
| **Salivary Immunoglobulina A** |   |   |   |
| 8 | in 6% (6/94) of Studies and 10% (1/5) of Parameters with Veracity Metrics | Concentration of IgA (n = 1) | CV = 3.7–91.0% (Small–Large) |
| **Blood Lactate Concentration** |   |   |   |
| 9 | in 5% (5/94) of Studies and 0% (0/8) of Parameters with Veracity Metrics |   |   |
| **Video-Computerized System** |   |   |   |
| 10 | in 5% (5/94) of Studies and 0% (0/11) of Parameters with Veracity Metrics |   |   |
| **Watch/Wristwatch** |   |   |   |
| 11 | in 5% (5/94) of Studies and 50% (1/2) of Parameters with Veracity Metrics | Trained/played minutes (min) (n = 1) | CV = 5.5% (Small) |
| **Body Composition** |   |   |   |
| 12 | in 3% (3/94) of Studies and 0% (0/3) of Parameters with Veracity Metrics |   |   |
| **Integrative Tool of Training Load Assessment** |   |   |   |
| 13 | in 3% (3/94) of Studies and 0% (0/3) of Parameters with Veracity Metrics |   |   |
| **Repeated Sprint Ability** |   |   |   |
| 14 | in 3% (3/94) of Studies and 36% (4/11) of Parameters with Veracity Metrics | Mean (s) (n = 1) | ICC = 0.61 (Moderate) |
|   |   | Best (s) (n = 1) | ICC = 0.69 (Moderate) |
|   |   | Decrement (%) (n = 1) | ICC = 0.34 (Poor) |
| **Salivary Cortisol** |   |   |   |
| 15 | in 3% (3/94) of Studies and 50% (1/2) of Parameters with Veracity Metrics | Concentration of cortisol (n = 1) | CV = 3.1% (Small) |
| **Salivary Testosterone** |   |   |   |
| 16 | in 3% (3/94) of Studies and 100% (1/1) of Parameters with Veracity Metrics | Concentration of testosterone (n = 1) | CV = 4.1% (Small) |
| **Submaximal Running Test** |   |   |   |
| 17 | in 3% (3/94) of Studies and 0% (0/2) of Parameters with Veracity Metrics |   |   |
| **Sprint Test** |   |   |   |
| 18 | in 2% (2/94) of Studies and 0% (0/6) of Parameters with Veracity Metrics |   |   |
| **Anaerobic Speed Reserve** |   |   |   |
| 19 | in 2% (2/94) of Studies and 36% (4/11) of Parameters with Veracity Metrics | Meters covered at maximal aerobic speed (m) (n = 1) | MDC = 116–144% |
|   |   | Time spent at maximal aerobic speed (min) (n = 1) | MDC = 112–174% |
|   |   | Meters covered at 30% anaerobic speed reserve (m) (n = 1) | MDC = 73–145% |
|   |   | Time spent at 30% anaerobic speed reserve (min) (n = 1) | MDC = 7–116% |
| **CK** |   |   |   |
| 20 | in 2% (2/94) of Studies and 33% (1/3) of Parameters with Veracity Metrics | CK absolute concentration (µ/L) (n = 1) | SEM = 2.3% (Good) |
| **Maximal Oxygen Uptake** |   |   |   |
| 21 | in 2% (2/94) of Studies and 0% (0/4) of Parameters with Veracity Metrics |   |   |
| **Visual Analogue Scale Questionnaire** |   |   |   |
| 22 | in 2% (2/94) of Studies and 0% (0/7) of Parameters with Veracity Metrics |   |   |
| **Total Quality Recovery** |   |   |   |
| 23 | in 2% (2/94) of Studies and 50% (1/2) of Parameters with Veracity Metrics | Recovery level (Morandi et al., 2020) (n = 1) | SEM = 30.0% (Poor) |
|   |   | CV = 12.0% (Medium) |
|   |   | ICC = 0.77 (Good) |
4. Discussion

The purpose of this systematic review is to evaluate data veracity for load monitoring in professional soccer. We describe here the metrics reported for tools and their associated parameters in a sample with 87% of athletes playing at top-level divisions or in their national teams. Thirty-nine different tools were found in the included studies; however, 73% of these studies did not report any veracity metric. Thus, data veracity was found for 54% of tools and 23% of associated parameters. Of note, some veracity metrics present
a great variability in the obtained metrics. The SEM and MDC ratings were between 0.3–89.7% (good–poor) and 48.0–305.0%, respectively. The ICC ratings were between poor and excellent reliability levels (i.e., 0.004–0.98), with the CV ratings from poor to great (i.e., 1.1–193.5%). For the remaining metrics, the results were between 2.0–10.0% for bias, 4.0–22.0% for SEE, and 3.0–10.0% for precision. In addition, some tools were used without the required accuracy. All this information provides a precise state of the art regarding the potential of current monitoring tools to be used with Big Data approaches.

4.1. The Impact of Accuracy and CV on the Practical Application of Tools and Parameters

The need for reporting accuracy can be better understood following an example reported by the United Nations Industrial Development Organization [24]. Thus, in practical terms, if the accuracy of a tool is ±3, this means that if 100 would be the reading displayed on a tool during a measurement, then the actual value could be anywhere between 97 and 103, including 97 and 103. In this regard, it is important to note that some tools did not present the required accuracy for the experimental design performed in each study. For example, the accuracy of an infrared camera is ±2 °C, and the applied protocol in the study provided for changes in body temperature ranged between 0.3 and 1.5 °C [98]. Moreover, in EPTS with a sampling rate of 10 Hz, the accuracy for high accelerations (>4 m/s²) is compromised, which may have an impact on the interpretation of results [79] because of the poor estimation of instantaneous velocity when performing these very high accelerations (>4 m/s²) during player tracking in team sports [135]. Therefore, tools with inadequate accuracy for the required interventions need to improve their accuracy or else they will generate an input noise, thus compromising the outputs of a Big Data approach.

A common concern of sports scientists as well as strength and conditioning coaches of high-level soccer clubs when using monitoring tools refers to reliability, which is one of the main factors related to the discrepancy between the expected and actual effectiveness of monitoring, thus representing a potential barrier for successful interventions [136]. For this purpose, the classic classification of the CV is considered for the decision-making process: >30% = large, <30% and >10% = medium, and <10% = small [40]. In the present systematic review, extremely large CVs, such as 193.5% for sprinting at speeds ≥25.2 km/h, are measured with EPTS [129]. Moreover, 137% for the time between peak power and peak force during a CMJ [137] and 679% for the variance from mean bedtime assessed by Actigraphy [138] have been found in the scientific literature. Large CVs make it extremely difficult to detect the real differences between moments after an intervention unless these differences were also very large [139]. Therefore, caution should be taken when using parameters with a CV greater than 30%.

4.2. ICC and SEM Should Always Be Reported Together

The ICC is probably the most popular reliability metric in the literature. Our analysis showed that the ICC was often included as a reliability metric but with large variability, ranging from poor = 0.004 (i.e., strain from session-RPE) [131] to excellent = 0.98 (i.e., contraction time measured with tensiomyography or jump height of a CMJ) [57,131]. Particularly, in the case of tensiomyography, although there is sufficient data in favor of its good-to-excellent relative reliability in the scientific literature (i.e., ICC = 0.70–0.99) [140] in agreement with our findings (i.e., ICC = 0.80–0.98), more evidence is necessary for identifying the accuracy of this tool, according to a current systematic review with meta-analysis [141], mainly because reliability based only on the ICC cannot be recommended [142]. This is because there is a relationship between the ICC and between-subjects variability, where the heterogeneity of the subjects is a determinant for the ICC obtained [25]. An excellent ICC can mask poor trial-to-trial consistency when between-subjects variability is high. Conversely, a poor ICC can be found even when trial-to-trial variability is low when the between-subjects variability is low. In this case, the homogeneity of the subjects’ means will make it difficult to differentiate between subjects even though the absolute measurement error is small. In other words, if individuals differ little from each other, the
ICC may be poor, even if the trial-to-trial variability is small [25]. Thus, an examination of the ICC in conjunction with the SEM is, therefore, recommended [142]. For example, an average HR, with ICC = 0.77 (good) and SEM = 3.0% (good), in one case [127] would be preferable than the same parameter with ICC = 0.89 (good) and SEM = 27% (poor) [124].

On the other hand, the SEM is not affected by between-subjects variability, as occurs with the ICC. The SEM has been used to define the boundaries in which a subject’s true score lies, and it can be calculated as both absolute and relative scores [25,26,37,38]. However, the range of variation of this veracity metric for the found parameters was high (i.e., between good = 0.3% for total distance covered, as measured with EPTS, and poor = 89.7% for accelerations >4 m/s$^2$, measured with EPTS) [64,73]. This would suggest that for the same tool, some parameters could be more recommended than others. In this regard, some researchers are using a limit of 10% of the SEM to determine which parameters can be used in subsequent analyses [35,87,143]. Although its use is strongly recommended by the specialized literature [25,26], some points of confusion can be highlighted. The first is that it is called either standard error of measurement (SEM), typical error of measurement (TEM), or typical error (TE), even with the same form of calculation. The second point, which is most impactful, is that the SEM is commonly referred to as the coefficient of variation (CV) and reported as a percentage, but the calculation is not performed by dividing the standard deviation by the mean and subsequently multiplied by 100 [144]. These facts can easily drive researchers and practitioners to misinterpretations. Furthermore, the SEM can also be used in the interpretation of individual scores as the minimum difference (MD) or minimum detectable change (MDC) [25,26]. In the present study, the MDC ranged from 48.0% for the distance at maximal sprint speed recorded with EPTS to 305.0% for the very high-speed running time obtained with EPTS [76] with regards to the group average. This extremely high variability could hinder the verification of real changes in an individual’s performance. Of note, the MDC in professional soccer players is not easily found in scientific literature. Apart from the included study in the current systematic review, another study reported an MDC between 1.0% and 30.0% for performance in a new agility test [145]. Therefore, the SEM is highly recommended for determining the reliability of any tool and parameter, while its interpretation can be facilitated when reported in both relative and absolute terms.

4.3. Needs, Limitations, and Potential of Big Data in Professional Soccer

There is limited information regarding the last three metrics of veracity. Bias was reported as absolute values in only two studies [127,129], and, in one study, it was reported as relative values ranging between 2.0–10.0%, with the SEE ranging between 4.0–22.0% for total and sprint distances between 19.8–25.2 and above 25.2 km/h, recorded with EPTS [133]. Precision ranged between 3.0–10.0% for urine metabolomic parameters in a single study [128]. In view of the definition itself, these three metrics are important for appropriately understanding veracity. Therefore, more studies with the use of these metrics in professional soccer athletes are warranted as well as more studies with professional female soccer players, who represented only 5% of the samples in the included studies. Moreover, 94% of the included articles are observational studies, and among the experimental study designs, none were randomized controlled trials. Despite the difficulty of carrying out experimental studies in professional soccer settings, we suggest that more of these studies are warranted to increase the veracity of monitoring tools and their associated parameters. The approach applied in this systematic review is based on one of the 4 Vs of Big Data. However, the other Vs also have an impact on soccer and should be the focus of future research. In this regard, the volume that describes the magnitude of the data [14] and is usually measured in terabytes (i.e., $10^{12}$), petabytes (i.e., $10^{15}$), zettabytes (i.e., $10^{21}$), and even yottabytes (i.e., $10^{24}$) [15] has increased exponentially, with a 300-fold increase in the last 15 years, from 2005 to 2020, reaching 40 zettabytes (i.e., $37.3$ trillion gigabytes; $10^{24}$) [15]. Therefore, this unprecedented volume of data is overwhelming, thus increasing the risks of it not being fully used to inform the practice [13]. For example, a
dataset from a Bundesliga season resulted in 400 gigabytes of tracking data [14]. Regarding this, variety refers to the heterogeneity of data, i.e., to different data formats and sources distinguished, which can be differentiated among structured (e.g., relational data), semi-structured, (e.g., XML data), and unstructured data (e.g., emails, pictures, videos, or social networking data) [14–17]. Thus, in soccer, data variety refers to position, video, fitness, training, skill performance, and health data [14]. Moreover, velocity describes the speed at which novel data is generated, processed, and analyzed [14–17]. Specifically for soccer, the velocity may vary between real-time streams from physiological and positional data to stored data tonotational analyses during training and competition [14]. All these three key concepts characterizing Big Data are highly relevant and should be considered with the suggestion by Lukoianova and Rubin [23], who have stated that Big Data can only have value when its veracity can be established and, thereby, the information quality confirmed.

To date, Big Data has been reported in the scientific literature of soccer in relation to tactical analysis only [14,146] but not in relation to load monitoring. Although data can be considered big based on the number of V2019s, there are no universal benchmarks for the number of Vs [14–18,21–23,146,147]. Therefore, the current systematic review establishes the initial basis for the implementation of Big Data approaches for load monitoring in professional soccer while identifying the strengths and limitations of the current evidence before moving on to the real applications of Big Data: data management, which involves processes and supporting technologies to acquire and store data and to prepare and retrieve it for analysis, and analytics, which refers to the techniques used to analyze and acquire intelligence from Big Data [146,147]).

Finally, we would like to make a brief statement and endorse an editorial on the current trend of having extreme positions in sportsscience, which is not recommended [148]. Thus, this review highlights the importance of data veracity analysis for objectively using tools and their associated parameters for load monitoring in professional soccer, considering equally traditional approaches and the forefront of technological advances.

5. Conclusions

In conclusion, a wide diversity of tools and parameters are used to monitor loads in elite professional soccer. However, it is not common to find data veracity for these tools and parameters in scientific literature. The reported veracity metrics will assist in the selection and use of the best monitoring tools and their associated parameters for load monitoring in professional soccer. Before looking for new tools and parameters, the current ones need to present adequate levels of data veracity (accuracy, reliability, and quality of data). Therefore, this information is warranted when aiming to use predictive analytics for structured Big Data in professional soccer.

Practical Applications

1. The use of Big Data approaches without appropriate data veracity can undermine the precision of the predictive analytics models and generate fatal errors with a high economic cost;
2. Up to the moment, data veracity is not commonly reported in scientific studies using tools and parameters for load monitoring in professional soccer;
3. It is necessary to more frequently analyze and share data veracity to perform the best data management and analytics when applying Big Data.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/app11146479/s1, Table S1: Risk of bias score; Table S2: Summary of the selected studies. Reference [149] is cited in supplementary materials.
Author Contributions: Conceptualization, J.G.C., C.A.C.F., D.B., and J.C.S.; methodology, J.G.C., C.A.C.F., D.B., A.L.-A., G.R.C., R.L.d.S.G., R.d.S.G., F.M.V., A.L.C.A., S.D.R., and J.C.S.; formal analysis, J.G.C., C.A.C.F., D.B., A.L.-A., G.R.C., R.L.d.S.G., R.d.S.G., F.M.V., A.L.C.A., S.D.R., J.A., and J.C.S.; writing—original draft preparation, J.G.C., C.A.C.F., D.B., A.L.-A., G.R.C., R.L.d.S.G., R.d.S.G., F.M.V., A.L.C.A., S.D.R., J.A., and J.C.S.; writing—review and editing, J.G.C., C.A.C.F., D.B., A.L.-A., S.D.R., J.A., and J.C.S.; supervision, J.C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: After publication, all data necessary to understand and assess the conclusions of the manuscript are available to any reader of Applied Sciences.

Acknowledgments: We would like to thank the “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/ Programa de Excelência Acadêmica” (CAPES/PROEX).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jaspers, A.; Brink, M.S.; Probst, S.G.M.; Frencken, W.G.P.; Helsen, W.F. Relationships between training load indicators and training outcomes in professional soccer. Sports Med. 2017, 47, 533–544. [CrossRef]

2. Boullosa, D.; Casado, A.; Claudino, J.G.; Jiménez-Reyes, P.; Ravé, G.; Castaño-Zambudio, A.; Lima-Alves, A.; de Oliveira, S.A.J.; Dupont, G.; Granacher, U.; et al. Do you play or do you train? Insights from individual sports for training load and injury risk management in team sports based on individualization. Front. Physiol. 2020, 11, 995. [CrossRef]

3. Ekstrand, J.; Waldén, M.; Hägglund, M. Hamstring injuries have increased by 4% annually in men’s professional football, since 2001: A 13-year longitudinal analysis of the UEFA Elite Club injury study. Br. J. Sports Med. 2016, 50, 731–737. [CrossRef]

4. Barnes, C.; Archer, D.T.; Hogg, B.; Bush, M.; Bradley, P.S. The evolution of physical and technical performance parameters in the english premier league. Int. J. Sports Med. 2014, 35, 1095–1100. [CrossRef]

5. Haugen, T.A.; Tønnessen, E.; Hem, E.; Leirstein, S.; Seiler, S. VO2max characteristics of elite female soccer players, 1989–2007. Int. J. Sports Physiol. Perform. 2014, 9, 515–521. [CrossRef] [PubMed]

6. Tønnessen, E.; Hem, E.; Leirstein, S.; Haugen, T.; Seiler, S. Maximal aerobic power characteristics of male professional soccer players, 1989–2012. Int. J. Sports Physiol. Perform. 2013, 8, 323–329. [CrossRef] [PubMed]

7. Alves, A.L.; Garcia, E.S.; Morandi, R.F.; Claudino, J.G.; Pimenta, E.M.; Soares, D.D. Individual analysis of creatine kinase concentration in Brazilian elite soccer players. Rev. Bras. Med. Esporte 2015, 21, 112–116. [CrossRef]

8. Sølen, T.; Chamari, K.; Castagna, C.; Wisloff, U. Physiology of soccer: An update. Sports Med. 2005, 35, 501–536. [CrossRef] [PubMed]

9. Moreira, A.; Bradley, P.; Carling, C.; Arruda, A.; Spigolon, L.M.P.; Franciscon, C.; Aoki, M.S. Effect of a congested match schedule on immune-endocrine responses, technical performance and session-RPE in elite youth soccer players. Int. J. Sports Physiol. Perform. 2014, 10, 995. [CrossRef]

10. Bourdon, F.C.; Cardinale, M.; Murray, A.; Gastin, P.; Kellmann, M.; Varley, M.; Gabbett, T.J.; Coutts, A.J.; Burgess, D.J.; Gregson, W.; et al. Monitoring athlete training loads: Consensus statement. Int. J. Sports Physiol. Perform. 2017, 12, S2–S161. [CrossRef] [PubMed]

11. Claudino, J.G.; Capanema, D.D.O.; De Souza, T.V.; Serra, M.C.; Pereira, A.C.M.; Nassis, G.P. Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: A systematic review. Sports Med. Open 2019, 5, 28. [CrossRef] [PubMed]

12. Düküng, P.; Achtzehn, S.; Holmberg, H.-C.; Sperlich, B. Integrated Framework of Load Monitoring by a Combination of smartphone applications, wearables and point-of-care testing provides feedback that allows individual adaptive adjustments to activities of daily living. Sensors 2018, 18, 1632. [CrossRef] [PubMed]

13. Passfield, L.; Hopker, J.G. A mine of information: Can sports analytics provide wisdom from your data? Int. J. Sports PhysiolPerform. 2017, 12, 851–855. [CrossRef] [PubMed]

14. Reif, R.; Memmert, D. Big data and tactical analysis in elite soccer: Future challenges and opportunities for sports science. SpringerPlus 2016, 5, 1410. [CrossRef] [PubMed]

15. IBM. The Four V’s of Big Data. Available online: https://www.ibmbigdatahub.com/infographic/four-vs-big-data (accessed on 9 August 2020).

16. Ristevski, B.; Chen, M. Big Data Analytics in Medicine and Healthcare. J. Integr. Bioinform. 2018, 15. [CrossRef]

17. Kalid, N.; Zaidan, A.A.; Zaidan, B.B.; Salman, O.; Hashim, M.; Muzammil, H. Based Real Time Remote Health Monitoring Systems: A Review on Patients Prioritization and Related “Big Data” Using Body Sensors information and Communication Technology. J. Med Syst. 2017, 42, 30. [CrossRef]
18. Reimer, A.; A Madigan, E. Veracity in big data: How good is good enough. *Health Inform. J.* 2019, 25, 1290–1298. [CrossRef]

19. Mohrer, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* 2009, 6, e1000097. [CrossRef]

20. Saw, A.E.; Main, L.C.; Gastin, P.B. Monitoring the athlete training response: Subjective self-reported measures trump commonly used objective measures: A systematic review. *Br. J. Sports Med.* 2016, 50, 281–291. [CrossRef] [PubMed]

21. Ramachandramurthy, S.; Subramaniam, S.; Ramasamy, C. Distilling Big Data: Refining quality information in the Era of Yottabytes. *Sci. World J.* 2015, 2015, 453997. [CrossRef] [PubMed]

22. Bellazzi, R. Big data and biomedical informatics: A challenging opportunity. *Yearb. Med. Inform.* 2014, 23, 8–13. [CrossRef] [PubMed]

23. Lukoianova, T.; Rubin, V.L.; Vashchilko, T. Veracity Roadmap: Is big data objective, truthful and credible? *Adv. Classif. Res. Online* 2014, 24, 4. [CrossRef]

24. United Nations Industrial Development Organization (UNIDO). Role of Measurement and Calibration in the Manufacture of Products for the Global Market: A Guide for Small and Medium-Sized Enterprises. Available online: https://www.unido.org/sites/default/files/2009-04/Role_of_measurement_and_calibration_pdf (accessed on 9 August 2020).

25. Weir, J.P. Quantifying test-retest reliability using the intraclass correlation coefficient and the sem. *J. Strength Cond. Res.* 2005, 19, 231–240. [CrossRef]

26. Hopkins, W.G. Measures of reliability in sports medicine and science. *Sports Med.* 2000, 30, 1–15. [CrossRef]

27. Tanner, R.K.; Gore, C.J. Quality assurance in Australian exercise-physiology laboratories—In quest of excellence. *Int. J. Sports Physiol. Perform.* 2006, 1, 58–60. [CrossRef]

28. Charter, R.A. Revisiting the standard errors of measurement, estimate, and prediction and their application to test scores. *Percept. Mot. Ski.* 1996, 82, 1139–1144. [CrossRef]

29. Farrance, I.; Frenkel, R. Uncertainty of measurement: A review of the rules for calculating uncertainty components through functional relationships. *Clin. Biochem. Rev.* 2012, 33, 49–75.

30. Currell, K.; Jeukendrup, A.E. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med.* 2008, 38, 297–316. [CrossRef]

31. Sokal, R.R.; Braumann, C.A. Significance tests for coefficients of variation and variability profiles. *Syst. Zool.* 1980, 29, 50–66. [CrossRef]

32. Hopkins, W.G. A New View of Statistics. 1997. Available online: http://www.sportsci.org/resource/stats/effectmag.html (accessed on 7 May 2021).

33. Dugdale, J.H.; Sanders, D.; Hunter, A.M. Reliability of change of direction and agility assessments in youth soccer players. *Sports 2020*, 8, 51. [CrossRef]

34. McInnes, S.; Carlson, J.; Jones, C.; McKenna, M.; McInnes, S.; Carlson, J.; Jones, C.; McKenna, M. The physiological load imposed on basketball players during competition. *J. Sports Sci.* 1995, 13, 387–397. [CrossRef]

35. Duthie, G.; Pyne, D.; Hooper, S. The reliability of video based time motion analysis. *J. Hum. Mov. Stud.* 2003, 44, 259–271.

36. Delaney, J.A.; Cummins, C.J.; Thornton, H.; Duthie, G. Importance, Reliability, and Usefulness of Acceleration Measures in Team Sports. *J. Strength Cond. Res.* 2018, 32, 3485–3493. [CrossRef] [PubMed]

37. Patterson, C.H. The Interpretation of the Standard Error of Measurement. *J. Exp. Educ.* 1955, 23, 247–252. [CrossRef]

38. Harvill, L.M. Standard error of measurement. *Educ. Meas.* 1991, 10, 33–41. [CrossRef]

39. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* 2016, 15, 155–163. [CrossRef] [PubMed]

40. Lande, R. On comparing coefficients of variation. *Syst. Zool.* 1977, 26, 214–217. [CrossRef]

41. Little, T.; Williams, A.G. Measures of exercise intensity during soccer training drills with professional soccer players. *J. Strength Cond. Res.* 2007, 21, 367–371. [CrossRef]

42. Moreira, A.; Arsati, F.; Cury, P.R.; Franciscon, C.; de Oliveira, P.; Araújo, V. Salivary Immunoglobulin A Response to a Match in Top-Level Brazilian Soccer Players. *J. Strength Cond. Res.* 2009, 23, 1968–1973. [CrossRef]

43. Gomez-Piriz, P.T.; Reyes, P.J.; Ruiz-Ruiz, C. Relation Between Total Body Load and Session Rating of Perceived Exertion in Professional Soccer Players. *J. Strength Cond. Res.* 2011, 25, 2100–2103. [CrossRef]

44. Akubat, I.; Patel, E.; Barrett, S.; Abt, G. Methods of monitoring the training and match load and their relationship to changes in fitness in professional youth soccer players. *J. Sports Sci.* 2012, 30, 1473–1480. [CrossRef]

45. Bara-Filho, M.G.; Freitas, D.S.; Moreira, D.; Matta, M.D.O.; De Lima, J.R.P.; Nakamura, F.Y. Heart rate variability and soccer training: A case study. *Matriz Revista Educação Física* 2013, 19, 171–177. [CrossRef]

46. Heisterberg, M.F.; Fahrenkrug, J.; Krstrup, P.; Storskov, A.; Kjær, M.; Andersen, J.L. Extensive Monitoring Through Multiple Blood Samples in Professional Soccer Players. *J. Strength Cond. Res.* 2013, 27, 1260–1271. [CrossRef] [PubMed]

47. Bujnovsky, D.; Maly, T.; Zahalka, F.; Maly, A. Analysis of physical load among professional soccer players during matches with respect to field position. *J. Phys. Educ. Sport.* 2015, 15, 569–576.

48. Gaudino, P.; Iaia, F.M.; Strudwick, A.J.; Hawkins, R.D.; Alberti, G.; Atkinson, G.; Gregson, W. Factors Influencing Perception of Effort (Session Rating of Perceived Exertion) during Elite Soccer Training. *Int. J. Sports Physiol. Perform.* 2015, 10, 860–864. [CrossRef] [PubMed]
49. Arcos, A.L.; Martínez-Santos, R.; Yanci, J.; Mendiguchia, J.; Méndez-Villanueva, A. Negative Associations between Perceived Training Load, Volume and Changes in Physical Fitness in Professional Soccer Players. *J. Sports Sci. Med.* 2015, 14, 394–401. [CrossRef]

50. Morcillo, J.A.; Reyes, P.J.; Cuadrado-Peñaflieal, V.; Lozano, E.; Ortega-Becerra, M.; Parraga, J. Relationships Between Repeated Sprint Ability, Mechanical Parameters, and Blood Metabolites in Professional Soccer Players. *J. Strength Cond. Res.* 2015, 29, 1673–1682. [CrossRef]

51. Morgans, R.; Owen, A.; Doran, D.; Drust, B.; Morton, J.P. Prematch Salivary Secretory Immunoglobulin A in Soccer Players from the 2014 World Cup Qualifying Campaign. *Int. J. Sports Physiol. Perform.* 2015, 10, 401–403. [CrossRef]

52. Thorpe, R.T.; Strudwick, A.J.; 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, 958–964. [CrossRef]

53. Torres-Ronda, L.; Gonçalves, B.; Marcelino, R.; Torrents, C.; Vicente, E.; Sampajo, J. Heart Rate, Time-Motion, and Body Impacts When Changing the Number of Teammates and Opponents in Soccer Small-Sided Games. *J. Strength Cond. Res.* 2015, 29, 2723–2730. [CrossRef]

54. Buchheit, M.; Cholley, Y.; Lambert, P. Psychometric and Physiological Responses to a Preseason Competitive Camp in the Heat With a 6-Hour Time Difference in Elite Soccer Players. *Int. J. Sports Physiol. Perform.* 2016, 11, 176–181. [CrossRef] [PubMed]

55. Coelho, D.B.; Cabido, C.E.T.; Ciminelli, V.A.L.; Coelho, L.G.M.; Becker, L.K.; De Oliveira, E.C.; Pereira, E.R.; Marins, J.C.B.; Garcia, E.S. Comparison of different ways of expressing creatine kinase concentration of soccer players during a competitive season. *Motriz Revista Educação Física* 2016, 22, 160–165. [CrossRef]

56. Ehrmann, P.F.; Duncan, C.S.; Sindsusak, D.; Franzsen, W.; Greene, D. GPS and Injury Prevention in Professional Soccer. *Int. J. Sports Physiol. Perform.* 2016, 30, 360–367. [CrossRef]

57. García-García, O.; Serrano-Gómez, V.; Hernández-Mendo, A.; Tapia-Flores, A. Assessment of the in-season changes in mechanical and neuromuscular characteristics in professional soccer players. *J. Sports Med. Phys. Fit.* 2015, 56, 714–723.

58. Maya, J.; Marquez, P.; Peñailillo, L.; Contreras-Ferrat, A.; Deldicque, L.; Zbinden-Foncea, H. Salivary Biomarker Responses to a Final Matches in Women’s Professional Football. *J. Sports Sci. Med.* 2016, 15, 365–371. [PubMed]

59. Moalla, W.; Fessi, M.S.; Farhat, F.; Nouira, S.; Wong, D.P.; Dupont, G. Relationship between daily training load and psychometric status of professional soccer players. *Res. Sports Med.* 2016, 24, 387–394. [CrossRef]

60. Owen, A.L.; Wong del, P.; Dunlop, G.; Groussard, C.; Kebi, W.; Dellall, A.; Morgans, R.; Zouhal, H. High-intensity training and salivary immunoglobulin A responses in professional top-level soccer players: Effect of training intensity. *J. Strength Cond. Res.* 2016, 30, 2460–2469. [CrossRef]

61. Rago, V.; Leo, J.; Pizzuto, F.; Costa, J.; Angeicchio, G.; Raiola, G. Variation of oxidative stress of elite football players during pre-season and in-season. *J. Phys. Educ. Sport* 2016, 16, 326–329.

62. Romagnoli, M.; Sanchis-Gomar, F.; Alis, R.; Risso-Ballester, J.; Bosio, A.; Graziani, R.L.; Rampinini, E. Changes in muscle damage, inflammation, and fatigue-related parameters in young elite soccer players after a match. *J. Sports Med. Phys. Fit.* 2015, 56, 1198–1205.

63. Bacon, C.S.; Mauger, A.R. Prediction of Overuse Injuries in Professional U18-U21 Footballers Using Metrics of Training Distance Players’ Match Differential Ratings of Perceived Exertion: A Preliminary Investigation. *Sports* 2018, 6, 13. [CrossRef]
