frontiers | Research Topics

Highlights in sports science, technology and engineering: 2021/22

Edited by
Billy Sperlich, Valentina Camomilla, Pietro Picerno and Laura Gastaldi

Published in
Frontiers in Sports and Active Living

ISSN 1664-8714
ISBN 978-2-83251-688-1
DOI 10.3389/978-2-83251-688-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Highlights in sports science, technology and engineering: 2021/22

Topic editors
Billy Sperlich — Julius Maximilian University of Würzburg, Germany
Valentina Camomilla — Università degli Studi di Roma Foro Italico, Italy
Pietro Picerno — University of eCampus, Italy
Laura Gastaldi — Polytechnic University of Turin, Italy

Citation
Sperlich, B., Camomilla, V., Picerno, P., Gastaldi, L., eds. (2023). *Highlights in sports science, technology and engineering: 2021/22*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83251-688-1

Table of contents

frontiers | Frontiers in Sports and Active Living

TYPE Editorial
PUBLISHED 01 February 2023
DOI 10.3389/fspor.2022.1117803

Editorial: Highlights in sports science, technology and engineering 2021/22

Peter Düking[1], Pietro Picerno[2], Valentina Camomilla[3], Laura Gastaldi[4] and Billy Sperlich[1]*

[1]Integrative and Experimental Exercise Science, Department of Sport Science, University of Würzburg, Würzburg, Germany, [2]SMART Engineering Solutions & Technologies (SMARTEST) Research Center, Università Telematica "e-Campus", Novedrate, CO, Italy, [3]Department of Movement, Human and Health Sciences, University of Rome Foro Italico, Rome, Italy, [4]Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Torino. Italy

KEYWORDS

wearable, sensor, markerless motion capture, muscle parameters, health monitoring, virtual reality, sport equipment

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Kamiar Aminian,
Swiss Federal Institute of Technology Lausanne, Switzerland

*CORRESPONDENCE
Billy Sperlich
billy.sperlich@uni-wuerzburg.de

SPECIALTY SECTION
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

RECEIVED 06 December 2022
ACCEPTED 13 December 2022
PUBLISHED 01 February 2023

CITATION
Düking P, Picerno P, Camomilla V, Gastaldi L and Sperlich B (2023) Editorial: Highlights in sports science, technology and engineering 2021/22.
Front. Sports Act. Living 4:1117803.
doi: 10.3389/fspor.2022.1117803

Editorial on the Research Topic
Highlights in Sports Science, Technology and Engineering: 2021/22

Introduction

Sports Technology and Engineering is a trending sub-category translating the innovative use of technology and engineering design into practice and encompasses many disciplines of sports science (e.g., sports medicine, prevention, rehabilitation, athletic development, etc.). Research in this field often aims to develop materials, sensors, algorithms, or full pieces of equipment to maintain and improve certain dimension of health, lifestyle and/or performance in different populations (e.g., able-bodied and disabled, sedentary, diseased, fitness oriented, competitive, or athletic). Such sports technologies are either developed from a research perspective to increase the understanding of operating principles and adaptation mechanisms (e.g., employing sensors and algorithms to monitor a variety of parameters in different settings) or from an applied perspective to provide technologies which optimize training, competition, or lifestyle activities.

Research around sports technology is growing rapidly and even outpaces other areas of interest in the field of sports science. **Figure 1** illustrates the number of yearly publications listed in the Pubmed database. In 2000, 168 articles were published in Pubmed for the search-term "Sports Technology" and for example "Endurance Training" or "Strength training" hit 299 and 437 numbers of publications, respectively. While the numbers of all these search terms increased since 2011 year-by-year annual publications for "Sports Technology" ($n = 702$ in 2011) have surpassed those for "Endurance Training" ($n = 678$ in 2011). Since 2020, the annual number of

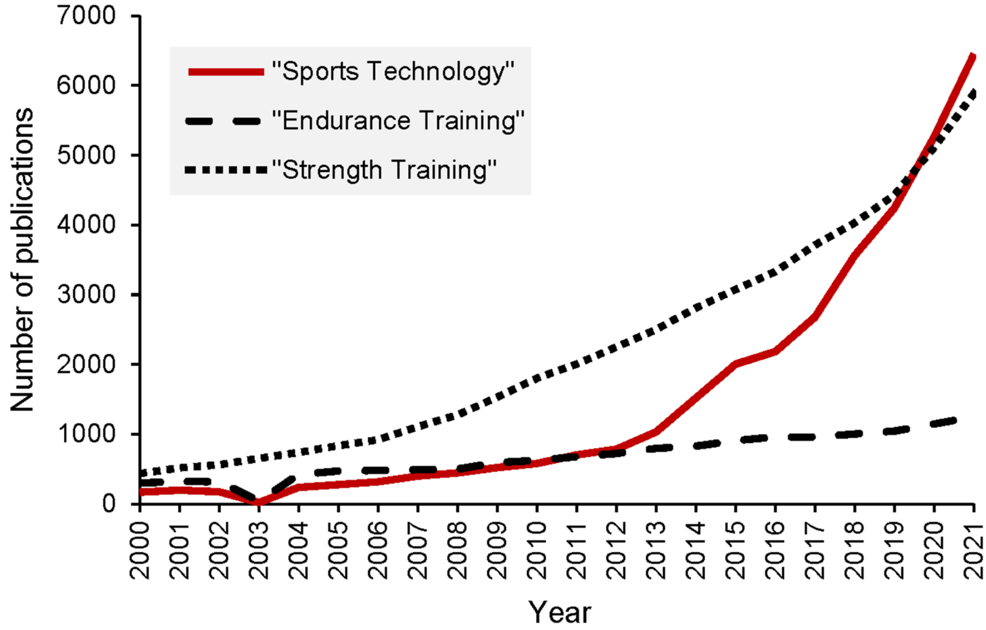

FIGURE 1
Number of publications for “Sports Technology”, “Endurance Training” and “Strength Training” in Pubmed.gov as of 07.11.2022.

publications for “Sports Technology” ($n = 5{,}264$ in 2020) is also higher than the yearly number of publications for “Strength Training” ($n = 5{,}114$ in 2020).

The variety of sports technology research is reflected in the present research topic including 10 articles from 54 authors out of which six are original research articles, one perspective article, one research report, one case report and one methodological article.

The aim of this research topic was to gather scientific contributions from the broad variety of research performed across the Sports Science, Technology and Engineering section in the years 2020–2021 to highlight the current main areas of interest, as well as emerging applications and trends. Here, the potential for sports technology and engineering research describes the potential to enhance “in the field” motion analysis, investigate mechanisms of adaptation, and enhance health monitoring and training in different populations and settings.

Two papers tested markerless motion capture for “in the field” sport motion analysis, through the assessment of upper and lower limbs kinematics. In this regard, Lahka and colleagues aimed at evaluating the concurrent validity of a bidimensional markerless motion capture system in assessing upper limbs kinematics of elite boxers while performing some typical in-ring boxing maneuvers Lahkar et al. while Pinheiro et al. tested the open-source OpenPose software for penalty kick analysis in elite football players from TV footage as compared to an observational analysis.

Two papers investigated technologies to assess muscle parameters during exercise: Puce et al. investigated the correlation between spectral parameters obtained from surface electromyography and variations of kinematic data and mechanical fatigue in elite swimmers. McPhail et al. and co-workers evaluated the within-session reliability of some force-related performance parameters during a novel unilateral isometric hex bar pull as performed by male and female elite freeski athletes on a force plate at the maximal voluntary contraction, while providing sex- and level specific reference values.

Health monitoring was the focus of three contributions: Ausland et al. proposed the use of a new mobile, long-term electrocardiogram (ECG) monitoring patch to assess automatic arrhythmia detection during endurance training in elite athletes. Bender et al. reported a device capable of automatically analyzing urine specific gravity (i.e., an index of hydration status of individuals) in real time. Finally, Fraysse et al. report physical activity cut-points for wrist-worn technologies in elderly populations (>70 years of age). Noteworthy, the authors investigate whether wear-side (i.e., dominant vs. non-dominant wrist) affects accuracy and conclude that wearing the technology on the dominant wrist might deliver more accurate data especially when individuals are active in low intensity zones. Schelling and co-workers presented a methodological framework for the design of a decision support systems for scheduling trips and training sessions in professional team sports.

Sports technology research contributes also to aspects of training, e.g., by augmenting the training environment using virtual reality, or adapting sport equipment to the specific athlete. In particular, the case report by Severin et al. and others compared the effects of adjusting seat and backrest angle on performance of an elite paralympic rower. In their perspective paper, McIlroy et al. shared their opinion concerning the use of virtual reality to train cyclists. The authors summarize strengths, weaknesses, as well as opportunities and threats of virtual online training platforms, with the attempt to enhance awareness of various aspects of virtual training technology and online cycling.

Throughout the reported research ideas and research outcomes, it remains clear that sports technology does not fulfill a purpose on its own, but it is a means to an end for sport scientists, applied practitioners, coaches and athletes. To increase our understanding of working principles and mechanisms, sports technology needs to be sound and reliable and must be flawlessly applied. In this regard, the field needs educated researchers and practitioners who understand how to employ and interact effectively and efficiently with technologies of different forms (1, 2). This interaction includes (i) having an evidence-based course of action on whether and when to apply technologies for a certain problem, (ii) the selection of appropriate technologies for a given purpose, population and setting, (iii) the appropriate management, handling and application of the technology and (iv) understanding of how to analyze and contextualize data correctly. Given the increasing availability of sports technologies, it seems necessary to constantly and critically evaluate new technologies and, at the same time, educate stakeholders in this area.

Author contributions

All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

References

1. Robertson PS. Man & machine: adaptive tools for the contemporary performance analyst. *J Sports Sci*. (2020) 38(18):2118–26.

2. Hughes GT, Camomilla V, Vanwanseele B, Harrison AJ, Fong DT, Bradshaw EJ. Novel technology in sports biomechanics: some words of caution. *Sports Biomech*. (2021):1–9. doi: 10.1080/14763141.2020.1869453

frontiers in Sports and Active Living

ORIGINAL RESEARCH
published: 15 January 2021
doi: 10.3389/fspor.2020.579278

Check for updates

Physical Activity Intensity Cut-Points for Wrist-Worn GENEActiv in Older Adults

François Fraysse[1], Dannielle Post[1], Roger Eston[1], Daiki Kasai[1], Alex V. Rowlands[2,3] and Gaynor Parfitt[1]*

[1] Alliance for Research in Exercise, Nutrition and Activity (ARENA), University of South Australia, Adelaide, SA, Australia, [2] National Institute for Health Research Leicester Biomedical Research Centre, Leicester, United Kingdom, [3] Diabetes Research Centre, University of Leicester, Leicester, United Kingdom

Purpose: This study aims to (1) establish GENEActiv intensity cutpoints in older adults and (2) compare the classification accuracy between dominant (D) or non-dominant (ND) wrist, using both laboratory and free-living data.

Methods: Thirty-one older adults participated in the study. They wore a GENEActiv Original on each wrist and performed nine activities of daily living. A portable gas analyzer was used to measure energy expenditure for each task. Testing was performed on two occasions separated by at least 8 days. Some of the same participants ($n = 13$) also wore one device on each wrist during 3 days of free-living. Receiver operating characteristic analysis was performed to establish the optimal cutpoints.

Results: For sedentary time, both dominant and non-dominant wrist had excellent classification accuracy (sensitivity 0.99 and 0.97, respectively; specificity 0.91 and 0.86, respectively). For Moderate to Vigorous Physical Activity (MVPA), the non-dominant wrist device had better accuracy (ND sensitivity: 0.90, specificity 0.79; D sensitivity: 0.90, specificity 0.64). The corresponding cutpoints for sedentary-to-light were 255 and 375 g · min (epoch independent: 42.5 and 62.5 mg), and those for the light-to-moderate were 588 and 555 g · min (epoch-independent: 98.0 and 92.5 mg) for the non-dominant and dominant wrist, respectively. For free-living data, the dominant wrist device resulted in significantly more sedentary time and significantly less light and MVPA time compared to the non-dominant wrist.

Keywords: accelerometer, dominant, non-dominant, sedentary, light, moderate

OPEN ACCESS

Edited by:
Emma Fortune, Mayo Clinic, United States

Reviewed by:
Jairo H. Migueles, University of Granada, Spain
Swati Chopra, University of Leeds, United Kingdom

****Correspondence:***
François Fraysse
francois.fraysse@unisa.edu.au

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: *02 July 2020*
Accepted: *07 December 2020*
Published: *15 January 2021*

Citation:
Fraysse F, Post D, Eston R, Kasai D, Rowlands AV and Parfitt G (2021) Physical Activity Intensity Cut-Points for Wrist-Worn GENEActiv in Older Adults. Front. Sports Act. Living 2:579278. doi: 10.3389/fspor.2020.579278

INTRODUCTION

Population-level measurement of physical activity (PA) and sedentary behaviors (SBs) is important for a number of reasons, including investigating relations with health outcomes (Osborn et al., 2018), quantifying the effect of PA interventions (Mitchell et al., 2019), and establishing secular trends in PA behaviors (Fraysse et al., 2019). Device-measured PA is most commonly carried out using accelerometers worn on the hip, wrist, or thigh. In recent years, most research-grade accelerometers have allowed direct access to raw acceleration data. This allows control over the whole data processing method, resulting in better transparency and reproducibility. Typically, to facilitate classification of activity intensity, acceleration magnitude is collapsed (summed or averaged) into epochs ranging from 1 to 60 s, and thresholds are applied to classify each waking wear epoch as sedentary, light PA (LPA), moderate PA (MPA), or vigorous PA (VPA). These thresholds, also commonly called cutpoints, are established in calibration studies (Evenson et al., 2008; Sasaki et al., 2011; Hildebrand et al., 2017) where acceleration data are recorded concurrently with energy expenditure (EE), obtained from measurement of VO_2/CO_2 using a metabolic cart (Bassett et al., 2012). Some studies have also used other means of estimating EE, such as direct observation, video recordings, or use of the PA compendium. EE is expressed relative to the standard unit of resting metabolism [metabolic equivalent (MET)], and typically, 1.5, 3, and 6 METs are considered the thresholds between sedentary, LPA, MPA, and VPA (Copeland and Esliger, 2009) (although some studies have used 4 and 7 METs for the latter two) (Gorman et al., 2014; Whitcher and Papadopoulos, 2014; Evenson et al., 2015).

These acceleration cutpoints are age-specific and wear-site–specific. The relation between EE and bodily movement changes with age. Physical fitness decreases with age, and as a result, performing the same activity requires higher EE. Because intensity cutpoints are based on fixed values of EE, the corresponding acceleration threshold (reflecting body movement) will tend to decrease the older the target population is. Previous studies have emphasized the need for activity cutpoints specific to older adults (Rejeski et al., 2015; Mankowski et al., 2017). Moreover, while different accelerometer brands generally show excellent agreement in terms of activity classification, they can differ in terms of raw acceleration output (Rowlands et al., 2017). GENEActiv and Axivity devices both use the ADXL345 accelerometer, and their raw acceleration outputs are practically identical, so that it is sensible to use the same cutpoints for both devices. It is not clear whether the same cutpoints can be applied for brands using different accelerometers; for instance, we know that ActiGraph GT9X exhibits overall lower accelerations than GENEActiv and Axivity.

Most cutpoint studies have focused on children or adults, including studies using raw acceleration data from the GENEActiv and ActiGraph GT3x+ devices (Esliger et al., 2011; Phillips et al., 2013; Schaefer et al., 2014; Hildebrand et al., 2017). ActiGraph devices prior to the GT3X provide results not in terms of raw acceleration, but in so-called *counts*, which are filtered signals, with the filtering parameters kept undisclosed by ActiGraph. For the hip-worn ActiGraph GT3X (ActiGraph, Pensacola, FL) accelerometer, ActiGraph *count* cutpoints have been compared in older (66.6 ± 2.9 years) and younger (21 ± 2.5 years) adults, with 824 and 2,207 counts $\cdot$ min^{-1}, respectively, associated with moderate-intensity (3 METs) activity (Whitcher and Papadopoulos, 2014). However, few studies have reported cutpoints for older adults (Gorman et al., 2014; Whitcher and Papadopoulos, 2014; Evenson et al., 2015).

For GENEActiv devices, studies by Duncan et al. (2019) and Sanders et al. (2019) have established acceleration cutpoints for older adults (55–77 and 60–86 years old, respectively). Duncan et al. tested the effect of wear site on activity classification, with devices worn on both wrists, waist, and ankle. The activities used were focused on the moderate activity level, and most of the moderate-intensity activities were walking activities, at different speeds. Sanders et al. investigated a non-dominant-wrist-worn GENEActiv and a waist-worn ActiGraph and produced two sets of cutpoints, one optimizing the overall classification accuracy and the other optimizing sedentary sensitivity and MVPA specificity. This study did not compare dominant vs. non-dominant wrist, and we know that acceleration, and therefore cutpoints, can be different between wrists for the same activity (Esliger et al., 2011; Phillips et al., 2013; Duncan et al., 2019). Moreover, both studies placed the emphasis on sedentary and moderate-vigorous intensities, at the expense of light intensity.

Older adults tend to have a lower exercise capacity, spend more time sedentary, and rarely engage in VPA relative to younger adults (Matthews et al., 2008; Troiano et al., 2008; Jefferis et al., 2019). For this reason, an increase in PA could result in improved health. In particular, it has been shown that even LPA is associated with a lower risk of death in the elderly (Ekelund et al., 2019; Klenk and Kerse, 2019). However, most accelerometer-based PA studies tend to focus on the ends of the PA spectrum, that is, sedentary and MVPA. Our study was designed with a focus on light-to-moderate activities, with a goal to achieving better discrimination between sedentary, LPA, and MPA in older adults.

Wrist-worn accelerometers are reported to be more acceptable and better tolerated by children (Fairclough et al., 2016), adolescents (Scott et al., 2017), and adults (Montoye et al., 2020), compared to a hip-worn device, although we do not know whether this is true for older adults for which daily activity patterns and usual clothes are usually very different from younger populations. The identification of PA intensity cutpoints, specific to older adults for wrist-worn accelerometers, and with a stronger focus on the light activity intensities, is warranted. A secondary purpose was to investigate the effect of accelerometer placement [dominant (D) or non-dominant (ND) wrist] on classification accuracy and test–retest reliability, with an emphasis on typical activities in LPA domain—such as grocery shopping, sweeping, washing dishes, gardening, and walking—in order to complement the results of Duncan et al. (2019). We also applied the cutpoints established in this study to a sample of free-living data from the same participants, in order to determine whether there were any differences between devices worn on each wrist.

METHODS

An opportunistic sample of 36 healthy, South Australia community-dwelling older adults was recruited for this study. Inclusion criteria were as follows: older than 70 years, fluent in English, and capable of undertaking general activities of daily living unassisted, such as walking and carrying shopping. Participants' characteristics are presented in **Table 1**. Twenty-four of the 36 participants were classified as overweight or obese, with a body mass index >25 kg/m^2. The protocol was approved by the University of South Australia Human Research Ethics Committee. Participants provided written, informed consent.

TABLE 1 | Participants' characteristics.

	Sample size	Age (y)	Height (cm)	Weight (kg)	Body mass index (kg/m^2)
Female	18	76 (4)	158 (6)	66 (11)	26.2 (4.1)
Male	18	78 (6)	174 (5)	86 (11)	28.3 (3.6)

Standard deviations are presented in brackets.

The experiment consisted of two laboratory visits at least a week apart, for test–retest data. The same protocol was repeated for the two visits.

Laboratory Sessions Protocol

Participants were fitted with one GENEActiv on each wrist (GENEActiv Original, Activinsights, UK). The devices were configured to record data at 100 Hz. Breath-by-breath online gas analysis was conducted via MetaMax 3B (Cortex Biophysik GmbH, Leipzig, Germany) with a face mask (Hans Rudolph Inc., Shawnee, KS, USA). Volume and gas calibration were conducted in accordance with the manufacturer's guidelines prior to each session. Heart rate was measured continuously using a wireless chest-strap telemetry system (RS400; Polar Electro Oy, Espoo, Finland).

They were then asked to perform a series of activities typical of activities of daily living for older adults. These were, in order, as follows:

1. light gardening (digging and removing objects from a sandpit) for 4 min,
2. sweeping the floor with a broom while standing for 4 min,
3. seated reading for 5 min,
4. walking overground at a self-paced comfortable speed for 4 min (a "comfortable, everyday walking pace"),
5. lying in a lateral recumbent position for 5 min,
6. washing and drying dishes while standing for 4 min,
7. walking overground at a self-paced brisk speed for 4 min (a "brisk pace"),
8. watching TV seated for 5 min, and
9. unpacking groceries while standing for 4 min.

There was a 1-min break between each activity. The researcher wrote down the start and end time of each activity with a 1-s resolution.

Free-Living Protocol

Between the two laboratory sessions, some participants ($n = 13$) wore a GENEActiv monitor on each wrist for 3 days continuously. They were instructed to keep the devices on at all times as much as feasible, including sleep, and to remove them only for prolonged water immersion. They were also asked to fill in a paper log every day with their bed time and get-up time. This free-living data were processed following the same method as the main study, following which sleep was isolated using self-report logs filled by the participants. Waking wear time was then classified using the cutpoints established in this study. t-Tests were performed for the daily average time spent in each of the three intensities in order to check for significant differences between wrists.

Data Processing

All processing was done in MATLAB (2018b, the MathWorks, Inc.), and the programs are available on request. GENEActiv data were downloaded and low-pass filtered with a cutoff frequency of 20 Hz. The 100-Hz data were collapsed in 5-s epochs by computing the signal vector magnitude (SVM), subtracting gravity, and summing magnitudes over a 5-s window:

$$\mathrm{SVM_{gs}} = \sum_{5s} \left| \sqrt{a_X^2 + a_Y^2 + a_Z^2} - g \right|$$

For each activity, the mean, median and standard deviation of the SVM over the central 3 min of activity were computed.

To determine the MET values for each activity, 30-s measured oxygen consumption (mL · kg^{-1} · min^{-1}) were averaged for the last 2 min of each activity and divided by 2.8 mL · kg^{-1} · min^{-1} [resting metabolic rate (1 MET) for older adults (Kwan et al., 2004)].

MET data were averaged for each activity. Test–retest reliability analyses were conducted for data collected for each activity across the two testing sessions. The activities were then coded into three categories: sedentary (<1.5 METs), light (1.5–2.99 METs), moderate (3.0–5.99 METs), and vigorous (≥6 METs) (Copeland and Esliger, 2009). The focus of the study was on the light–moderate region of PA, and as such, very few instances of vigorous activity were observed. For the purpose of establishing cutpoints, we therefore opted to group moderate and vigorous activities into a single moderate-to-vigorous (MVPA) level corresponding to METs ≥3.

ROC Analysis

The goal of the receiver operating characteristic (ROC) curve analysis was to find the cutpoints of accelerometer SVM that most accurately classified each of the three considered activity levels. We followed the same method as Esliger et al. (2011) and Phillips et al. (2013).

The two SVM cutpoints between sedentary and LPA, and between LPA and MVPA, were varied from 1 to 100 g · s (for sedentary to light) and 1–1,000 g · s (for LPA to MVPA) in increments of 1 g · s.

The true- and false-positive, and true- and false-negatives, were then computed, with the definition for true positives as follows:

- Sedentary: if the accelerometer SVM classified the activity as sedentary, and the METs for the activity were <1.5;
- MVPA: if the accelerometer SVM classified the activity as MVPA, and the METs for the activity were at or >3.0.
- Consequently, LPA classification is the one that is above sedentary and below MVPA. A true-positive for LPA classification is therefore as follows: MET values are between 1.5 and 3.0, and acceleration magnitude between the sedentary-to-light and light-to-MVPA thresholds.

Sensitivity and specificity were then computed, and the ROC curves for each of the two cutpoints were created. The optimal cutpoints were defined as the SVM threshold values that maximized the product of sensitivity and specificity. In order to allow comparison with other studies, these cutpoints are presented in two ways:

- scaled to a 60-s epochs equivalent by multiplying the values by 12 (5 s × 12 = 60 s) to allow direct comparison with the cutpoints (g · min) of Esliger et al.;
- averaged over the 5-s epoch length, which makes the resulting cutpoints independent of epoch length.

RESULTS

Acceleration SVM as a Function of Activity Intensity

Of the 36 initial participants, five were excluded from analysis: one did not complete all tasks, two withdrew before completion, one had issues with GENEActiv data extraction (identical data for dominant and non-dominant wrists, likely due to operator error during configuration or extraction), and one had mismatches between GENEActiv and VO_2 data time stamps. Thus, 31 participants (14 female) were included in the analysis.

The MET values and acceleration SVM for each activity, averaged across participants, are presented in **Table 2**. Standard deviations and 95% confidence intervals are also presented. There were no significant differences ($p > 0.05$) for either MET values or acceleration between the two time points; we therefore decided to merge the two time points for the subsequent ROC analysis in order to obtain more robust cutpoint estimates. Generally speaking, sitting, lying recumbent, and watching TV were sedentary activities (MET <1.5); light gardening and doing dishes were LPA (1.5≤ MET <3.0); and walking and unpacking groceries were MPA (3.0≤ MET <6.0) with a few participants performing brisk walking as VPA (MET ≥6.0). Paired t-tests resulted in significant differences between acceleration magnitudes (SVM) for the dominant and non-dominant wrists for gardening ($p < 0.001$), sweeping ($p < 0.001$), and doing dishes ($p < 0.001$) for both timepoints, with the dominant wrist exhibiting larger SVM for all these. For all other activities, there were no significant differences at either timepoint.

Figure 1 presents the SVM vs. METs for each participant and activity. As can be seen, acceleration SVM increased reasonably linearly with METs. Pearson correlation was $r^2 = 0.650$ and $r^2 = 0.628$ for dominant and non-dominant wrists, respectively. Dominant wrist SVM was overall higher than non-dominant; however, the difference was only statistically significant for LPA ($1.5 \leq$ METs < 3.0, $p < 0.001$) and MVPA (METs ≥ 3.0, $p < 0.05$).

In order to check whether participants had reached steady state for each activity when recording started, we calculated the MET values during the 1-min period preceding the 2 min recording for each activity. As shown in **Table 2B**, those MET values indicate steady state was reached for each activity. Moreover, data sampled 1 min before each sedentary activity started support the fact that

TABLE 2A | MET and SVM for the nine activities and both timepoints, average (SD) and 95% CI across 31 participants.

		Gard		Sweep		Seat		WalkS		Lie		Dishes		WalkF		TV		Grocer	
		Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI	Mean (SD)	95% CI
T1	MET	2.79 (0.74)	0.25	3.86 (0.86)	0.29	1.41 (0.24)	0.08	3.68 (0.69)	0.23	1.26 (0.19)	0.07	2.48 (0.40)	0.13	4.94 (1.05)	0.35	1.29 (0.20)	0.07	3.55 (0.61)	0.20
	ND SVM	**26.9 (13.6)**	4.7	**105.9 (33.3)**	11.5	12.4 (4.1)	1.4	76.8 (24.4)	8.5	11.9 (12.5)	4.3	**56.8 (16.4)**	5.7	130.4 (53.3)	18.5	9.3 (3.6)	1.2	55.4 (11.9)	4.1
	D SVM	**80.7 (26.0) <0.001**	9.0	**135.9 (41.1) <0.001**	14.2	14.0 (6.5)	2.3	76.2 (24.7)	8.6	13.8 (15.3)	5.3	**76.5 (20.2) <0.001**	7.0	135.7 (55.4)	19.2	9.6 (5.1)	1.8	60.8 (15.5)	5.4
T2	MET	2.75 (0.58)	0.20	3.85 (0.81)	0.27	1.41 (0.25)	0.08	3.47 (0.91)	0.31	1.23 (0.31)	0.10	2.47 (0.39)	0.13	4.82 (1.33)	0.45	1.27 (0.22)	0.07	3.49 (0.78)	0.26
	ND SVM	**27.7 (14.4)**	5.0	**96.8 (31.6)**	10.9	13.9 (4.9)	1.7	70.2 (25.4)	8.8	11.5 (10.8)	3.8	**55.3 (14.6)**	5.1	131.7 (56.8)	19.7	8.8 (5.1)	1.8	54.6 (11.2)	3.9
	D SVM	**81.1 (29.9) <0.001**	10.4	**127.8 (36.2) <0.001**	12.5	14.5 (4.8)	1.7	71.1 (27.7)	9.6	11.8 (13.0)	4.5	**74.2 (20.4) <0.001**	7.1	135.4 (60.7)	21.0	7.1 (3.3)	1.1	59.9 (13.6)	4.7

Bold SVM values indicate significant differences between dominant (D) and non-dominant wrists (ND). T1, timepoint 1; T2, timepoint 2; gard, light gardening; sweep, sweeping the floor; seat, seated reading; walkS, walking at self-paced comfortable speed; lie, lying in a lateral recumbent position; dishes, washing and drying dishes; walkF, walking at self-paced brisk speed; TV, watching TV seated; grocer, unpacking groceries while standing. Green shading indicates sedentary activities; yellow shading indicates light activities; orange shading indicates moderate activity.

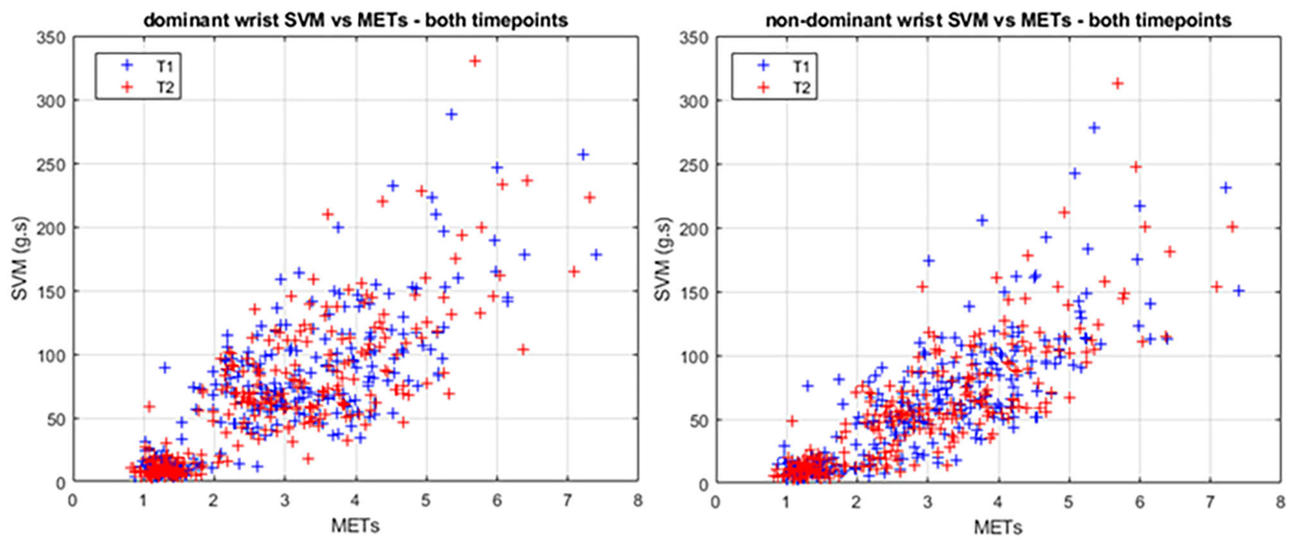

FIGURE 1 | SVM vs. METs for the dominant **(left)** and non-dominant **(right)** wrists. Each point represents one activity and one participant. Timepoint 1 and 2 are shown in blue and red, respectively.

TABLE 2B | MET values for the nine activities and both timepoints, average (SD) and 95% CI across 31 participants, compared to the MET values for the same activities, in the minute prior to the recording period.

		Gard	Sweep	Seat	WalkS	Lie	Dishes	WalkF	TV	Grocer
		Mean (*SD*)	**Mean (*SD*)**	**Mean (*SD*)**	**Mean (*SD*)**	**Mean (*SD*)**	**Mean (*SD*)**	**Mean (*SD*)**	**Mean (*SD*)**	**Mean (*SD*)**
T1	MET	2.79 (0.74)	3.86 (0.86)	1.41 (0.24)	3.68 (0.69)	1.26 (0.19)	2.48 (0.40)	4.94 (1.05)	1.29 (0.20)	3.55 (0.61)
T2	MET	2.75 (0.58)	3.85 (0.81)	1.41 (0.25)	3.47 (0.91)	1.23 (0.31)	2.47 (0.39)	4.82 (1.33)	1.27 (0.22)	3.49 (0.78)
Minute before recording	MET	2.78 (0.76)	3.78 (0.89)	1.49 (0.34)	3.59 (1.09)	1.42 (0.33)	2.19 (0.44)	4.65 (1.07)	1.40 (0.33)	3.34 (0.75)

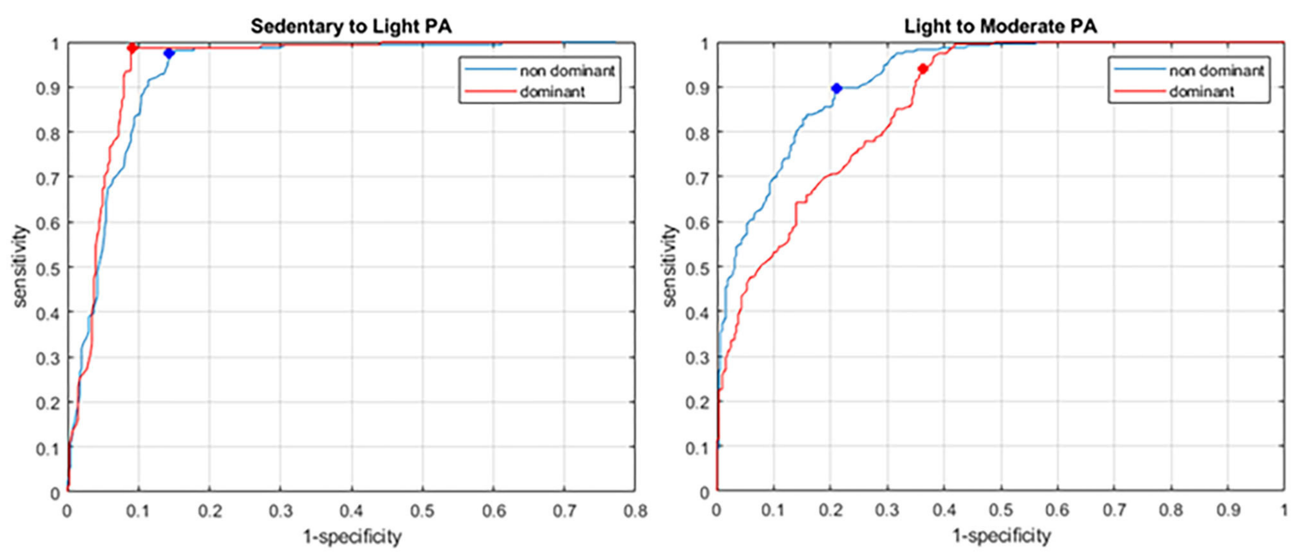

FIGURE 2 | ROC for the sedentary to LPA **(left)** and LPA to MPA **(right)** cut points. Dominant and non-dominant wrists are shown in red and blue, respectively. The dots show the selected cut-points maximizing the product of sensitivity and 1-specificity.

participants had recovered, with MET values of 1.49 (*SD*, 0.34), 1.42 (*SD*, 0.33), and 1.40 (*SD*, 0.33) prior to the three sedentary activities.

Intensity Cutpoints

Figure 2 shows the results of the ROC analysis for the sedentary to LPA, and the light to MPA intensity cutpoints. For both wrists,

TABLE 3 | Intensity cutpoints and associated sensitivity and specificity for sedentary to light PA, and light to moderate PA, for the present study and the studies of Esliger et al., Duncan et al., and Sanders et al. cutpoints from both studies were converted to 60-s epoch equivalents at 100 Hz.

Sedentary to light	Cutpoint (g · min) 60-s epochs	Cutpoint (mg) epoch independent	Sensitivity	Specificity
ND—this study	**255**	**42.5**	**0.97**	**0.86**
ND—Esliger	271	45.2	0.97	0.95
ND—Duncan		19.2	0.74	0.81
ND—Sanders		57	0.43	0.99
D—this study	**375**	**62.5**	**0.99**	**0.91**
D—Esliger	483	80.5	0.99	0.96
D—Duncan		20.2	0.88	0.84
Light to moderate				
ND—this study	**588**	**98.0**	**0.90**	**0.79**
ND—Esliger	806	134.3	0.95	0.72
ND—Duncan		89.7	0.67	0.81
ND—Sanders		104	0.81	0.65
D—this study	**555**	**92.5**	**0.94**	**0.64**
D—Esliger	550	91.7	1	0.56
D—Duncan		113.8	0.65	0.8

TABLE 4 | Average daily time spent in each activity intensity for the 3-day free-living data, calculated for the dominant (D) and non-dominant (ND) wrists, using the respective cutpoints.

		n	Mean (min/day)	Standard deviation (min/day)
Sedentary	D	13	630*	75
	ND	13	538*	90
Light PA	D	13	103*	25
	ND	13	209*	57
MVPA	D	13	141	50
	ND	13	134	44

Asterisks indicate significant differences (p < 0.05) between dominant and non-dominant wrists.

the sedentary to light cutpoint presented excellent sensitivity (0.987 and 0.974 for D and ND, respectively) and relatively lower, although still very acceptable, specificity (0.908 and 0.856, respectively). Corresponding intensity cutpoints for 60-s epoch data were 375 and 255 g · min for the dominant and non-dominant wrists, respectively. The light to MPA cutpoints displayed good sensitivity (0.940 and 0.898 for D and ND, respectively) but lower specificity (0.638 and 0.789, respectively) than the sedentary to LPA cutpoint. Corresponding 60-s intensity cutpoints were 555 and 588 g · min for D and ND wrists, respectively. **Table 3** summarizes these findings and presents the results from Esliger et al. (2011), Duncan et al. (2019), and Sanders et al. (2019) for comparison. Age range for our participants was 70–91 years, and that of Esliger et al. was 40–63 years.

Free-Living Data

In order to assess the practical differences in intensity estimates between wrists, we used data from a 3-day free-living sample collected on 13 participants that were also part of the main study. Average daily time spent in each intensity level (sedentary, LPA, and MVPA) is presented in **Table 4**.

The dominant wrist and associated cutpoints resulted in significantly more ($p < 0.01$) sedentary time (+92 min/day average), and conversely, significantly less LPA time (−106 min/day average). There was no significant difference ($p = 0.42$) for MVPA time.

DISCUSSION

This study fills a gap in the literature regarding intensity cutpoints for older adults when using the GENEActiv accelerometer, with an emphasis on the LPA and MPA intensity. Additionally, it examines the benefits of wearing the accelerometer on the dominant or non-dominant wrist in terms of activity intensity classification accuracy.

There are a number of published cutpoints for the GENEActiv on adult populations (Esliger et al., 2011; Hildebrand et al., 2017; Duncan et al., 2019; Sanders et al., 2019). However, Hildebrand et al. (2017) used a different method of data processing, in that they rounded the negative acceleration values to zero rather than taking the absolute value. Therefore, our results can only be compared with those of Esliger et al. (2011). In this regard, our cutpoints are generally lower than those of Esliger et al. (2011) for both sedentary to LPA and LPA to MPA, with the exception of the LPA to MPA cutpoint for the dominant wrist. In comparison to Esliger et al. (2011), the corresponding intensity cutpoints for sedentary-to-light were 6% lower for the non-dominant (255 vs. 271 g · min) and 22% lower for the dominant wrist (375 vs. 483 g · min), and the light-to-moderate cutpoint was 27% lower on the non-dominant wrist (588 vs. 806 g · min), but similar for the dominant wrist (555 vs. 550 g · min). The population of Esliger et al. (2011) was younger than ours, so these findings are in line with the fact that cutpoints become lower for the same intensity with increasing age. This is most notable when comparing children and adult cutpoints, and our study indicates that this trend continues between adults and elderly, in agreement with the findings of Whitcher and Papadopoulos (2014).

Usually, accelerometers are worn on the non-dominant wrist, as it is believed to be a better estimate of overall body movement and therefore activity intensity. In particular, sedentary activities involving mostly dominant arm movement, such as writing, eating, or smoking, could result in the activity being erroneously classified at a higher intensity than it really is. Our ROC results indicate that classification accuracy is comparable for sedentary to LPA, with very high sensitivity and specificity for both dominant and non-dominant wrists, and only slightly lower specificity for the dominant wrist. Note, however, that in this case the specificity is >0.9, which indicates a very good accuracy. Esliger et al. (2011) report similar sensitivities but higher specificities; this may be due to the fact that our study has a continuous range of METs and acceleration SVM (**Figure 2**), whereas the activities used in the study of Esliger et al. (2011) resulted in a more clustered distribution, which may have contributed to higher specificity (fewer false-positives). Finally, posture (standing vs. sitting) was not considered in the reference method; therefore, the sedentary-to-light cutpoint may be misclassifying light activities as sedentary, if the person is standing up but not moving his/her arms much. Additionally, we used a base value for resting METs for all participants instead of measuring each participant's base resting METs individually, which may have affected the final cutpoint values at all levels.

For the LPA to MPA cutpoint, classification accuracy drops for both wrists; in particular, the specificity is much lower, at 0.638 and 0.789 for dominant and non-dominant wrists, respectively. This was also observed in the results of Esliger et al. (2011). Lower specificity indicates more false-positives, i.e., the accelerometer classified the activity as MPA (or VPA), whereas the activity was actually lower than MPA according to MET values. One reason for this may be that some LPA-level activities involve large or rapid arm movement, while overall the body is standing still: washing dishes and light gardening would be examples from our study. This would also explain the fact that the non-dominant wrist device showed better specificity (0.789) than the dominant one (0.638); as mentioned above, some LPAs primarily involve dominant arm movement; therefore, the risk of misclassifying these as MPA is higher with a dominant wrist device, resulting in more false-positives. **Table 2** confirms this: our LPA-level activities showed significantly higher accelerations of the dominant wrist compared to the non-dominant.

Cutpoints established by Duncan et al. (2019) and Sanders et al. (2019) are more difficult to compare directly to ours, because these studies round negative magnitudes to 0 rather than take their absolute value, which causes overall acceleration magnitudes to be lower and therefore cutpoints to be lower. Nevertheless, it can be seen that the sedentary-to-light cutpoints from Duncan et al. are lower than ours by a factor of 2 and 3 for the non-dominant and dominant wrists, respectively (and Duncan having very similar cutpoints for both wrists). The Sanders cutpoints reported in **Table 2** are those established using their so-called "Se" method, which aims at maximizing detection of sedentary and MVPA levels, at the expense of LPA. The Sanders cutpoint for sedentary-to-light, at the non-dominant wrist, is 57 mg, approximately 27% higher than ours, which is consistent with their methodology of minimizing false-negatives for sedentary time (and therefore results in a higher cutpoint value).

The target population for this study was older adults, who generally have relatively high levels of sedentary time (Matthews et al., 2008). In this population, a shift in PA from sedentary to light is expected to have positive health outcomes (Stamatakis et al., 2018; Jefferis et al., 2019). In that regard, having a better discrimination between sedentary and LPA should be seen as positive as it allows a finer detection of improvements in PA behaviors. Note that the UK Biobank study used the dominant wrist (Doherty et al., 2017), whereas the majority of other large-scale studies have used the non-dominant wrist [e.g., NHANES (Matthews et al., 2008), LSAC Checkpoint (Fraysse et al., 2019)].

Our study seems to indicate that a dominant-wrist-worn device achieves a better discrimination between sedentary and LPA intensities. However, the difference between dominant and non-dominant wrists remains small, and the present study is by design limited to a small number of laboratory-based activities. Of note, the sedentary-level activities we tested involve about equal movement of the right and left hands (lying, seated reading, and watching TV), whereas our light-intensity activities (gardening and washing dishes) involve mostly the dominant hand. This could have caused the difference in classification we see here between the two devices.

Similarly, for the light-to-moderate threshold, walking is not associated with large differences between dominant and non-dominant accelerations, whereas activities such as gardening are. The proportion of activities that are predominantly performed with the dominant hand will determine the magnitude of this effect. Comparing acceleration magnitudes from both wrists using free-living data has provided more insight into this issue with a younger population (Rowlands et al., 2019), but remains to be done with data for an older population. Cross-validation of the cutpoints using an independent sample would provide better ecological validity to these findings. Additionally, reallocation of LPA into MPA or VPA also has positive effects, and in this case, the non-dominant wrist provides better discrimination. Overall, it is still unclear which of the dominant or non-dominant wrist provides better estimates of activity intensities. Finally, a recent study indicates that temporal patterns of PA are associated with health outcomes in older adults (Li et al., 2019); in this regard, obtaining cutpoints that allow good separation of PA levels is even more critical.

When the dominant and non-dominant cutpoints were applied to free-living data, results showed that the dominant wrist resulted in significantly more time spent sedentary, at the expense of time spent in LPA (**Table 3**). This result is expected considering the higher sedentary-to-light cutpoint for the dominant wrist. It is, however, unclear which wrist is a better estimate of sedentary and LPA intensities. One possibility is that the laboratory study exacerbated differences between wrists that are not as large in free-living. A recent study by Migueles et al. (2019) found indeed that the dominant wrist exhibited overall larger acceleration magnitudes; however, the resulting difference they found in cutpoints (50 and 45 mg for dominant and non-dominant, respectively) is smaller than the one we found here, suggesting the possibility that our laboratory activities favored motion of the dominant arm. It is also worth noting the large difference in sedentary-to-light cutpoints found by Sanders et al. (2019) when optimizing for either overall best intensity discrimination (20 mg) compared to optimizing for sedentary detection (57 mg).

On a side note, the MET values for the comfortable and brisk walks were significantly different ($p < 0.001$), and both were in the range of MPA (>3.0 METs). While current guidelines advocate walking at a "brisk" pace for health benefits (American College of Sports Medicine, 2017), our data indicate that a self-selected "comfortable" walking speed should be enough in this population, with associated MET values >3.0. Instructions to participants to "walk at a comfortable pace" indeed resulted in moderate-intensity activity according to MET recordings. Moreover, the fact that comfortable and brisk paces were significantly different in terms of energy spent indicates that using self-selected walking speeds in a laboratory study is a feasible method, and more ecologically valid than treadmill walking.

In summary, we provide modified cutpoints for sedentary, LPA, and MVPA in older adults. However, the question of accelerometer wear site (dominant or non-dominant wrist) still remains. In particular, if studying a relatively sedentary population for which most of the PA will be LPA, such as the retired older adults in our study, it may be beneficial to use dominant-wrist-worn devices as our data suggest they provide more accurate estimates of time spent in LPA vs. MPA.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of South Australia Human Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GP, FF, and DP conceptualized the study. DP and DK led the data collection. FF, DP, and DK analyzed the data. FF led the writing. GP, AR, and RE provided expert advice and critical review of this manuscript. All authors reviewed the manuscript.

FUNDING

AR was supported by the NIHR Leicester Biomedical Research Centre, and the Collaboration for leadership in Applied Health Research and Care (CLAHRC) East Midlands. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or Department of Health.

REFERENCES

American College of Sports Medicine (2017). *ACSM's Guidelines for Exercise Testing and Prescription. 10th Edn*. Philadelphia, PA: Wolters Kluwer.

Bassett, D. R. Jr., Rowlands, A., and Trost, S. G. (2012). Calibration and validation of wearable monitors. *Med. Sci. Sports Exerc. 44(1 Suppl* 1), S32–S8. doi: 10.1249/MSS.0b013e3182399cf7

Copeland, J. L., and Esliger, D. W. (2009). Accelerometer assessment of physical activity in active, healthy older adults. *J. Aging Phys. Act*. 17, 17–30. doi: 10.1123/japa.17.1.17

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M. H., et al. (2017). Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study. *PLoS ONE* 12:e0169649. doi: 10.1371/journal.pone.0169649

Duncan, M. J., Rowlands, A., Lawson, C., Leddington Wright, S., Hill, M., Morris, M., et al. (2019). Using accelerometry to classify physical activity intensity in older adults: what is the optimal wear-site? *Eur. J. Sports Sci*. 20, 1131–1139. doi: 10.1080/17461391.2019.1694078

Ekelund, U., Tarp, J., Steene-Johannessen, J., Hansen, B. H., Jefferis, B., Fagerland, M. W., et al. (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ* 366:l4570. doi: 10.1136/bmj.l4570

Esliger, D. W., Rowlands, A. V., Hurst, T. L., Catt, M., Murray, P., and Eston, R. G. (2011). Validation of the GENEA accelerometer. *Med. Sci. Sports Exerc*. 43, 1085–1093. doi: 10.1249/MSS.0b013e31820513be

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., and McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *J. Sports Sci*. 26, 1557–1565. doi: 10.1080/02640410802334196

Evenson, K. R., Wen, F., Herring, A. H., Di, C., LaMonte, M. J., Tinker, L. F., et al. (2015). Calibrating physical activity intensity for hip-worn accelerometry in women age 60 to 91 years: the women's health initiative OPACH calibration study. *Prev. Med. Rep*. 2, 750–756. doi: 10.1016/j.pmedr.2015.08.021

Fairclough, S., Noonan, R., Rowlands, A., Van Hees, V., Knowles, Z., and Boddy, L. (2016). Wear compliance and activity in children wearing wrist and hip mounted accelerometers. *Med. Sci. Sports Exerc*. 48, 245–253. doi: 10.1249/MSS.0000000000000771

Fraysse, F., Grobler, A. C., Muller, J., Wake, M., and Olds, T. (2019). Physical activity and sedentary activity: population epidemiology and concordance in Australian children aged 11–12 years and their parents. *BMJ Open* 9(Suppl. 3), 136–146. doi: 10.1136/bmjopen-2018-023194

Gorman, E., Hanson, H. M., Yang, P. H., Khan, K. M., Liu-Ambrose, T., and Ashe, M. C. (2014). Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. *Eur. Rev. Aging Phys. Act*. 11, 35–49. doi: 10.1007/s11556-013-0132-x

Hildebrand, M., Hansen, B. H., van Hees, V. T., and Ekelund, U. (2017). Evaluation of raw acceleration sedentary thresholds in children and adults. *Scand. J. Med. Sci. Sports* 27, 1814–1823. doi: 10.1111/sms.12795

Jefferis, B. J., Parsons, T. J., Sartini, C., Ash, S., Lennon, L. T., Papacosta, O., et al. (2019). Objectively measured physical activity, sedentary behaviour and all-cause mortality in older men: does volume of activity matter more than pattern of accumulation? *Br. J. Sports Med*. 53, 1013–1020. doi: 10.1136/bjsports-2017-098733

Klenk, J., and Kerse, N. (2019). Every step you take. *BMJ* 366:l5051. doi: 10.1136/bmj.l5051

Kwan, M., Woo, J., and Kwok, T. (2004). The standard oxygen consumption value equivalent to one metabolic equivalent (3.5 ml/min/kg) is not appropriate for elderly people. *Int. J. Food Sci. Nutr*. 55, 179–182. doi: 10.1080/09637480410001725201

Li, P., Lim, A. S., Gao, L., Hu, C., Yu, L., Bennett, D. A., et al. (2019). More random motor activity fluctuations predict incident frailty, disability, and mortality. *Sci. Transl. Med.* 11:eaax1977. doi: 10.1126/scitranslmed.aax1977

Mankowski, R. T., Anton, S. D., Axtell, R., Chen, S. H., Fielding, R. A., Glynn, N. W., et al. (2017). Device-measured physical activity as a predictor of disability in mobility-limited older adults. *J. Am. Geriatr. Soc.* 65, 2251–2256. doi: 10.1111/jgs.15037

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., et al. (2008). Amount of time spent in sedentary behaviors in the united states, 2003–2004. *Am. J. Epidemiol.* 167, 875–881. doi: 10.1093/aje/kwm390

Migueles, J. H., Cadenas-Sanchez, C., Rowlands, A. V., Henriksson, P., Shiroma, E. J., Acosta, F. M., et al. (2019). Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cutpoints for the assessment of physical activity in adults. *Sci. Rep.* 9:18235. doi: 10.1038/s41598-019-54267-y

Mitchell, B. L., Smith, A. E., Rowlands, A. V., Fraysse, F., Parfitt, G., Lewis, N. R., et al. (2019). Promoting physical activity in rural Australian adults using an online intervention. *J. Sci. Med. Sport* 22, 70–75. doi: 10.1016/j.jsams.2018.07.002

Montoye, A. H., Clevenger, K. A., Pfeiffer, K. A., Nelson, M. B., Bock, J. M., Imboden, M. T., et al. (2020). Development of cut-points for determining activity intensity from a wrist-worn actigraph accelerometer in free-living adults. *J. Sports Sci.* 38, 2569–2578. doi: 10.1080/02640414.2020.1794244

Osborn, W., Simm, P., Olds, T., Lycett, K., Mensah, F., Muller, J., et al. (2018). Bone health, activity and sedentariness at age 11-12 years: cross-sectional Australian population-derived study. *Bone* 112, 153–160. doi: 10.1016/j.bone.2018.04.011

Phillips, L. R., Parfitt, G., and Rowlands, A. V. (2013). Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J. Sci. Med. Sport.* 16, 124–128. doi: 10.1016/j.jsams.2012.05.013

Rejeski, W. J., Marsh, A. P., Brubaker, P. H., Buman, M., Fielding, R. A., Hire, D., et al. (2015). Analysis and interpretation of accelerometry data in older adults: the LIFE study. *J. Gerontol. A Biol. Sci. Med. Sci.* 71, 521–528. doi: 10.1093/gerona/glv204

Rowlands, A. V., Mirkes, E. M., Yates, T. E., Clemes, S. A., Davies, M. J., Khunti, K., et al. (2017). Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Med. Sci. Sports Exerc.* 50, 257–265. doi: 10.1249/MSS.0000000000001435

Rowlands, A. V., Plekhanova, T., Yates, T., Mirkes, E. M., Davies, M., Khunti, K., et al. (2019). Providing a basis for harmonization of accelerometer-assessed physical activity outcomes across epidemiological datasets. *J. Measure. Phys. Behav.* 2, 131–142. doi: 10.1123/jmpb.2018-0073

Sanders, G. J., Boddy, L. M., Sparks, S. A., Curry, W. B., Roe, B., Kaehne, A., et al. (2019). Evaluation of wrist and hip sedentary behaviour and moderate-to-vigorous physical activity raw acceleration cutpoints in older adults. *J. Sports Sci.* 37, 1270–1279. doi: 10.1080/02640414.2018.1555904

Sasaki, J. E., John, D., and Freedson, P. S. (2011). Validation and comparison of actigraph activity monitors. *J. Sci. Med. Sport* 14, 411–416. doi: 10.1016/j.jsams.2011.04.003

Schaefer, C. A., Nigg, C. R., Hill, J. O., Brink, L. A., and Browning, R. C. (2014). Establishing and evaluating wrist cutpoints for the GENEActiv accelerometer in youth. *Med. Sci. Sports Exerc.* 46, 826–33. doi: 10.1249/MSS.0000000000000150

Scott, J. J., Rowlands, A. V., Cliff, D. P., Morgan, P. J., Plotnikoff, R. C., and Lubans, D. R. (2017). Comparability and feasibility of wrist- and hip-worn accelerometers in free-living adolescents. *J. Sci. Med. Sport* 20, 1101–1106. doi: 10.1016/j.jsams.2017.04.017

Stamatakis, E., Kelly, P., Strain, T., Murtagh, E. M., Ding, D., and Murphy, M. H. (2018). Self-rated walking pace and all-cause, cardiovascular disease and cancer mortality: individual participant pooled analysis of 50 225 walkers from 11 population British cohorts. *Br. J. Sports Med.* 52, 761–768. doi: 10.1136/bjsports-2017-098677

Troiano, R. P., Berrigan, D., Dodd, K. W., Masse, L. C., Tilert, T., and McDowell, M. (2008). Physical activity in the united states measured by accelerometer. *Med. Sci. Sports Exerc.* 40, 181–188. doi: 10.1249/mss.0b013e31815a51b3

Whitcher, L., and Papadopoulos, C. (2014). Accelerometer derived activity counts and oxygen consumption between young and older individuals. *J. Aging Res.* 2014:184693. doi: 10.1155/2014/184693

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

frontiers
in Sports and Active Living

CASE REPORT
published: 11 February 2021
doi: 10.3389/fspor.2021.625656

Check for updates

Case Report: Adjusting Seat and Backrest Angle Improves Performance in an Elite Paralympic Rower

Anna Cecilia Severin[1]*, *Jørgen Danielsen*[1], *Jørgen Falck Erichsen*[2], *Sindre Wold Eikevåg*[2,3], *Martin Steinert*[3], *Gertjan Ettema*[1] *and Julia Kathrin Baumgart*[1]

[1] Department of Neuromedicine and Movement Science, Center for Elite Sports Research, Norwegian University of Science and Technology, Trondheim, Norway, [2] Department of Civil and Environmental Engineering, Center for Sports Facilities and Technology, Norwegian University of Science and Technology, Trondheim, Norway, [3] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

OPEN ACCESS

Edited by:
Laura Gastaldi,
Politecnico di Torino, Italy

Reviewed by:
Gunnar Treff,
University of Ulm, Germany
Frédérique Hintzy,
Université Savoie Mont Blanc, France

***Correspondence:**
Anna Cecilia Severin
cecilia.severin@ntnu.no
orcid.org/0000-0003-0404-1072

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: 03 November 2020
Accepted: 18 January 2021
Published: 11 February 2021

Citation:
Severin AC, Danielsen J, Falck Erichsen J, Wold Eikevåg S, Steinert M, Ettema G and Baumgart JK (2021) Case Report: Adjusting Seat and Backrest Angle Improves Performance in an Elite Paralympic Rower. Front. Sports Act. Living 3:625656. doi: 10.3389/fspor.2021.625656

Paralympic rowers with functional impairments of the legs and trunk rely on appropriate seat configurations for performance. We compared performance, physiology, and biomechanics of an elite Paralympic rower competing in the PR1 class during ergometer rowing in a seat with three different seat and backrest inclination configurations. Unlike able-bodied rowers, PR1 rowers are required to use a seat with a backrest. For this study, we examined the following seat/backrest configurations: conA: 7.5°/25°, conB: 0°/25°, and conC: 0°/5° (usually used by the participant). All data was collected on a single day, i.e., in each configuration, one 4-min submaximal (100 W) and one maximal (all-out) stage was performed. The rowing ergometer provided the average power and (virtual) distance of each stage, while motion capture provided kinematic data, a load cell measured the force exerted on the ergometer chain, and an ergospirometer measured oxygen uptake ($\dot{V}O_2$). Where appropriate, a Friedman's test with *post-hoc* comparisons performed with Wilcoxon signed-ranked tests identified differences between the configurations. Despite similar distances covered during the submaximal intensity (conA: 793, conB: 793, conC: 787 m), the peak force was lower in conC (conA: 509, conB: 458, conC: 312 N) while the stroke rate (conA: 27 conB: 31, conC: 49 strokes·min^{-1}) and $\dot{V}O_2$ (conA: 34.4, conB: 35.4, conC: 39.6 mL·kg^{-1}·min^{-1}) were higher. During the maximal stage, the virtual distances were 7–9% longer in conA and conB, with higher peak forces (conA: 934 m, 408 N, conB: 918 m, 418 N, conC: 856 m, 331 N), and lower stroke rates (conA: 51, conB: 54, conC: 56 strokes·min^{-1}), though there was no difference in $\dot{V}O_{2peak}$ (~47 ml^{-1}·kg^{-1}·min^{-1}). At both intensities, trunk range of motion was significantly larger in configurations conA and conB. Although fatigue may have accumulated during the test day, this study showed that a more inclined seat and backrest during ergometer rowing improved the performance of a successful Paralympic PR1 rower. The considerable increase in ergometer rowing performance in one of the top Paralympic rowers in the world is astonishing and highlights the importance of designing equipment that can be adjusted to match the individual needs of Paralympic athletes.

Keywords: kinematics, paraplegia, elite athlete, equipment modification, rowing ergometer

INTRODUCTION

Paralympic rowers compete in three classes; PR1 for athletes with no leg function, minimal/no trunk function, and poor sitting stability, PR2 for athletes with limited/no leg function and functional use of the trunk, and PR3 for athletes with residual leg function (https://bit.ly/370Scrz, accessed December 4, 2020). While Paralympic rowers compete over the same 2000-m distance as Olympic rowers, the current world records for male and female PR1 rowers are around 3 min slower (∼7 vs. 10 min) than the world record for able-bodied rowers (www.worldrowing.com/events/statistics, accessed October 22, 2020). The faster times in able-bodied rowers are mainly because of the ability to utilize their whole body during the rowing task (Baudouin and Hawkins, 2002; Maestu et al., 2005; Van Soest and Hofmijster, 2009). In addition, while both PR1 and PR2 rowers use a fixed seat, PR1 rowers have less sitting stability than PR2 rowers, and are thus required to be strapped into their seat during competition. Therefore, PR1 rowers rely predominantly on their arms and shoulders to generate the boat speed (Cutler et al., 2017).

Regardless of whether rowers are able to actively utilize their legs or not, the purpose of the sport is to cover the race distance as fast as possible. The boat speed is dependent on the propulsive force produced (Baudouin and Hawkins, 2004), which in turn depends on the physical capabilities and technique of the rower, and the configuration and design of the equipment (Baudouin and Hawkins, 2002; McGregor et al., 2004). Burkett (2010) highlighted that the seat can be modified to match the individual needs of the athlete in Paralympic rowing, and the World Rowing Federation (WRF) currently has few restrictions with regard to seat configurations. The only regulations state that that PR1 athletes must have a backrest on their seat and use a trunk strap for safety purposes with specifications on how these straps should be formed and function (Rolland and Smith, 2017). While using and adapting equipment to match the requirements of the individual Paralympic rower may have a large effect on performance, such effects have not been reported in the literature.

To date, most research on seat modifications for Paralympic performance has been conducted on wheelchair sports (e.g., Costa et al., 2009; Vanlandewijck et al., 2011; Van Der Slikke et al., 2018). Vanlandewijck et al. (2011) found that utilizing a more posteriorly inclined seat can benefit seating stability but highlighted that it may also have negative effects on performance. This was because the increased hip flexion angle and pelvic posterior tilt appeared to reduce the trunk and shoulder range of motion (ROM). Contrary to wheelchair propulsion, rowing propulsion is comprised of a backward pull and thus relies more on trunk extension. It is therefore possible that adjusting the inclination of the backrest, and thereby allow more trunk extension, may compensate for an inclined seat. This may, in turn, allow the athlete to regain some of the restricted motion and improve performance. However, it remains unknown if this applies to Paralympic rowers with minimal trunk function. This case report therefore aimed at assessing the effects of a more inclined seat and backrest on rowing performance in a multiple Paralympic PR1 world champion.

CASE DESCRIPTION

The participant was an elite female Paralympic PR1 rower (age: 30 years, height: 1.80 m, body mass: 60 kg), who acquired an incomplete spinal cord injury in 2008 at the level of the 10th thoracic vertebra, leaving her with minimal trunk function and reduced sitting stability (see https://bit.ly/370Scrz for a description of the tests performed during classification). At the time of the data collection she did not have any additional injuries, was in good health, and trained ∼28 h per week. Written informed consent was obtained prior to data collection, and the testing complied with the declaration of Helsinki.

LABORATORY TESTING AND MEASURED VARIABLES

The participant attended the laboratory on 2 consecutive days, with pilot testing and familiarization on day 1, and the data collection on day 2. A custom-made test seat replaced the original seat on a Concept2 rowing ergometer (Concept2, Morrisville. VT. USA). The seat and backrest inclinations were adjustable but the seat itself was stationary (non-sliding). Based on the pilot testing from day 1, the three seat configurations analyzed on day 2 were: seat 7.5° (from horizontal) and backrest 25° (from vertical) (conA), seat 0° and backrest 25° (conB), and her usual configuration seat 0° and backrest 5° (conC) (**Figure 1**). The participant was strapped into the seat with one strap across her upper thighs, and one strap around her lower trunk, similar to her competition set-up.

The Concept2 software provided the virtual rowing distance covered (henceforth referred to as distance) and the average power output, while a Futek Miniature Load Cell (Futek LCM200; capacity, 250 lbs.; nonlinearity 0.5%; hysteresis 0.5%; weight 17 g; Futek Inc., Irvine, CA) was used to record the instantaneous force exerted by the participant on the chain of the ergometer (200 Hz). The load cell was calibrated against a range of forces of known magnitude employing calibrated weights (linear correlation $r^2 = 0.999$). Kinematics were collected by a 10-camera system (Oqus, Qualisys AB, Gothenburg, Sweden) recording at 100 Hz. Bilateral symmetry was assumed, and retroreflective markers were attached to the participants left side on the 2nd toe, lateral malleolus, lateral femoral epicondyle, greater trochanter, iliac crest, the spinous processes of the T10 and C7 vertebrae, acromion process, lateral epicondyle of the humerus, and styloid process of the radius. One additional marker was placed on the ergometer handle and one on the flywheel, allowing for identifying strokes. Rate of oxygen uptake ($\dot{V}O_2$) was recorded using an ergospirometer with a mixing chamber (Oxycon Pro, Jaeger GmbH, Hoechberg, Germany) and a mouthpiece (Hans Rudolph Inc, Kansas City, MO, USA). Prior to testing, the gas analyzer was calibrated against a known mixture of gases (15% O_2 and 5% CO_2) and ambient air. Calibration of the flow transducer was manually performed with a 3L high precision syringe (Hans Rudoph Inc., Kansas City, MO, USA). Heart rate (HR) was monitored using an H10 Polar heart rate monitor (Polar Electro Inc., Kempele, Finland). Blood lactate

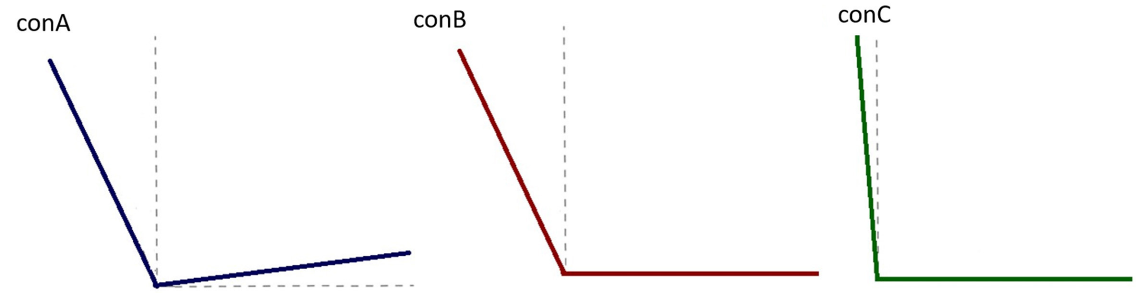

FIGURE 1 | Illustration of the three seat configurations used during this study. conA had a 25° incline of the backrest from the vertical plane, with a 7.5° incline of the seat to the horizonal plane. conB had the same backrest incline as conA (25° to the vertical) and a flat seat (0° to the horizontal). conC, which was the participant's usual set up, had the backrest inclined to 5° from the vertical plane and a flat seat (0° from the horizontal plane).

concentration (BLa) was assessed with the Lactate pro 2 (Arkray Inc., Kyoto, Japan). Subjective rate of perceived exertion (RPE) was measured on a 6–20 Borg scale (Borg, 1982).

The data collection protocol consisted of three 4-min stages performed at 80 W, 100 W, and an all-out effort in each of the three seat configurations. During the all-out stage, the participant was instructed to row as hard as she could for the 4 mins and pace herself so that she reached exhaustion toward the end of the stage. Maximal exhaustion was considered reached if 2 of the 3 following criteria were met: (1) the self-reported max heart rate from the participant, (2) respiratory exchange ratio over 1.15, and (3) an RPE of 18 or higher. The participant was allowed 2–3 mins rest between stages and 30 mins between the different configurations. The 80 W stages were considered familiarization stages and were not included in the analysis. The 100 W (SUBMAX) stages provided steady-state responses while the all-out (MAX) stages provided peak responses. Performance (i.e., distance covered during MAX), biomechanical, and physiological data were recorded throughout the 4-min stages. RPE was recorded after each stage and a BLa was measured from the earlobe directly after SUBMAX, and 1 and 3 min after MAX.

SUBMAX steady-state $\dot{V}O_2$ and HR data were calculated by averaging the final 60 s of each stage. For MAX, the data was analyzed using a 30 s moving average for the $\dot{V}O_2$ and 30 s for the HR, and the peak value was identified as $\dot{V}O_{2peak}$ and HR_{peak}, respectively. Kinematic and force data were analyzed using custom MATLAB code (MATLAB 2019b, Matworks Inc., Nantick, MA, USA). Marker and force data were low-pass filtered using a 4th order Butterworth filter with cut-offs of 7 and 50 Hz, respectively. Elbow and shoulder joint and trunk angles were calculated from marker positions (**Figure 2**).

The start of each rowing stroke was defined as the point where the handle marker was closest to the flywheel marker. The stroke was divided into a drive and a recovery phase (Cutler et al., 2017), with the end of drive identified as when the handle was farthest away from the flywheel. Eighteen strokes in the middle of each stage were extracted for analysis. For each stroke, timeseries data for the joint angles (trunk, shoulder, and elbow) and force data from the load cell were time-normalized to 101 data points (0–100% of each stroke). In addition, the following discrete biomechanical variables were extracted for the 18 cycles: maximal and minimal joint angles, peak force, impulse (the integral of force over time), the drive phase duration (expressed as % of stroke), the stroke rate, and stroke length (i.e., distance the handle moved during the drive phase).

The joint angles, peak force, impulse, drive phase duration, stroke rate, and stroke length were analyzed in SPSS version 26 (IBM Inc., Armonk, NY, USA). All variables violated the assumptions of a repeated-measures one-way ANOVA so differences between configurations were determined using a Friedmans test with a subsequent Wilcoxon singed-ranks tests for *post-hoc* comparisons. Statistical significance of all *post-hoc* tests was accepted at an alpha level of 0.017 ($0.05/3 \approx 0.017$ following Bonferroni adjustments). Cohen's D was used to indicate effect size, and was considered small if $d < 0.5$, moderate if $0.5 < d < 0.8$, and large if $d > 0.8$ (Cohen, 1988). The drive phase of the normalized time series were analyzed using Statistical Parametric Mapping (SPM) (github.com/0todd0000/spm1dmatlab, accessed March 17, 2020) by employing paired samples *T*-tests (Pataky et al., 2016). The time interval used for the SPM analysis was chosen as the drive phase duration for conA and conB that occurred last (38% for SUBMAX from conA, and 51% for MAX from conB). Statistical significance for the SPM analyses was accepted at an alpha level of 0.05.

OUTCOMES AND RESULTS

The Friedmans tests indicated significant main effects of seat configurations on all tested variables ($\chi^2(2) < 36.000$, $p < 0.05$). **Table 1** shows the results of the *post-hoc* comparisons for the biomechanical variables along with descriptive data for the performance and physiological variables. During SUBMAX, conC had significantly lower peak force and impulse coupled with higher $\dot{V}O_2$ and significantly higher stroke rate than conA and conB. Further, although the distance, $\dot{V}O_2$, and RPE were similar between conA and conB, conB had higher HR and significantly lower peak force and impulse. During MAX, longer distances were covered in conA (+78 meters) and conB (+60 meters), compared to conC (**Table 1**, **Figure 3**). Peak force was significantly higher and stroke rate was significantly lower in

TABLE 1 | Performance, physiological, and biomechanical variables presented as single values or mean ± SD for the three seat configurations tested on day two with statistical comparisons for the three configurations.

	100 W						MAX					
	conA	**conB**	**conC**	**conA v conB**	**conA v conC**	**conB v conC**	**conA**	**conB**	**conC**	**conA v conB**	**conA v conC**	**conB v conC**
Distance (m)	793	793	787	–	–	–	934	918	856	–	–	–
Power output (W)	101	101	99	–	–	–	165	157	127	–	–	–
$\dot{V}O_2$ ($mL{\cdot}kg^{-1}{\cdot}min^{-1}$)	34.4	35.4	39.6	–	–	–	46.3	46.2	47.4	–	–	–
HR (bpm)	157	166	176	–	–	–	188	188	187	–	–	–
BLa ($mmol{\cdot}L^{-1}$)	3.6	6.4	11.3	–	–	–	21.8	23.5	18.4	–	–	–
RPE	11	12	14	–	–	–	20	19	19	–	–	–
Peak Force (N)	509 ± 40	458 ± 30	312 ± 28	0.003$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	408 ± 17	418 ± 29	331 ± 23	0.170	<0.001$^{\alpha}$	<0.001$^{\alpha}$
Impulse (N·s)	172.2 ± 8.1	157.9 ± 8.7	97.0 ± 8.9	<0.001$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	153.5 ± 5.9	129.9 ± 10.0	102.9 ± 6.5	<0.001$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$
Drive phase duration (%)	32 ± 1	37 ± 1	51 ± 1	<0.001$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	52 ± 1	54 ± 1	56 ± 1	0.001$^{\alpha}$	<0.001$^{\alpha}$	0.001$^{\alpha}$
Stroke rate (spm)	26.5 ± 0.7	30.8 ± 0.8	48.6 ± 1.4	<0.001$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	50.9 ± 0.8	54.1 ± 1.2	56.0 ± 0.7	<0.001$^{\alpha}$	<0.001$^{\alpha}$	0.001$^{\alpha}$
Stroke length (cm)	80.5 ± 2	77.7 ± 1.2	67.2 ± 1.3	0.002$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	78.7 ± 1.2	77.3 ± 1.7	68.4 ± 1.1	0.006$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$
Trunk flexion (°)	68.7 ± 0.8	70.9 ± 0.9	69.3 ± 1.2	<0.001$^{\alpha}$	0.170$^{\beta}$	<0.001$^{\alpha}$	69.8 ± 0.8	71.6 ± 1.1	70.2 ± 1.1	<0.001$^{\alpha}$	0.053	0.004$^{\alpha}$
Trunk extension (°)	123.9 ± 1.5	124.5 ± 1.1	106.4 ± 0.7	0.102	<0.001$^{\alpha}$	<0.001$^{\alpha}$	127.4 ± 1.1	129.6 ± 0.8	109.4 ± 0.9	<0.001$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$
Trunk ROM (°)	55.2 ± 1.8	53.6 ±1.2	37.2 ± 1.1	0.011$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	57.7 ± 1.3	58.0 ± 1.5	39.2 ±1.3	0.446	<0.001$^{\alpha}$	<0.001$^{\alpha}$
Shoulder flexion (°)	82.4 ± 1.4	81.7 ± 1.5	80.1 ± 2.0	0.148	0.004$^{\alpha}$	0.025$^{\alpha}$	84.7 ± 1.4	76.7 ± 3.5	75.1 ± 1.8	<0.001$^{\alpha}$	<0.001$^{\alpha}$	0.094$^{\beta}$
Shoulder extension (°)	51.8 ± 2.7	50.3 ± 2.2	45.5 ± 2.1	0.064$^{\beta}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	48.5 ± 1.6	38.9 ± 3.2	44.2 ± 2.4	<0.001$^{\alpha}$	0.002$^{\alpha}$	0.001$^{\alpha}$
Shoulder ROM (°)	134.1 ± 3.5	132.0 ± 2.6	125.6 ± 3.2	0.043$^{\beta}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	133.1 ± 1.7	115.7 ± 4.5	119.3 ± 3.2	<0.001$^{\alpha}$	<0.001$^{\alpha}$	0.018$^{\alpha}$
Elbow flexion (°)	140.1 ± 2.2	139.5 ± 1.8	134.8 ± 1.7	0.446	<0.001$^{\alpha}$	<0.001$^{\alpha}$	139.0 ± 1.7	130.9 ± 3.0	130.4 ± 1.8	<0.001$^{\alpha}$	<0.001$^{\alpha}$	0.586
Elbow extension (°)	54.4 ± 1.8	58.6 ± 3.0	62.5 ± 2.0	0.002$^{\alpha}$	<0.001$^{\alpha}$	0.002$^{\alpha}$	59.9 ± 2.3	68.6 ± 2.0	64.4 ± 2.1	<0.001$^{\alpha}$	0.001$^{\alpha}$	<0.001$^{\alpha}$
Elbow ROM (°)	85.7 ± 3.1	80.9 ± 3.5	72.5 ± 2.6	0.010$^{\alpha}$	<0.001$^{\alpha}$	<0.001$^{\alpha}$	79.1 ± 2.3	62.5 ± 4.3	65.7 ± 2.3	<0.001$^{\alpha}$	<0.001$^{\alpha}$	0.013$^{\alpha}$

bpm, beats per min; cm, centimeters; d, Cohen's D; m, meters; N, Newtons; p, p-value; spm, strokes per min; $\dot{V}O2$, Oxygen uptake; W, watt.
Drive phase duration indicates the timing of when the participant transitioned from the drive phase to the recovery phase.
$^{\alpha}$denotes a large effect size ($d > 0.8$ or $d < -0.8$), $^{\beta}$denotes a moderate effect size ($0.5 < d < 0.8$ or $-0.8 < d < -0.5$).

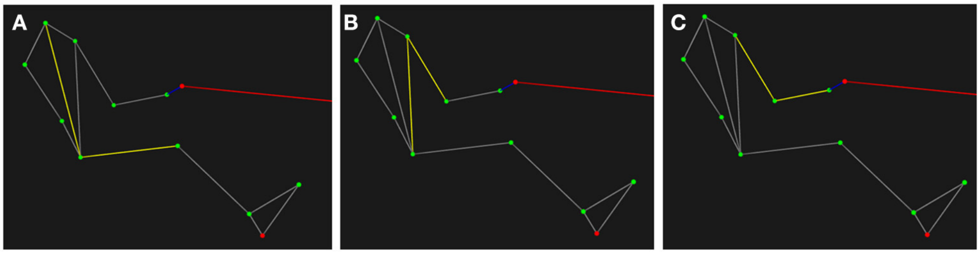

FIGURE 2 | Definition of the angles calculated from the kinematic data (indicated by the yellow lines). **(A)** Trunk: angle created between the thigh and neck using markers on the lateral femoral epicondyle, the greater trochanter, and the spinous process of C7. **(B)** Shoulder: angle created between the greater trochanter, acromion process, and lateral humeral epicondyle. **(C)** Elbow: angle created between the acromion process, lateral humeral epicondyle, and radial styloid process.

conA and conB compared to conC also at MAX, although these differences were smaller than at SUBMAX.

Trunk extension was significantly less in conC compared to conA and conB during both intensities (**Table 1**, **Figure 3**). The SPM analysis showed significant differences between configurations conA and conB (shaded areas in **Figure 3**) in both their force profiles and joint kinematics throughout a large part of the drive phase during both intensities. While the elbow and shoulder joints showed a similar pattern during all three configurations, the timing of their peak flexion significantly differed between configurations during SUBMAX (**Figure 3**). Though the differences in peak shoulder and elbow flexion and extension angles between conC and the other configurations were small (less than 6.0°), the consistent movement pattern of the participant resulted in these differences reaching statistical significance (**Table 1**).

DISCUSSION

In this case report, the performance of an elite Paralympic rower improved substantially during an all-out maximal effort on a rowing ergometer when using adjusted seat configurations. Compared to her usual setup (conC), the configurations with an increased back angle (conA and conB) showed 7–9% improved performance (virtual distance covered) along with significantly higher peak force production, larger impulse, increased trunk motion and longer stroke length coupled with lower stroke frequency.

During SUBMAX, the participant was able to maintain the target power (100 W) in all three configurations, and therefore covered similar distances. However, she had to employ a higher stroke rate in her usual setup (49 strokes·min^{-1}) compared to conA and conB, which is considerably higher than what usually is reported for able-bodied rowers (20–36 strokes·min^{-1}) (McGregor et al., 2004; Hofmijster et al., 2007). This was to compensate for the lower peak force, lower impulse, and shorter stroke length. The higher stroke rate was achieved predominantly through a shorter recovery phase in conC (conA: 1.53 s, conB: 1.23 s, and conC: 0.60 s), which required an active contribution from the participant to return the handle toward the flywheel before the next stroke. In addition, with a high stroke rate, the participant moved faster and changed the direction of movement more frequently, which required her to continuously overcome larger linear momentum. Further, while the drive phase durations only differed by < 0.1 s between all three configurations, the stroke lengths in conA and conB were ∼10 cm longer than conC during SUBMAX. It has been shown in able-bodied rowers, that the amount of positive work done per stroke during rowing is mainly dependent on stroke length (Hofmijster et al., 2007). Consequently, the high stroke rate with a shorter recovery phase, and shorter stroke lengths, would be disadvantageous for producing work. A high stroke rate has further been linked to increased respiratory demands (Saltin et al., 1998; Lindinger and Holmberg, 2011), which is supported by the current study in that the higher $\dot{V}O_2$ in conC indicates a lower efficiency than in conA and conB.

Surprisingly, both peak force and impulse were higher at SUBMAX (work rate 100 W) than at MAX for both conA and conB (work rate 165 and 157 W, respectively). In addition, her stroke rate increased considerably for conA and conB (+83 and +57%, respectively) at MAX compared to SUBMAX, while conC only increased with 15%. It is also noteworthy that the stoke rate only differed by ∼10% between the three configurations at MAX, which suggests that she has a "default" stroke rate during "all-out" effort bouts. It seems the participant adopted this "default" stroke rate when performing an "all-out" effort during the testing, which subsequently shortened the time per cycle, and caused the lower impulse and peak force at MAX. Importantly, if the participant adopted this "default" stroke rate, it suggests the "all-out" instruction triggered a rowing technique that was different from the one used when instructed to maintain a target power (i.e., SUBMAX). So, while the participant covered a longer distance at MAX in conA and conB, this was done with a less powerful drive phase. Our findings therefore suggest that she may be able to perform even better if she can maintain a rowing technique at MAX, with a more powerful drive phase and lower stroke rate.

Further, while stroke rate and drive phase duration differed considerably between conC and the other configurations at SUBMAX, they were more similar at MAX, suggesting that the participants "all-out" effort strategy was similar regardless of the seat configuration. Conversely, the gains the participant achieved in peak force/impulse when transitioning from SUBMAX to

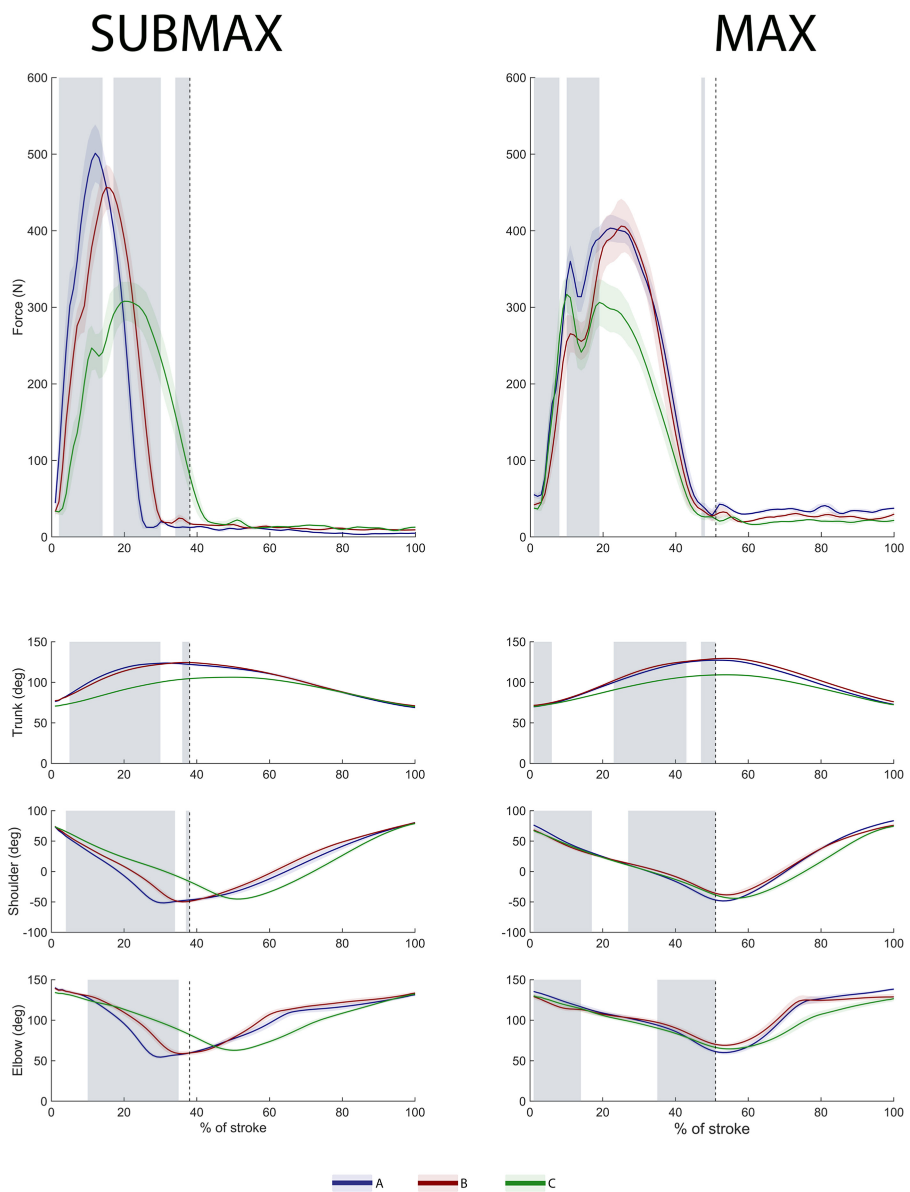

FIGURE 3 | Mean ± standard deviation of the handle force and joint angles for the 3 seat configurations (conA, conB, and conC) for 0–100% of a stroke. The vertical dashed line indicates the point of phase shift (between drive and recovery phase), up until which the SPM analysis was performed. The shaded area in gray indicates when there was a significant difference during the drive phase between conA and conB since conC was statistically different to the other configurations during most of the drive phase.

MAX were noticeably smaller in conC than conA and conB. This was perhaps since she was not able to increase the already high stroke rate much further in conC (+15%), which also resulted in a shorter distance covered. In line with the lower efficiency during SUBMAX, the performance was poorer in conC during MAX, despite similar levels of volitional exhaustion ($\dot{V}O_{2peak}$, HR_{peak} and RPE) in all configurations.

Even in a participant with minimal residual trunk function, the increased performance in conA and conB was likely associated with the increased trunk motions due to the inclined backrest (**Table 1**). This supports previous research that linked increased power production to increased trunk ROM in able-bodied male rowers (McGregor et al., 2004). The increased trunk motion likely also triggered the arm movements earlier in the stroke during SUBMAX (**Figure 3**), which subsequently allowed the more rapid force development and the longer recovery phases (**Table 1**). The backrest inclination was the same for conA and conB (25°), so the marginally better performance in conA may in part be due to the increased seat inclination (conA: 7.5°, conB: 0°). Speculatively, the inclined seat may have prevented the participant from sliding forward during the strokes and thereby increased her stability. In wheelchair athletes, an inclined seat has been cautioned to have negative effects on performance since it creates "closed" posture with reduced trunk ROM (Vanlandewijck et al., 2011). However, the difference in seat angle between conA and conB caused only minimal differences in trunk angles (flexion: 2.2° and 1.8°, extension: 0.6° and 2.2° in SUBMAX and MAX, respectively), and the more closed posture in conA did not have a negative effect on her performance. Furthermore, wheelchair propulsion and rowing are opposite movements, and it is therefore likely that rowing performance would be more affected by the range of trunk extension and is not as affected by limited trunk flexion as wheelchair performance. This further highlights the importance of allowing trunk extension even for Paralympic rowers with minimal residual trunk function. Overall, this data shows the importance of designing individualized equipment to match the very heterogenous physical capabilities of Paralympic athletes.

SUBJECT PERSPECTIVE

Following this experiment, the athlete chose to employ conA and conB during training and has, after a few months, settled with conA. During the experiment, the athlete commented that conA and conB felt "easier and more effective," and that she "didn't have to use so much energy." The coach also observed that athlete seemed more relaxed in these adjusted configurations and particularly noticed the lower stroke rate.

LIMITATIONS

A limitation of this case report is that because of time restriction, all configurations were tested on 1 day, with conA first, then conB, and conC last. Despite measures to prevent fatigue, it is possible that the results may in part be attributed to accumulating fatigue. However, the differences between conC and the other configurations are so large that it seems unlikely that these effects would disappear completely even if fatigue was avoided. Some of the significant differences in kinematics between the configurations were also very small (e.g., trunk angle, **Figure 3**). This was caused by the very consistent movement patterns from the single, experienced rower in this case report, resulting in small standard deviations. However, in elite sport, even such small changes may still affect the athlete's chances of winning a medal or finishing off the podium. Finally, even though differences exist between indoor ergometers and on-water rowing (Shaharudin et al., 2014), ergometers are frequently used by high performance rowers during testing and training (Bjerkefors et al., 2007; Van Soest and Hofmijster, 2009; Cutler et al., 2017). On the ergometer, we saw a 7–9% performance improvement in conA and conB, compared to the participants usual set up. Although non-standardized on-water pilot testing has indicated performance improvements with the adjusted seat, the extent of these during competitions remains to be investigated.

CONCLUSION

This case study showed that adjustments to the seat and backrest improved performance by 7–9% in an elite Paralympic PR1 rower compared to her usual configuration during land-based ergometer rowing. The two configurations with increased backrest inclination allowed longer virtual distances, higher peak forces, larger impulses, increased trunk motions, longer stroke lengths, and lower stroke rates compared to the participants usual set-up. It should be acknowledged that the design of the study, where the participant performed three "all-out" tests on a single day, may have resulted in accumulating fatigue and thus affected the results. However, the differences between her usual set up and the adjusted configurations were so large that is seems unlikely that they would disappear completely if fatigue was avoided.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author, AS. The data are not publicly available due to their containing information that could compromise the privacy of the research participant.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Norwegian Centre for Research Data (ID 689366). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

AS participated in the conceptualizing of the study, the data processing, statistical analyses, and writing of the manuscript. JF, SW, and MS designed the seat and modified the Concept2 used in the study, participated in the conceptualization of the study, the data collection, and the writing of the manuscript. JD participated in the conceptualization of the study, the data collection, and the writing of the manuscript. GE contributed to the design of the study, the data analysis, and writing of the manuscript. JB participated in the conceptualization of the study, the data collection, statistical analysis, and the writing of the manuscript. All authors have read and approved the final version of the manuscript and agree with the order of presentation of the authors.

FUNDING

NeXt Move was funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The laboratory facilities and equipment were provided by NeXt Move, Norwegian University of Science and Technology (NTNU).

ACKNOWLEDGMENTS

The authors thank the participant and her coach for their participation and contribution to this study, master students Magnus Kjærnet Bjølseth and Andris Kvam who contributed to developing the seat, as well as Aurora Aasrud Haug and Tonje Ludvigsen for support during the data collection.

REFERENCES

Baudouin, A., and Hawkins, D. (2002). A biomechanical review of factors affecting rowing performance. *Br. J. Sports Med.* 36, 396–402; discussion 402. doi: 10.1136/bjsm.36.6.396

Baudouin, A., and Hawkins, D. (2004). Investigation of biomechanical factors affecting rowing performance. *J. Biomech.* 37, 969–976. doi: 10.1016/j.jbiomech.2003.11.011

Bjerkefors, A., Carpenter, M. G., and Thorstensson, A. (2007). Dynamic trunk stability is improved in paraplegics following kayak ergometer training. *Scand. J. Med. Sci. Sports* 17, 672–679. doi: 10.1111/j.1600-0838.2006.00621.x

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med. Sci. Sports Exerc.* 14, 377–381. doi: 10.1249/00005768-198205000-00012

Burkett, B. (2010). Technology in Paralympic sport: performance enhancement or essential for performance? *Br. J. Sports Med.* 44, 215–220. doi: 10.1136/bjsm.2009.067249

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Costa, G. B., Rubio, M. P., Belloch, S. L., and Soriano, P. P. (2009). Case study: effect of handrim diameter on performance in a paralympic wheelchair athlete. *Adapt. Phys. Activ. Q.* 26, 352–363. doi: 10.1123/apaq.26.4.352

Cutler, B., Eger, T., Merritt, T., and Godwin, A. (2017). Comparing para-rowing set-ups on an ergometer using kinematic movement patterns of able-bodied rowers. *J. Sports Sci.* 35, 777–783. doi: 10.1080/02640414.2016.1189587

Hofmijster, M. J., Landman, E. H., Smith, R. M., and Knoek Van Soest, A. (2007). Effect of stroke rate on the distribution of net mechanical power in rowing. *J. Sports Sci.* 25, 403–411. doi: 10.1080/02640410600718046

Lindinger, S. J., and Holmberg, H. C. (2011). How do elite cross-country skiers adapt to different double poling frequencies at low to high speeds? *Eur. J. Appl. Physiol.* 111, 1103–1119. doi: 10.1007/s00421-010-1736-8

Maestu, J., Jurimae, J., and Jurimae, T. (2005). Monitoring of performance and training in rowing. *Sports Med.* 35, 597–617. doi: 10.2165/00007256-200535070-00005

McGregor, A. H., Bull, A. M., and Byng-Maddick, R. (2004). A comparison of rowing technique at different stroke rates: a description of sequencing, force production and kinematics. *Int. J. Sports Med.* 25, 465–470. doi: 10.1055/s-2004-820936

Pataky, T. C., Robinson, M. A., and Vanrenterghem, J. (2016). Region-of-interest analyses of one-dimensional biomechanical trajectories: bridging 0D and 1D theory, augmenting statistical power. *PeerJ* 4:e2652. doi: 10.7717/peerj.2652

Rolland, J., and Smith, M. (2017). *Para Rowing Competition Regulations Event Regulations and/or Departures From the FISA Rules of Racing.* Lausanne: WR Federation.

Saltin, B., Radegran, G., Koskolou, M. D., and Roach, R. C. (1998). Skeletal muscle blood flow in humans and its regulation during exercise. *Acta Physiol. Scand.* 162, 421–436. doi: 10.1046/j.1365-201X.1998.0293e.x

Shaharudin, S., Zanotto, D., and Agrawal, S. (2014). Muscle Synergies of untrained subjects during 6 min maximal rowing on slides and fixed ergometer. *J. Sports Sci. Med.* 13, 793–800.

Van Der Slikke, R. M., De Witte, A. M., Berger, M. A., Bregman, D. J., and Veeger, D. J. H. (2018). Wheelchair mobility performance enhancement by changing wheelchair properties: What is the effect of grip, seat height, and mass? *Int. J. Sports Physiol. Perf.* 13, 1050–1058. doi: 10.1123/ijspp.2017-0641

Van Soest, A. J., and Hofmijster, M. (2009). Strapping rowers to their sliding seat improves performance during the start of ergometer rowing. *J. Sports Sci.* 27, 283–289. doi: 10.1080/02640410802495336

Vanlandewijck, Y. C., Verellen, J., and Tweedy, S. (2011). Towards evidence-based classification in wheelchair sports: impact of seating position on wheelchair acceleration. *J. Sports Sci.* 29, 1089–1096. doi: 10.1080/02640414.2011.576694

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

frontiers in Sports and Active Living

ORIGINAL RESEARCH
published: 19 February 2021
doi: 10.3389/fspor.2021.644765

Surface Electromyography Spectral Parameters for the Study of Muscle Fatigue in Swimming

Luca Puce[1]*, *Ilaria Pallecchi*[2], *Lucio Marinelli*[1,3], *Laura Mori*[1,3], *Marco Bove*[3,4], *Daniele Diotti*[1], *Piero Ruggeri*[4,5], *Emanuela Faelli*[4,5], *Filippo Cotellessa*[1] and *Carlo Trompetto*[1,3]

[1] Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Genoa, Italy, [2] National Research Council (CNR), SPIN institute, Department of Physics, Genoa, Italy, [3] Department of Neuroscience, IRCCS Ospedale Policlinico San Martino, Genoa, Italy, [4] Department of Experimental Medicine, Section of Human Physiology, University of Genoa, Genoa, Italy, [5] Centro Polifunzionale di Scienze Motorie, University of Genoa, Genoa, Italy

The purpose of this study was to assess validity, stability and sensitivity, of 4 spectral parameters–median frequency (F_{med}), mean frequency (F_{mean}), Dimitrov index (DI), and mean instant frequency (F_{mi})–in measuring localized muscle fatigue in swimming and to investigate their correlation with the variations of kinematic data and mechanical fatigue. Electrophysiological measures of muscle fatigue were obtained in real-time during a 100 m front crawl test at maximum speed in 15 experienced swimmers, using surface electromyography in six muscles employed in front crawl, while kinematic data of swimming was measured from video analysis. Mechanical fatigue was measured as the difference between muscle strength prior to and immediately after the 100 m front crawl in a dry-land multi-stage isometric contraction test. Statistically significant fatigue ($p < 0.0001$) was found for all spectral parameters in all muscles. F_{med} and F_{mean} varied between 10 and 25%, DI between 50 and 150%, and F_{mi} between 5 and 10%. Strong correlation (Pearson $r \geq 0.5$) with mechanical fatigue was found for all spectral parameters except for F_{mi} and it was strongest for F_{med} and F_{mean}. From our study, it turns out that F_{med} and F_{mean} are more valid and stable parameters to measure fatigue in swimming, while DI is more sensitive.

Keywords: electromyography, spectral parameters, fatigue, swimming, master swimmers, video analysis

Edited by:
Valentina Camomilla, Università degli Studi di Roma Foro Italico, Italy

Reviewed by:
Igor Štirn, University of Ljubljana, Slovenia
Kerstin Witte, Otto von Guericke University Magdeburg, Germany

***Correspondence:**
Luca Puce
luca1puce@gmail.com

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: *21 December 2020*
Accepted: *25 January 2021*
Published: *19 February 2021*

Citation:
Puce L, Pallecchi I, Marinelli L, Mori L, Bove M, Diotti D, Ruggeri P, Faelli E, Cotellessa F and Trompetto C (2021) Surface Electromyography Spectral Parameters for the Study of Muscle Fatigue in Swimming. Front. Sports Act. Living 3:644765. doi: 10.3389/fspor.2021.644765

INTRODUCTION

Fatigue has been defined as "*a reduction in force output that occurs during sustained voluntary activity*" (Bigland-Ritchie et al., 1983), and more recently as "*any exercise-induced loss of ability to produce force with a muscle or muscle group*" (Taylor et al., 2006). The phenomenon of fatigue is a common experience in sports, particularly complex as it varies with the change in the type of exercise performed. In particular swimming is a dynamic task and it requires coordinated activation of lower limbs, core, and upper body muscles in each stroke cycle. In addition, the water environment does not offer a fixed fulcrum to exert a maximal force, indeed muscle force at each pulling stroke is only about 50% of the maximal voluntary contraction (Stirn et al., 2011). The decay of velocity and the variations of the kinematic parameters are widely used methods to monitor fatigue in swimming (Stirn et al., 2011; Ikuta et al., 2012; Figueiredo et al., 2013; Conceição et al., 2014; Puce et al., 2018). Although relatively simple to determine, these methods are neither direct nor muscle-specific measurements of fatigue. Real-time monitoring of localized muscle fatigue during the execution of a task is possible through surface electromyography (EMG). During a prolonged muscle contraction, as consequence of the physiological mechanisms of fatigue, the spectral weight of the EMG shifts from high to low frequencies (Dimitrov et al., 2006; González-Izal et al., 2012). For this reason, the time evolution of power spectrum parameters such as the median frequency (F_{med}), the mean frequency (F_{mean}), the Dimitrov index (DI) (Dimitrov et al., 2006) and the mean instant frequency (F_{mi}) could be used to detect the electrophysiological signs of localized muscle fatigue in swimming. Two studies monitored the variation of F_{mean} (Stirn et al., 2011; Conceição et al., 2014). Stirn et al. reported that at the end of a 100 m front crawl, F_{mean} decreased significantly by 20–25% in the upper body muscles. In the study of Conceição et al., despite the evolution of kinematic and physiological parameters reflected the development of muscle fatigue during 200 m breaststroke, only a non-significant trend of F_{mean} decrease in the upper body muscles was reported. Front crawl and breaststroke are technically different in terms of functional involvement of muscles, and consequently fatigue. The upper body muscles, monitored in both studies, are more active in front crawl than in breaststroke, as in breaststroke most of the propulsion is provided by lower body muscles. For this reason, a direct comparison of these apparently conflicting results is not possible.

In order to investigate the evolution of fatigue in a 200-m front crawl, Figueiredo et al. (2013) used DI, which is thought to be more sensitive than F_{med} and F_{mean} as a measure of fatigue in sub-maximal contractions (Dimitrov et al., 2006; González-Izal et al., 2012). The results of this work showed a significant increase of DI by 40–60% for upper limb muscles, but no significant variation for those in lower limbs. Spectral analyses of EMG signal that rely on the Fourier transform (F_{med}, F_{mean} and DI), are based on the assumption that the signal is stationary during the analyzed 0.5–2.0 s intervals (Cifrek et al., 2009). This may not be the case for EMG signals associated to dynamic contractions, especially involving fast movements, where myoelectric signal bursts are often shorter than 500 ms (Bonato et al., 1996). On the other hand, alternative analysis models have been used for non-stationary EMG signals, such as the short-time Fourier transformation (STFT), where the FFT is applied to short overlapping stationary intervals, to the detriment of frequency resolution, the autoregressive or autoregressive–moving-average methods (Witte et al., 2006), the Wavelet methods based on intensity analysis, the methods based on transforms that work well for non-stationary and non-linear data. Among other time-frequency distributions that do not require the hypothesis of stationarity of the EMG signal, are the Cohen class time-frequency transforms (Cifrek et al., 2009), which may thus be more suitable for the spectral analysis in the case of dynamic contractions (Bonato et al., 1996, 2001; González-Izal et al., 2010). Caty et al. (2006) calculated F_{mi} using the Choi-Williams transform, belonging to the Cohen's class transforms, in a 4 × 50 m front crawl and observed a decrease in F_{mi} for the *extensor carpi ulnaris* and *flexor carpi ulnaris* muscles by 11 and 9%, respectively. These variations of F_{mi} were sizably smaller than the variations of spectral parameters based on the Fourier transform (Stirn et al., 2011; Figueiredo et al., 2013; Conceição et al., 2014). From the above mentioned reports, it appears that further studies are necessary to gain insight on the validity of the methods of spectral analysis of EMG signal, for the assessment of localized muscle fatigue in swimming. Yet, valid, stable and sensitive methods to measure fatigue could be useful to assess the level of performance, prevent injuries (Matthews et al., 2017) and adjust training methods (Puce et al., 2018). In this work, we present the evolution of the 4 spectral parameters F_{med}, F_{mean}, DI and F_{mi} during a 100 m front crawl and their correlation with the variations of the kinematic data and the peak torque. Our final aim is to assess validity, stability, and sensitivity of each spectral parameter in measuring the localized muscle fatigue in swimming.

METHODS

Subjects

Fifteen elite masters swimmers (two women; mean ± standard deviation age 33.0 ± 9.7 years; weight 71.9 ± 9.5; height 177.8 ± 8.6 cm) took part in the research study, after 10 days of tapering phase. The swimmers were front crawl specialists, even if some of them were not sprinters. Their average technical index was considered high (612 ± 43) (Santos et al., 2020).

The study was carried out in accordance with the code of ethics of the World Medical Association (Declaration of Helsinki 2014) for experiments involving humans. A written informed consent was obtained from all participants prior to participation in the study. The project was approved by the local ethics committee (University of Genova, Italy. N. 2020/21).

Study Design

Electrophysiological measures of muscle fatigue and kinematic data were obtained in real-time during a 100 m front crawl, using EMG and video analysis. Mechanical fatigue was measured as the difference between muscle strength prior to and immediately after the 100 m front crawl in a dry-land multi-stage Isometric Contraction Test (MICT). After the MICT performed at rest (pre-MICT), a 30 min time was used for recovery and for application of EMG electrodes and adhesive markers. Then the 100 m front crawl Swimming Fatigue Test (SFT) was carried out. The second MICT (post-MICT) was performed immediately after the SFT. It must be stressed that the time elapsed between the end of the SFT and the post-MICT was kept at minimum ($<$ 10 s), in such a way that the level of muscular strength measured in the post-MICT was representative of the fatigue experienced in the SFT.

The outcome measures of this study were the variations of the four spectral parameters of the EMG signal, F_{med}, F_{mean}, DI and F_{mi}, measured during the SFT, the variations of the peak torque measured before and after the SFT and the variations of the kinematics data, velocity, stroke length, and stroke index, measured during the SFT.

Validity of the spectral parameters was determined on the basis of the correlation of electrophysiological signs of fatigue with mechanical fatigue and kinematic parameters. Stability of the spectral parameters was estimated in terms standard deviation of data within individual 100 m SFTs. Sensitivity of the spectral parameter was determined in terms of range of variation in the SFTs.

Swimming Fatigue Test (SFT)

Measurements were performed in a 50-m indoor swimming pool. After the pre-MICT and 30 min recovery, the swimmers performed an individual warm-up. After the warm-up, the swimmers were instructed to perform a 100 m front crawl at the highest level of self-perceived exertion. Due to the measuring equipment attached to the body, the underwater turn was allowed but the dive start was not.

Multi-Stage Isometric Contraction Test (MICT)

MICT had a total duration of 43 s and was carried out on the pool deck. It consisted of six isometric contractions lasting 3 s interspersed with 5 s (change of exercise), each one involving different muscles, carried out using a cable cross over apparatus (model Technogym Cable Stations Ercolina Rehab, Cesena, Italy) with a load cell connected to a Digital Force Indicator display (model C2S-AMP, Modena, Italy). For the pre-MICT and post-MICT tests to be equivalent, participants were asked to express strength at their maximum in the contractions. It must be pointed out that these cannot be considered as maximum voluntary contractions (MVC), which are commonly used in fatigue experiments, due to the short duration (3 s) of the effort. Yet, this duration was chosen to minimize the difference in recovery times among the six successive contractions in the post-MICT tests, as well as to avoid an overall fatigued state that developed in the participants through the MICT tests, in case of contraction durations of 5 s or longer. The maximum value that was maintained on the display for a time at least of the order of 1 s was considered as the peak force value. This force value was converted into a torque (Nm) for each exercise and the difference between pre-MICT and post-MICT torques was normalized to the pre-MICT torque value. The swimmer was harnessed to a chair and changed their body position for each contraction in such a way as to be biomechanically able to recruit a specific muscle (**Figure 1**). The first muscle used was the *pectoralis major* (PM), followed by the *triceps lateralis* (TL), *latissimus dorsi* (LD), *anterior deltoid* (AD), *biceps femoris* (BF) and *rectus femoris* (RF). The succession was structured to optimize the timing. The athletes carried out a pre-test warm-up and a month-long training to learn to compensate as little as possible with other muscles, keep a correct standard position, stay within the contraction/exercise change times, and express the maximum strength during the test.

EMG Data Collection

EMG signals from PM, TL, LD, AD, BF, and RF of the dominant side were measured through bipolar surface electrodes using a waterproof wireless EMG equipment (Cometa srl, Milan, Italy) operating at 2,000 Hz, according to SENIAM guidelines (Hermens et al., 2000). To avoid alterations induced by underwater recording, a water resistant adhesive tape over the electrodes was applied (Rainoldi et al., 2004). The above six muscles were selected according to their relevance in front crawl (Clarys, 1983) and were the same as those assessed in the MICT. Data analysis was performed using the open source software Python distributed by Anaconda Inc. In each EMG trace, the activation interval $[t_{in}, t_{fin}]$ of each stroke was identified where the envelope of the rectified signal around the maximum amplitude exceeded 20% of the maximum amplitude itself, following the same criterion of Stirn et al. (2011). Once the starting and ending times of the activation interval of each stroke were identified this way, spectral analysis was carried out on each activation interval. The EMG intervals were filtered with a band-pass Butterworth filter of 4-th order in the range of 20–500 Hz and then analyzed in the frequency domain. The four spectral parameters F_{med}, Fmean, DI and Fmi were calculated.

F_{med} of the power spectrum was calculated as the frequency that divides the power spectrum into two parts having the same spectral weight, according to the following equations:

$$\int_{f_{c1}}^{F_{med}} PSD(f)\,df = \int_{F_{med}}^{f_{c2}} PSD(f)\,df = \frac{1}{2}\int_{f_{c1}}^{f_{c2}} PSD(f)\,df \quad (1)$$

where $f_{c1} = 20$ Hz and $f_{c2} = 500$ Hz are the cut-off frequencies of the high-pass and low-pass filters applied to the spectra and PSD(f) is the power spectral density.

F_{mean} of the power spectrum was calculated as the momentum of order 1 of the power spectrum:

$$Fmean = \frac{\int_{f_{c1}}^{f_{c2}} f \cdot PSD(f)\,df}{\int_{f_{c1}}^{f_{c2}} PSD(f)\,df} \quad (2)$$

DI was calculated as the ratio of the momentum of order −1 of the power spectrum to the momentum of order 5:

$$DI = \frac{\int_{f_{c1}}^{f_{c2}} \frac{1}{f} \cdot PSD(f)\,df}{\int_{f_{c1}}^{f_{c2}} f^5 \cdot PSD(f)\,df} \quad (3)$$

The instant frequency IF(t) was calculated using the Choi–Williams time-frequency distribution, with intermediate value of the kernel parameter (O'Toole and Boashash, 2013). F_{mi} was then calculated by averaging IF(t) over each activation interval between times t_{in} and t_{fin}:

$$IF(t) = \frac{\int_{f_{c1}}^{f_{c2}} f \cdot \left|CW(t,f)\right|^2 df}{\int_{f_{c1}}^{f_{c2}} \left|CW(t,f)\right|^2 df}, \quad Fmi = \frac{\int_{t_{in}}^{t_{fin}} IF(t)\,dt}{t_{fin} - t_{in}} \quad (4)$$

Here $|CW(t,f)|^2$ are the squared time-frequency components of the Choi-Williams transform.

Electrophysiological detection of muscle fatigue were obtained as time variation of spectral index, quantified by the slope of the linear regressions of spectral index vs. time, with uncertainty given by the standard deviation of the linear regression. A negative slope for F_{med}, F_{mean}, F_{mi} and positive slope for DI indicates fatigue. The slopes were finally normalized

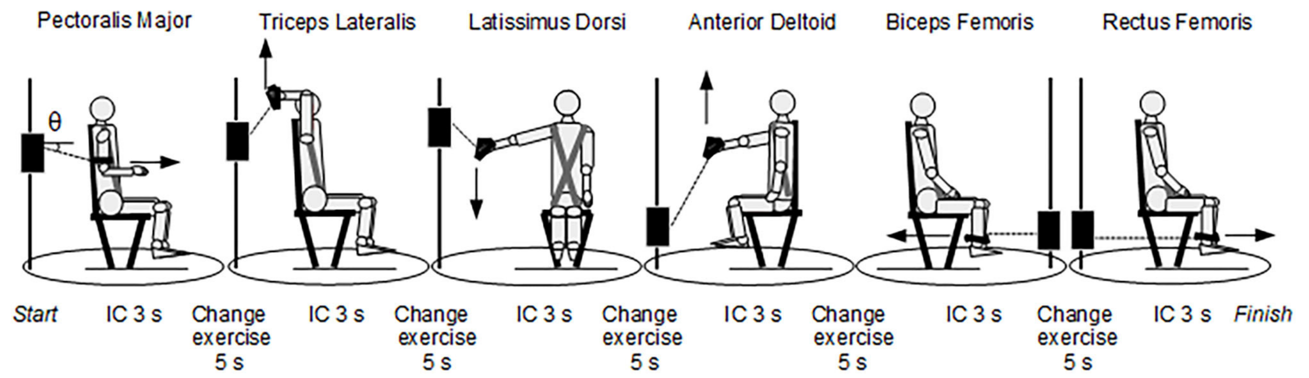

FIGURE 1 | Multi-stage isometric contraction test (MICT). In the left-most position, θ indicated the angle between the cable and the horizontal axis. The joint and cable angles in the six positions are: Pectoralis Major: shoulder flexed in neutral position (0°) and elbow flexed by 90°, $\theta = -15°$; Triceps Lateralis: shoulder flexed by 180° and elbow flexed by 90°, $\theta = +55°$; Latissimus Dorsi: shoulder abducted by 75 °, $\theta = -45°$; Anterior Deltoid: shoulder flexed by 75°, $\theta = +60°$; Biceps Femoris: knee flexed by 90°, $\theta = 0°$; Rectus Femoris: knee flexed by 90°, $\theta = 0°$. IC = isometric contraction.

to the initial value of the regression line for each EMG trace. Average values of normalized slopes of each spectral parameter for each muscle were calculated over the 15 participants. The average of each spectral parameter x over the participants was calculated by weighting each value with the inverse variance $\frac{1}{\sigma_i^2}$ obtained from the linear regression:

$$x_{average} = \frac{\sum_{i=1}^{15} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{15} \frac{1}{\sigma_i^2}} \tag{5}$$

The error bars on these averages were calculated as:

$$\sigma_x = \sqrt{\frac{1}{\sum_{i=1}^{15} \frac{1}{\sigma_i^2}}} \tag{6}$$

Kinematic Data Collection

The measurements of kinematic data were carried out by analyzing video recordings (Kinovea 0.8.25), acquired on sagittal plane using two cameras (model GoPro Hero 8, GoPro, San Mateo, CA, USA) one above the water surface and one below. The cameras were fixed to a pushcart which was moved at the same speed as the swimmer speed. Precise information on the absolute position of body and limbs was obtained by applying adhesive markers on the joints of the lower and upper limbs and synchronizing the biomechanical analysis with the EMG signal. Specifically, triggering of video recording and EMG signal was done by tapping a spare EMG probe at the start. Swimming velocity (SV), stroke length (SL), and stroke index (SI) were evaluated. SI was calculated as the product of SV and SL, and it was used as an index of the swimming efficiency (Costill et al., 1985). The SV was calculated as the ratio of space swum to chronometric time. The SL was calculated as the ratio of space swum to the corresponding number of strokes. To avoid the influence of the start and turn phases, all three parameters were calculated in the free-swimming segment, that is between 15th and 45th m of the pool length. Finally, the variations of SV, SL and SI were evaluated as the linear time derivatives $\left(\frac{dSI}{dt}, \frac{dSV}{dt}, \frac{dSL}{dt}\right)$ in the free swimming segments and normalized to the respective values at the instant t_0 corresponding to the 15th m of the first length, $SI(t = t_0), SV(t = t_0), SL(t = t_0)$.

Statistical Analysis

Correlation between electrophysiological signs of fatigue (normalized slopes of EMG spectral index) and mechanical fatigue (normalized difference in torque between pre-MICT and post-MICT) and between electrophysiological signs of fatigue and decay of kinematic parameters (normalized slopes of kinematic parameters) was evaluated by the Pearson coefficient r. It was assumed that correlation was *strong* for $0.5 \le |r| \le 1$, *moderate* for $0.3 \le |r| \le 0.49$, *low* for $|r| \le 0.29$, *null* for $|r| = 0$. One-way ANOVA test was used to evaluate the significance of differences of electrophysiological signs of fatigue between different muscles. Statistical significance was also evaluated in terms of *p*-value for the average normalized slope of each spectral parameter for each muscle and for the average variation of kinematic parameters. Statistical significance was set at $p < 0.05$.

RESULTS

Electrophysiological Signs of Fatigue

In **Figure 2**, we present a typical EMG acquisition during the underwater phase of a stroke, where the activation intervals of the 6 muscles and their relative shift are shown. The upper body muscles had one activation interval per stroke, while lower limbs have either two or three activation intervals per stroke, depending on the participant. PM has the longest activation interval.

F_{med}, F_{mean}, and F_{mi} for all the participants and muscles exhibited a decreasing trend over time while DI exhibited an increasing trend. **Figure 3** shows recordings of a representative set of these parameters.

The average normalized slopes of each parameter and each muscle over the 15 participants are shown in **Figure 4**. All these values were statistically significant ($p < 0.0001$).

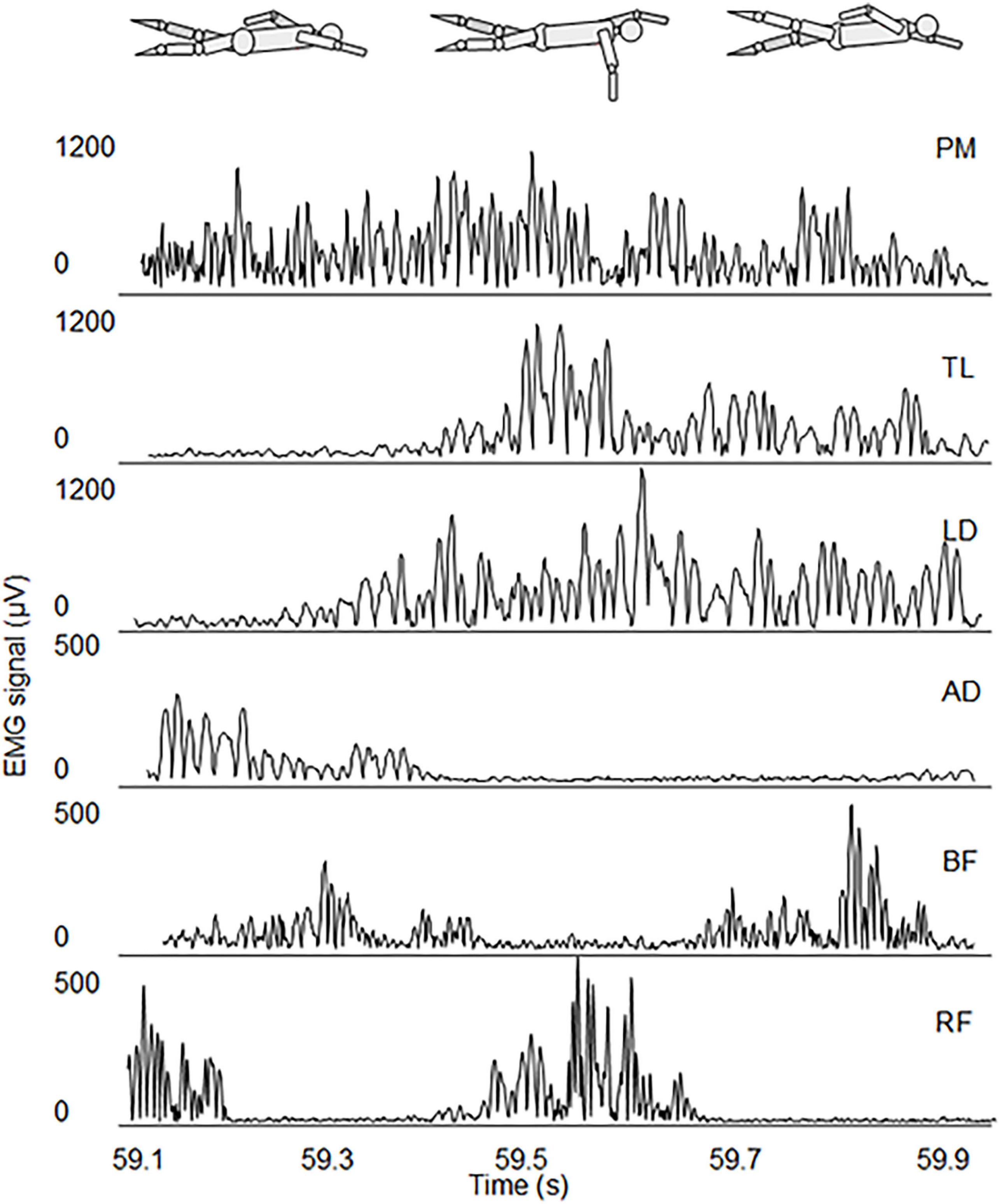

FIGURE 2 | EMG activation of the underwater phase stroke of the 6 muscles of one representative participant during the SFT.

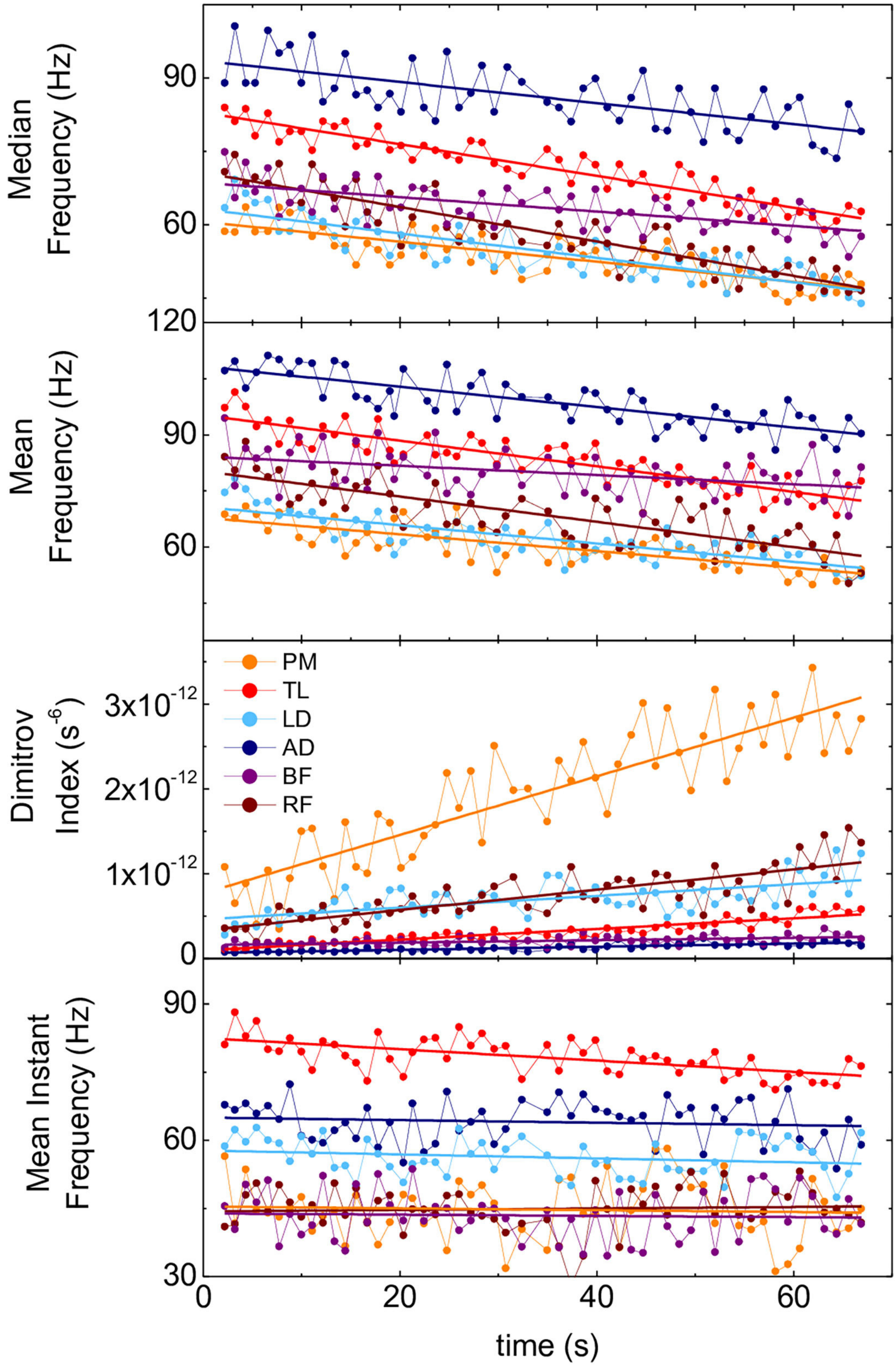

FIGURE 3 | Time evolution of spectral parameters of the 6 muscles of one representative participant during the SFT. Panels from top to bottom: F_{med}, F_{mean}, DI, F_{mi}. PM, pectoralis major; TL, triceps lateralis; LD, latissimus dorsi; AD, anterior deltoid; BF, biceps femoris; RF, rectus femoris.

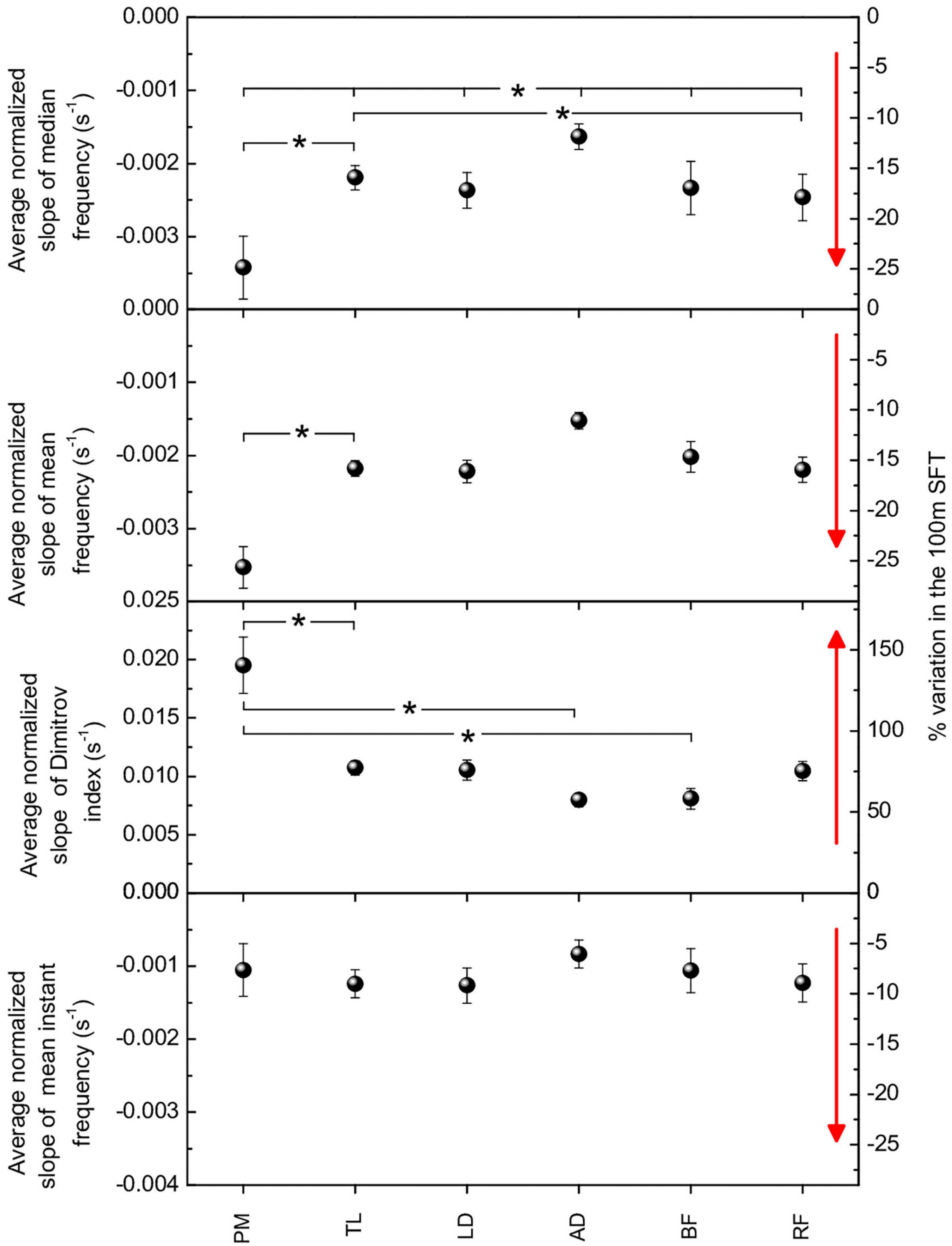

FIGURE 4 | Weighted averages of the normalized slopes of spectral parameters for the six muscles. Panels from top to bottom: F_{med}, F_{mean}, DI, F_{mi}. All these values are statistical significant ($p < 0.0001$). Right-hand axes indicate the percent variation in the FST. Asterisks indicate statistical significance ($p < 0.05$) of difference between either two or all six muscles. Red arrows point to the direction of increasing fatigue for each spectral parameter.

F_{med} and F_{mean} varied between 10 and 25%, DI between 50 and 150%, and F_{mi} between 5 and 10%. Regarding scattering of data for each parameter in individual SFTs (see sets of data in **Figure 3** as an example), in absolute terms, the standard deviation of DI was on average seven times larger than the standard deviation of F_{med}, 13 times larger than the standard deviation of F_{mean}, and nine times larger than the standard deviation of F_{mi}. In relative terms, normalizing to average values, the standard deviation of DI was still the largest (∼70%), as compared to ∼60% for F_{med} and F_{mi} and ∼35% for F_{mean}.

For all spectral parameters, larger relative changes among upper body muscles was observed in PM and LD, while the least relative change was observed in AD. In the lower limb muscles, the relative change was comparable for BF and RF, except in the case of DI, which showed larger variation for RF.

Regarding the differences between muscles, statistical significance ($p < 0.05$) was observed for the PM-TL and TL-RF couples and within the group of all the 6 muscles for F_{med}; only for the PM-TL couple for F_{mean}; for the PM-TL, PM-AD and PM-BF couples for DI.

Mechanical Fatigue

Mechanical fatigue, evaluated as normalized difference between post and pre MICT varied between 8% for RF and 17% for LD, as shown in **Figure 5**. Due to the scattering of data, statistical significance was found only for TL and LD ($p < 0.05$).

Kinematic Data

The average swimming time achieved by the swimmers was longer than their personal best by as little as 3.2%. Hence, considering the fact that in SFT the participants did not perform the dive start, did not wear a racing suit and had to cope with the burden of the equipment, the performance can be considered performed at maximum effort. Moreover, for all participants a progressive decrease of SV, SL, and SI was observed. **Figure 6** presents the normalized slope of the three kinematic parameters across the FST. SV, SL, and SI exhibited variations by 15, 9, and 22%, respectively. Statistical significance applied to SV and SI ($p < 0.05$).

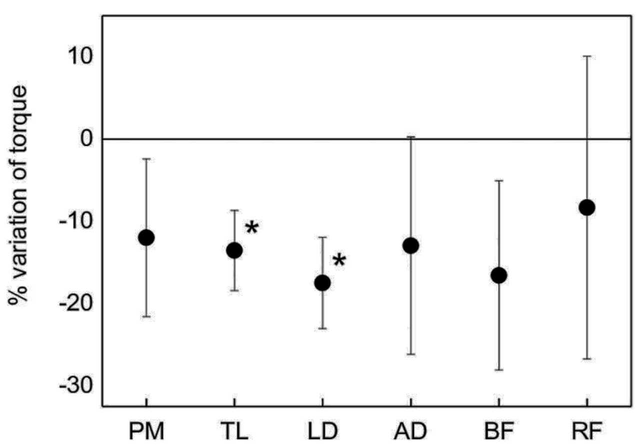

FIGURE 5 | Average percent variation of the torque for each muscle from the pre-MICT to the post-MICT, representing mechanical fatigue. Asterisks indicate statistical significance ($p < 0.05$).

Correlation

For increasing fatigue F_{med}, F_{mean}, and F_{mi} should decrease and DI increase, and torque should decrease as well. In these conditions, all the characteristic spectral frequencies (F_{med}, F_{mean}, and F_{mi}) correlate positively with the changes in torque, while DI correlates negatively with the changes in torque. Indeed, this is just what comes out from the Pearson coefficients r, reported in **Table 1**, used to evaluate the correlation between electrophysiological signs of fatigue, assessed through different spectral parameters, and mechanical fatigue.

The negative variations of F_{med} and F_{mean} exhibited *strong* positive correlation with mechanical fatigue for all the muscles, except RF. Statistical significance of these correlations was found for TD, LD, AD, and BF. On the contrary, the negative variations

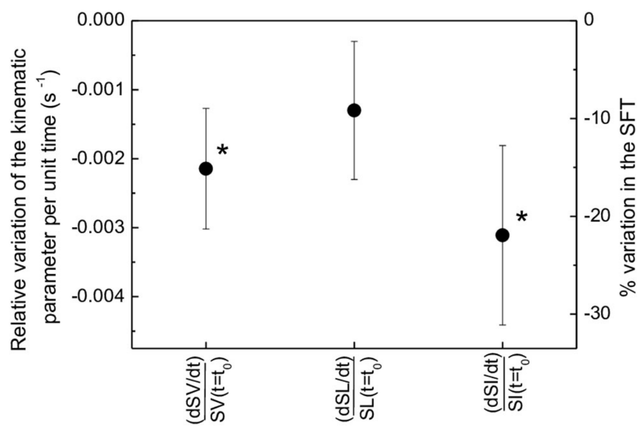

FIGURE 6 | Average relative variation of the kinematic parameters (SV, SL, SI) per unit time in the FST, expressed as linear time derivative of each kinematic parameter ($\frac{dSI}{dt}$, $\frac{dSV}{dt}$, $\frac{dSL}{dt}$, respectively) normalized to the value of the kinematic parameter at the beginning of the free swimming segment [$SI(t = t_0), SV(t = t_0), SL(t = t_0)$, respectively]. Asterisks indicate statistical significance ($p < 0.05$).

TABLE 1 | Pearson correlation coefficient $r_{Pearson}$ coefficients between normalized variations of spectral parameters (electrophysiological fatigue) and normalized variations of torque (mechanical fatigue).

Muscle	Spectral parameter			
	F_{med}	F_{mean}	DI	F_{mi}
PM	0.61	0.59	−0.42	−0.37
TL	0.88*	0.83*	−0.63*	0.20
LD	0.59*	0.66*	−0.03	−0.60
AD	0.90*	0.95*	−0.53	−0.15
BF	0.75*	0.78*	−0.68*	0.26
RF	−0.26	−0.33	0.57	−0.53

Asterisks indicate statistical significance of the correlation ($p < 0.05$).

F_{mi} exhibited positive correlation with mechanical fatigue only for TL and BF, but this correlation was *low* and not significant.

The positive variations of DI exhibited *strong* negative correlation with mechanical fatigue for TD, AD, and BF, with statistical significance for TL and BF, while for PM only moderate *correlation* was found.

Not significant correlation was found between the changes in spectral parameters and the changes in all kinematic parameters SV, SL, and SI.

DISCUSSION

In this study, we presented the evolution of the four spectral parameters during 100 m front crawl and their correlation with the variation of the torque and kinematic data to assess the validity and sensitivity of each spectral parameter that measures fatigue in swimming. From our study, it turned out that F_{med} and F_{mean} are more stable and valid parameters to measure fatigue in swimming, while DI is more sensitive.

Electrophysiological Signs of Fatigue

The relative difference in fatigue between different muscle is qualitatively similar for all the spectral parameters, in agreement with results of Dimitrov studies (Dimitrova et al., 2005; Dimitrov et al., 2006, 2008). The widest range of variation was observed for DI revealing about six times larger sensitivity to fatigue than F_{med} and F_{mean} and 15 times larger than F_{mi}. In other studies, DI was found to be up to 150 times larger than F_{med} and F_{mean} during electrically evoked contractions (Dimitrova et al., 2005) and 50 times larger during voluntary isometric contractions (Dimitrov et al., 2008).

Larger sensitivity of DI is due to definition as the ratio of momentum of order −1 of the power spectrum to the momentum of order five, which better describes the shift of spectral weight from high to low frequencies with increasing fatigue, due to different mechanisms. The spectral moment of order (−1) emphasizes the increase in low and ultralow frequencies in the EMG spectrum due to increased negative after-potentials during fatigue. The spectral moment of order five emphasizes the effect of decreases in high frequencies, due to increments in the duration of the intracellular action potentials and decrements in the action potential propagation velocity (Dimitrov et al., 2006). As a counterpart of such larger sensitivity, DI showed the largest standard deviations of the data within individual 100 m SFTs, likely due to numerical instability, originating from the stochastic nature of the EMG signal and the DI definition itself in terms of higher-order momenta of the power spectrum.

Comparing fatigue in different muscles, the largest fatigue was observed for PM. Indeed, in studies that analyse the EMG amplitude normalized of MVC (Clarys, 1983; Pink et al., 1991), PM was observed to produce the most propulsive force in front crawl, together with LD and TB. Moreover, from our results, the duration of its EMG activity in the front crawl stroke was the longest among the investigated muscles, lasting throughout the whole underwater phase (**Figure 2**). This may be due to the fact that PM, in synergy with other respiratory muscles, is also responsible for inspiration. LD and TB showed large fatigue, as well. LD is known as the "*swimmers muscle*" due to the major role it plays in the successful completion of each of the swim styles (Laudner and Williams, 2013) and, together with TB, it is considered the key muscle in maintaining swimming speed in fatigued conditions (Ikuta et al., 2012). AD was the least fatigued upper body muscle. The contribution of this muscle to propulsion is limited (**Figure 2**), and its main function is bringing the shoulder over the head during the recovery phase of the stroke (Pink et al., 1991). Although lower limb muscles only contribute about 15% to propulsion in front crawl (Stirn et al., 2011), fatigue in BB and RF was similar to upper body muscles. In front crawl the biomechanics of lower limbs do not rely on a good support surface, as in the case of breaststroke, however they have twice or three times as many activations intervals as the upper body muscles in each stroke (**Figure 2**).

Correlation

The electrophysiological manifestation of fatigue obtained as change in spectral parameters over time is associated with phenomena that occur in the muscle, prior to occurrence of mechanical fatigue. Indeed, the variations of these parameters reflect physiological phenomena that will only subsequently degrade the mechanical performance of the muscle. Therefore, it is interesting to inspect the correlation between the parameters that reflect electrophysiological signs of fatigue and parameters that identify mechanical fatigue.

An important finding of the present work was that F_{med} and F_{mean} showed the highest correlation between electrophysiological signs of fatigue and mechanical fatigue. Hence, these spectral parameters proved to be the most valid in dynamic contractions, whose bursts are often shorter than 500 ms (**Figure 2**), a time interval in which the problem of possible non-stationarity of the EMG signal may arise.

Regarding F_{mi} it exhibited the least correlation between electrophysiological and mechanical fatigue and the least sensitivity to fatigue in different muscles. It must be noted that we observed a negligible dependence of results on the choice of the Choi–Williams parameter over a wide range of values. This parameter suppresses the cross-terms in the frequency spectrum, i.e., terms originating from the product of different frequency components, which appear as a consequence of non-stationarity of the signal. The negligible influence of this parameter on the results indicates that the problem of non-stationarity is not relevant in our signal. In this situation, the use of the Choi-Williams distribution in place of the Fourier transform is not appropriate, as it causes a smoothening of the frequency spectrum, thus also altering the frequency components of the signal and not just the mixed terms.

We remark that the correlation between electrophysiological and mechanical fatigue observed in our study was not obvious a priori. Indeed, dynamic and static contractions have different patterns of neural activations (Cheng and Rice, 2005).

The kinematic parameters were not correlated with electrophysiological measures of fatigue, represented by any of the spectral parameters. This result can be explained by the unaware tendency of swimmers to maintain constant velocity in fatigued conditions, through modification of arm coordination (Cheng and Rice, 2005) and muscle activation (Stirn et al., 2011; Ikuta et al., 2012). This unaware compensation strategy may be a further reason why fatigue occurs differently for different muscles.

PRACTICAL IMPLICATIONS

This study provides information on the use of the most appropriate spectral parameters in terms of validity, stability and sensitivity for the assessment of fatigue in swimming. Our results show that F_{med} and F_{mean} are the most valid and stable parameters and are thus recommended, particularly in tests where maximum effort is required. DI is the most sensitive, and it may be more suitable in tests where low intensity muscle contractions are required, although it is intrinsically more liable to numerical instability, due to the stochastic nature of the EMG signal.

LIMITATIONS OF THIS STUDY

We point out two limitations of this study. First, the positions of the MICT were optimized in neither singling out the effort of a specific muscle, nor in simulating the in-water stroke movement. However, the MICT provided a good measure of mechanical fatigue as long as it was performed in identical conditions prior to and after the swimming fatiguing test. A second limitation regards the measurement of the mechanical force, which was carried out by visual inspection of the dynamometer display, rather than by recording and mathematic averaging the output signal of this instrument. However, attention was paid to ensure reproducibility of the method.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Local Ethics Committee (University of Genova, Italy. N. 2020/21). The participants provided their written informed consent to participate in this study. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LP ideated and participated at the experiment and wrote the manuscript. CT ideated the experiment and wrote the manuscript. FC, DD, and LMa participated in the experiment and in the literature search. EF performed the EMG experiment and in the literature search. PR performed the EMG experiment and collaborated in revising the manuscript. MB ideated the experiment and collaborated in revising the paper. LMo and IP analyzed the data and collaborated in revising the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

Bigland-Ritchie, B., Johansson, R., Lippold, O. C., and Woods, J. J. (1983). Contractile speed and EMG changes during fatigue of sustained maximal voluntary contractions. *J. Neurophysiol.* 50, 313–324. doi: 10.1152/jn.1983.50.1.313

Bonato, P., Gagliati, G., and Knaflitz, M. (1996). "Analysis of myoelectric signals recorded during dynamic contractions," in *IEEE Engineering in Medicine and Biology Magazine*. 15 (New York, NY: IEEE publisher), 102–111. doi: 10.1109/51.544517

Bonato, P., Roy, S. H., Knaflitz, M., and De Luca, C. J. (2001). Time-frequency parameters of the surface myoelectric signal for assessing muscle fatigue during cyclic dynamic contractions. *IEEE Trans. Biomed. Eng.* 48, 745–753. doi: 10.1109/10.930899

Caty, V. Y., Rouard, A. H., Hintzy, F., Aujouannet, Y. A., Molinari, F., and Knaflitz, M. (2006). Time–frequency parameters of wrist muscles EMG after an exhaustive freestyle test. *Revista Portuguesa de Ciencias do Desporto* 6, 28–30.

Cheng, A. J., and Rice, C. L. (2005). Fatigue and recovery of power and isometric torque following isotonic knee extensions. *J. Appl. Physiol.* (1985) 99, 1446–1452. doi: 10.1152/japplphysiol.00452.2005

Cifrek, M., Medved, V., Tonković, S., and Ostojić, S. (2009). Surface EMG based muscle fatigue evaluation in biomechanics. *Clin.. Biomech.* 24, 327–340. doi: 10.1016/j.clinbiomech.2009.01.010

Clarys, J. P. (1983). "A review of EMG in swimming: explanation of facts and/or feedback information," in *Biomechanics and Medicine in Swimming: Proceedings of the Fourth International Symposium of Biomechanics in Swimming and the Fifth International Congress on Swimming Medicine*, eds A. P. Hollander (Amsterdam, Champaign, IL: Human Kinetics Publishers), 123–135.

Conceição, A., Silva, A. J., Barbosa, T., Karsai, I., and Louro, H. (2014). Neuromuscular fatigue during 200 m breaststroke. *J. Sports Sci. Med.* 20, 200–210.

Costill, D. L., Kovaleski, J., Porter, D., Kirwan, J., Fielding, R., and King, D. (1985). Energy expenditure during front crawl swimming: predicting success in middle-distance events. *Int. J. Sports Med.* 6, 266–270. doi: 10.1055/s-2008-1025849

Dimitrov, G. V., Arabadzhiev, T. I., Hogrel, J. Y., and Dimitrova, N. A. (2008). Simulation analysis of interference EMG during fatiguing voluntary contractions. Part II–changes in amplitude and spectral characteristics. *J. Electromyogr. Kinesiol.* 18, 35–43. doi: 10.1016/j.jelekin.2006.07.002

Dimitrov, G. V., Arabadzhiev, T. I., Mileva, K. N., Bowtell, J. L., Crichton, N., and Dimitrova, N. A. (2006). Muscle fatigue during dynamic contractions assessed by new spectral indices. *Med. Sci. Sports Exerc.* 38, 1971–1979. doi: 10.1249/01.mss.0000233794.31659.6d

Dimitrova, N. A., Hogrel, J. Y., Arabadzhiev, T. I., and Dimitrov, G. V. (2005). Estimate of M-wave changes in human biceps brachii during continuous stimulation. *J. Electromyogr. Kinesiol.* 15, 341–348. doi: 10.1016/j.jelekin.2005.01.005

Figueiredo, P., Rouard, A., Vilas-Boas, J. P., and Fernandes, R. J. (2013). Upper- and lower-limb muscular fatigue during the 200-m front crawl. *Appl. Physiol. Nutr. Metab.* 38, 716–724. doi: 10.1139/apnm-2012-0263

González-Izal, M., Malanda, A., Gorostiaga, E., and Izquierdo, M. (2012). Electromyographic models to assess muscle fatigue. *J. Electromyogr. Kinesiol.* 22, 501–512. doi: 10.1016/j.jelekin.2012.02.019

González-Izal, M., Malanda, A., Navarro-Amézqueta, I., Gorostiaga, E. M., Mallor, F., Ibañez, J., et al. (2010). EMG spectral indices and muscle power fatigue during dynamic contractions. *J. Electromyogr. Kinesiol*. 20, 233–240. doi: 10.1016/j.jelekin.2009.03.011

Hermens, H. J., Freriks, B., Disselhorst-Klug, C., and Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. *J. Electromyogr. Kinesiol*. 10, 361–374. doi: 10.1016/S1050-6411(00)00027-4

Ikuta, Y., Matsuda, Y., Yamada, Y., Kida, N., Oda, S., and Moritani, T. (2012). Relationship between decreased swimming velocity and muscle activity during 200-m front crawl. *Eur. J. Appl. Physiol*. 112, 3417–3429. doi: 10.1007/s00421-012-2321-0

Laudner, K. G., and Williams, J. G. (2013). The relationship between latissimus dorsi stiffness and altered scapular kinematics among asymptomatic collegiate swimmers. *Phys. Ther. Sport*. 14, 50–53. doi: 10.1016/j.ptsp.2012.02.007

Matthews, M. J., Green, D., Matthews, H., and Swanwick, E. (2017). The effects of swimming fatigue on shoulder strength, range of motion, joint control, and performance in swimmers. *Phys. Ther. Sport*. 23, 118–122. doi: 10.1016/j.ptsp,.2016.08.011

O'Toole, J. M., and Boashash, B. (2013). Fast and memory-efficient algorithms for computing quadratic time–frequency distributions. *Appl. Comput. Harmon. Anal*. 35, 350–358. doi: 10.1016/j.acha.2013.01.003

Pink, M., Perry, J., Browne, A., Scovazzo, M. L., and Kerrigan, J. (1991). The normal shoulder during freestyle swimming. An electromyographic and cinematographic analysis of twelve muscles. *Am. J. Sports Med*. 19, 569–576. doi: 10.1177/036354659101900603

Puce, L., Marinelli, L., Pierantozzi, E., Mori, L., Pallecchi, I., Bonifazi, M., et al. (2018). Training methods and analysis of races of a top level Paralympic swimming athlete. *J. Exerc. Rehabil*. 14, 612–620. doi: 10.12965/jer.1836254.127

Rainoldi, A., Cescon, C., Bottin, A., Casale, R., and Caruso, I. (2004). Surface EMG alterations induced by underwater recording. *J. Electromyogr. Kinesiol*. 14, 325–331. doi: 10.1016/j.jelekin.2003.10.002

Santos, K. B., Bento, P. C. B., Payton, C., and Rodacki, A. L. F. (2020). Symmetry in the front crawl stroke of different skill level of able-bodied and disabled swimmers. *PLoS ONE* 15:e0229918. doi: 10.1371/journal.pone.0229918

Stirn, I., Jarm, T., Kapus, V., and Strojnik, V. (2011). Evaluation of muscle fatigue during 100-m front crawl. *Eur. J. Appl. Physiol*. 111, 101–113. doi: 10.1007/s00421-010-1624-2

Taylor, J. L., Todd, G., and Gandevia, S. C. (2006). Evidence for a supraspinal contribution to human muscle fatigue. *Clin. Exp. Pharmacol. Physiol*. 33, 400–405 doi: 10.1111/j.1440-1681.2006.04363.x

Witte, K., Heller, M., Ganter, N., Edelmann-Nusser, J., and Schwab, K. (2006). "Time-variant spectral analysis of surface EMG – applications in sports practice," in *Proceedings of XXIV International Symposium on Biomechanics in Sports, Salzburg, Institute of Medical Statistics, Computer Sciences and Documentation* (Konstanz).

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

frontiers in Sports and Active Living

PERSPECTIVE
published: 04 March 2021
doi: 10.3389/fspor.2021.631101

Virtual Training of Endurance Cycling – A Summary of Strengths, Weaknesses, Opportunities and Threats

*Benjamin McIlroy [1,2], Louis Passfield [3,4], Hans-Christer Holmberg [5,6] and Billy Sperlich [1]**

[1] Department of Sport Science, Integrative and Experimental Exercise Science, University of Würzburg, Würzburg, Germany, [2] Department of Sport and Public Services, Brooklands College, Weybridge, United Kingdom, [3] Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada, [4] School of Sport and Exercise Sciences, University of Kent, Canterbury, United Kingdom, [5] Department of Physiology and Pharmacology, Biomedicum C5, Karolinska Institute, Stockholm, Sweden, [6] Department of Engineering Sciences and Mathematics, Luleå University of Technology, Luleå, Sweden

Virtual online training has emerged as one of the top 20 worldwide fitness trends for 2021 and continues to develop rapidly. Although this allows the cycling community to engage in virtual training and competition, critical evaluation of virtual training platforms is limited. Here, we discuss the strengths, weaknesses, opportunities and threats associated with virtual training technology and cycling in an attempt to enhance awareness of such aspects. Strengths include immersive worlds, innovative drafting mechanics, and versatility. Weaknesses include questionable data accuracy, inadequate strength and reliability of power-speed algorithms. Opportunities exist for expanding strategic partnerships with major cycling races, brands, and sponsors and improving user experience with the addition of video capture and "e-coaching." Threats are present in the form of cheating during competition, and a lack of uptake and acceptance by a broader community.

Keywords: algorithms, cycling, e-coach, e-health, ergometer, simulation, virtual training, SWOT

OPEN ACCESS

Edited by:
Pietro Picerno,
University of eCampus, Italy

Reviewed by:
Enzo Iuliano,
eCampus University, Italy
Puneeta Ajmera,
University of Delhi, India

***Correspondence:**
Billy Sperlich
billy.sperlich@uni-wuerzburg.de

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: *19 November 2020*
Accepted: *28 January 2021*
Published: *04 March 2021*

Citation:
McIlroy B, Passfield L, Holmberg H-C and Sperlich B (2021) Virtual Training of Endurance Cycling – A Summary of Strengths, Weaknesses, Opportunities and Threats. Front. Sports Act. Living 3:631101. doi: 10.3389/fspor.2021.631101

INTRODUCTION

In a recent survey, virtual and online training ranked in the top 6 worldwide fitness trends for 2021 (Thompson, 2021). Development of strategic digital live-streaming or pre-recorded sessions of group, individual, or instructional programs allows exercise to be performed at home (Thompson, 2021). This is particularly important at this time, since the national or local lockdowns used to manage the COVID-19 pandemic in many parts of the world, including temporary closure of gyms, has forced many athletes to engage in ergometer training at home.

In the case of cycling, advances in technology have improved indoor training equipment, providing novel simulation trainers equipped with power measuring capability connected online with new 2- or 3D virtual training and competition applications. The virtual environment can be achieved with wearable technology, such as a virtual reality (VR) headset, or through a figure on the screen (normally referred to as an avatar) whose movements the player controls. The most important feature of VR is effective immersion, making the individual feel fully present in the virtual environment (Witmer and Singer, 1998; Radianti et al., 2020).

In the rapidly evolving field of virtual online training, Zwift (https://www.zwift.com) is currently one of the most popular platforms, with more than 2.5 million registered app users in 190 countries (Long, 2020) and an all-time high of more than 30,000 users cycling at the same time (Schlange, 2020b). Other communities such as Peloton (https://www.onepeloton.com/), Real Grand Tours (RGT; https://www.rgtcycling.com/), Rouzy (https://rouvy.com/), and others are growing. The Zwift community continues to expand, with the formation of racing teams, first informally and more recently with support from sponsors, and competition monitored through a third party (Zwiftpower; https://zwiftpower.com/).

Here, we would like to share our experience concerning cycling in virtual reality by elite and amateur athletes worldwide. In this context we summarize the strengths, weaknesses, as well as opportunities and threats of virtual online training platforms (i.e., especially, but not exclusively Zwift) in attempt to enhance awareness of various aspects of virtual training technology and online cycling. This description is also intended to act as a starting point for discussion and planning of future research on this new and rapidly evolving type of sport.

STRENGTHS

Availability at a Wide Range of Costs

With the Zwift online training platform, for example, power, speed, cadence, and heart rate are monitored indoors by specific sensors on a bicycle set up as a static trainer. Most popular are turbo trainers, free-standing rollers, or specially designed indoor bicycles. Turbo trainers vary in their level of technology, starting with wheel-on trainers (i.e., attaching the rear bicycle wheel to a weighted fly wheel), which are usually unpowered and unable to provide data, requiring external devices for this purpose (and therefore nicknamed "dumb" trainers). Top-of-the-line ("smart") trainers are direct-drive (requiring the rear bicycle wheel to be removed and the bicycle chain to be attached directly to the trainer) and are usually electronic and capable of simulating conditions such as incline and changes in the road surface, while monitoring performance data.

To date, a wheel-on trainer plus external speed and cadence monitors costs ∼£150/€165/$195 (not including the bicycle) in addition to the monthly Zwift subscription fee of £12/€13/$15. The cost can rise to more than £2,000/€2,200/$2,500 when purchasing a high-quality trainer, with incline simulation and a secondary power-meter for back-up and verification. Although the latter may seem to be relatively expensive, it is cheaper than other commercially available virtual training platforms utilizing VR headsets (Duking et al., 2018), as well as many of the cycle training camps organized in different parts of the world. This high-end set-up can provide an almost unlimited number of simulated training environments, routes, and races.

Novel Strategies for Team Management

The adaptable nature of virtual cycling platforms allows preparation for many different kinds of competitions, decreasing the need for athletes and coaches to travel to different training venues, thereby avoiding jetlag and fatigue (Fowler et al., 2017), reducing the time lost to periodization and tapering, and costs normally associated with travel (Le Meur et al., 2012). In addition, coaching staff can assist athletes remotely, regardless of location or time zones. Cyclists can train and compete in greater comfort in their own homes and/or other familiar surroundings. Moreover, fewer mechanics will be required as the risk of problems with an indoor trainer is relatively low.

Realistic Simulation of Many Different Racing Situations and Conditions

New routes and training environments are being developed continuously, with the most recent updates encompassing simulations of different stages of the Tour de France, including the world-famous sprint finish along the Champs-Élysées. Currently, more than 70 racing courses are available, ranging from short sprints (<5 km) to endurance courses (longer than 100 km). World-famous climbs, such as the Alpe d'Huez and Mont Ventoux, can be simulated in virtual reality. This versatility allows greater training specificity than is possible with more traditional indoor cycling.

Furthermore, virtual cycling platforms can simulate drafting effects that mimic those experienced outdoors, i.e., a cyclist is able to conserve energy by riding behind another cyclist on-screen. This drafting effect allows for basic adaptations based on differences in the cyclist's height and mass, the weight and model of the bicycle and wheel selection (in the Zwift simulation), the size of riding group, and inclines, even shallow climbs of up to 3 degrees. When cycling downhill, the rider can free-wheel and still maintain speed while in different positions, most notably the "super-tuck" position, an extremely aerodynamic position which every avatar will assume when free-wheeling at or above a certain speed. During actual cycling outdoors on varying terrain, freewheeling is quite common and, accordingly, virtual training platforms can simulate a range of cadences.

In addition, virtual cycling platforms allow simulation of different road surfaces, including tarmac, gravel, and dirt, each with its own resistance and riding experience. Thus, with only one type of bicycle at home, the athlete can train and compete in a greater variety of scenarios or categories than would otherwise be possible.

The most recent innovation is the introduction of steering capabilities via a steering platform fixed to the bicycle at home, which allows for an even greater level of immersion in the virtual environment.

Safety

Globally, rates of traffic-related cycling injuries vary from 174 to 1,329 per 100,000 registered cyclists (Ag, 2019), resulting in significant costs – in the case of minor injuries, averaging 841 € in time lost from work, medical treatment, and costs for replacement of equipment (Aertsens et al., 2010). In addition, fear of, e.g., heavy traffic, darkness and/or bad weather, being attacked by strangers and bicycle theft is often a barrier to engaging in cycling (Heesch et al., 2012).

The ability to participate in simulated races in different disciplines and in large group races without fear of accidents is particularly useful for those recovering from an injury or who are anxious when cycling in groups. Nervous and inexperienced cyclists can also join a race on virtual cycling platforms without having to deal with the potentially intimidating experience of traveling to an outdoor event and negotiating the start of a mass participation event.

Athletes can also conduct high-intensity training sessions without encountering traffic or, e.g., having to stop at traffic lights, allowing training loads to be standardized. In fact, the sense of pressure and urgency that can be created in connection with virtual cycling can increase both the intensity and enjoyment of high-intensity interval cycling by untrained individuals (Farrow et al., 2019).

An additional advantage is that the cyclist does not have to worry about detrimental environmental factors, such as extreme temperatures, rain, snow, strong winds or air pollution (Heesch et al., 2012). While training indoors, a cyclist can control the temperature and humidity and even simulate different altitudes with hypoxia-inducing procedures.

Gamification

The gamification of indoor cycling, with feedback loops commonly employed in video games, has lead to a myriad of possibilities for interactive usage that enhances engagement (Beatty, 2013). With the virtual training platform, successful performance is rewarded with special currency, experience points and levels that can be used to make in-game purchases, e.g., bike frames and wheelsets with properties (better aerodynamics or lighter weight) that can improve performance. As has been shown in connection with many exercise tasks (Van Der Kooij et al., 2019; Van Mastrigt et al., 2020), such rewards may motivate users and encourage them to exercise at higher speeds, climb more meters or ride for longer periods to accumulate even greater rewards

In addition, at random points along the course, virtual cycling platforms offer temporary events, called power-ups, that can boost performance, ranging from a reduction in drag to a decrease in the cyclist's body mass, a feature similar to those in many video games.

This can both attenuate the perceived level of exertion, thereby promoting more prolonged and/or intense cycling, and make the experience more versatile and enjoyable (Farrow et al., 2019).

Moreover, the gaming nature of this program may attract new participants by including music and social interactions (e.g., multiplayer options that allow friends to be included or guidance to be received from experienced players), as well as reducing frustration due to poor-quality graphics and overly complex controls and display functions that may evoke motion sickness (Faric et al., 2019).

Finally, virtual forms of training may allow players to engage in more physical activity thereby reducing screen time and self-efficacy (Staiano et al., 2017).

WEAKNESSES

Accuracy

Questionable accuracy has been one of the most obvious weaknesses of Zwift (Whiting, 2018). The many types and models of trainers involved require multiple ways of measuring power output. Some trainers have built-in power meters; others require external devices; and some require speed and cadence sensors which use Zwift's own algorithms to estimate power output. Alternatively, meters on the crank-arm, wheel hub or pedal of the cycle, each with its own level of accuracy, can be used to monitor power.

Zwift applies an algorithm to convert this measured power output to in-game speed. This offers a somewhat crude estimation of actual speed since, as explained in more detail elsewhere, it is based on several factors, including the cyclist's mass, height and choice of bike (Schlange, 2020a).

In this context, aerodynamics, which have a considerable influence on outdoor cycling (Atkinson et al., 2003), are only measured in basic terms of height and mass, with adaptations for specific in-game bicycle and wheel choices. The cyclist's body size and shape are not considered, nor is their riding positions. Cyclists with superior technique and flexibility may be able to assume more aerodynamic positions than others, but this has no impact in-game.

At present, for the devices commonly employed to measure power, manufacturers report a variance in accuracy of ±1–3% (TacX, Wahoo, Elite, 4iiii and Stages). Although this may not be important to a recreational rider, for a competitive cyclist it could well mean the difference between winning and losing. Therefore, for appropriate simulation and interpersonal racing in Zwift, this accuracy must be improved. At present, elite cyclists must verify their Zwift power data with a secondary measuring device, which entails additional expense and technical experience.

The cyclists using "dumb" trainers, with only speed and cadence monitors and no power measurement device, make use of Zwift's alternative algorithms for estimating power (Zwift Power or ZP). Within the racing community these are not considered reliable, and many races exclude riders using these algorithms when reporting results. This could lead to simulated high-level racing becoming an elitist sport.

Indoor vs. Outdoor Load Metrics

Many recreational and competitive cyclists train both indoors and outdoors over the course of a season. Depending on the technology involved, cyclists may perceive these two types of training differently. In fact, power output and heart rate during cycling outdoors and indoors may differ (Mieras et al., 2014). Thus, internal and external load metrics associated with indoor and outdoor cycling cannot be applied interchangeably.

Inaccurate Data Entry and "Cyber-Doping"

For the estimated power and actual power algorithms offered by Zwift to function, the user must provide body mass and height to establish an individual drag coefficient for drafting, riding solo, leading groups, or riding up- or downhill. Some cyclists may not know their actual mass and therefore enter incorrect data. False data can affect performance outcomes, since Zwift utilizes watts per kilogram body mass as the main determinant of avatar speed. Moreover, entering an incorrect height would alter the drag coefficient, the second determinant of avatar speed. More concerningly, the cyclist may deliberately enter an incorrect body mass and/or height to improve apparent performance, a practice nicknamed "cyber-doping" and seen by the Zwift community as analogous to real-life doping. There have also been cases of gender swapping, most frequently by men, who then participate in competitions for women only.

So far, these practices have been policed by the community itself, with users flagging suspect performances or requesting verification of power data and/or body mass through external forums. The most common and supposedly robust enforcement involves suspending a cyclist until his/her power data is verified by a secondary power source, although this approach is not always readily available and entails additional cost. Riders can also be suspended until a verified weigh-in video is provided, but this is rare and more questionable, as weighing scales are often poorly calibrated.

Such factors can lead to confusion when reporting results. The results of traditional (non-virtual) elite races are released almost immediately, allowing athletes, teams, and sponsors to celebrate their successes. Zwift's requirements for verification of performances that are suspect could reduce confidence in the results and undermine public perception of the races.

System Failure

Another weakness of the Zwift system are dropouts, i.e., shorter- or longer-term loss of Bluetooth or Ant+ connectivity between power meters, trainers or computing devices used for simulation. The racing community calls these events "cyber mechanicals," in analogy to the mechanical failures seen in non-virtual bike races. Dropouts are relatively rare, sometimes only lasting a matter of seconds, but since they may occur at any time, these can still exert a considerable impact on apparent performance, especially during a race. Faulty hardware, problems with software including bugs and/or hosting, and human error (such as not charging devices) can all lead to dropouts. Regardless of the cause, dropouts constitute a risk that sponsors and/or athletes may find unpalatable.

The Human Component

The very nature of the simulation may reduce the skillful technique and bike handling needed for success in elite non-virtual races. Because of the way it is constructed, sprinting maximally on an ergometer is different to sprinting outdoors. The platform simulates cornering, so there is no need for the user to do so. Furthermore, it is not necessary to understand body positioning while descending or braking and distance management within a group of cyclists. Onscreen avatars and power data make it difficult to determine whether attacks or changes in pace will have the desired effect. Crowds, which can provide emotional support and a sense of gratification when successful, are absent. This may reduce the enthusiasm of both the competitor and sponsors. Furthermore, the overall performance of elite cyclists can be affected by the skills and characteristics of their teammates (e.g., cyclists often try to help the team leader win at the cost of their own chances) (Torgler, 2007). In general, virtual racing may attenuate the intuitive feelings of real-life racing.

OPPORTUNITIES

A New "Normal"

The ongoing Covid-19 pandemic is causing more and more individuals to incorporate virtual platforms into their daily lives. Online exercise and virtual personal training are becoming more common (Thompson, 2021). At the same time, many global sports competitions have been postponed or canceled, opening opportunities for a viable and stable virtual platform to offer alternatives to professional athletic competitions. As long as the availability of live sports events remains limited, the numbers of viewers may increase, and new audiences may be captured.

New Event Formats, Sponsors, and Teams

Collaboration between event organizers and commercial brands is on the rise, with the first virtual Tour de France in July 2020 (www.letour.fr, 2020). Now teams (some associated with professional teams who compete in Grand Tour races) that focus solely on virtual racing through Zwift are being formed. In addition to the virtual world championships on the Zwift program each year, the three Grand Tours of cycling (Tour de France, Giro d'Italia, Vuelta a España) could also conduct virtual races. The high-profile one-day races could be added as well. In this way elite cyclists could compete year-round with teams and brands exposed to new audiences.

Moreover, the new technology in combination with the pandemic situation presents opportunities for changing the traditional structure of road cycling teams. Cyclists who normally sacrifice personal chances of success by drafting a team leader could race more aggressively, potentially leading to more exciting races and new cycling stars. Furthermore, with shorter races and the ability of each cyclist to prepare nutrition and hydration in advance and keep these close at hand, less time will be lost in this respect and fewer employees assigned to such tasks will be required.

The Crossover Athlete, Talent Identification, and Coaching

Cyclists who normally specialize in one type of event could try racing in different competitions, e.g., road cyclists could compete in virtual mountain bike races, BMX riders could try gravel racing, etc. Virtual online platforms could also expand to include track cycling disciplines, BMX racing, cyclocross and fixed gear racing. World championship in all-round categories could be offered.

This situation could well-lead to the identification of new talents, with cyclists being particularly successful in virtual disciplines they have not competed in previously. In this manner virtual online cycling platforms could become a testing agency for National Governing Bodies and Olympic Federations. As an example, this is currently operated through an academy, which partners with a professional cycling team to offer a male and female rider a development contract (Norman, 2020), but it could be expanded upon.

Partnering with high-level coaches to provide a greater variety of in-game training plans is another opportunity. With data collected being fed back to the coach and alterations being made where necessary to accommodate training adaptations and responses (Duking et al., 2018, 2020), it is possible that digital online coaching could reduce incidents of overtraining and injury. This could also be accomplished through additional algorithms that automatically adjust the resistance of "smart" trainers when an athlete is training at an intensity that is too high (Duking et al., 2020). Outdoor cycling data could also be added to these algorithms to create a more complete training plan. Such digital coaching framework (Duking et al., 2018) could allow athletes to exercise and train with a quality they may not normally have access to. There are currently external platforms providing this service (e.g., Today's Plan and others), but this approach could be included within the virtual online training platform itself, thereby increasing usage and control.

Finally, virtual platforms provide athletes with physical or cognitive disabilities with opportunities they do not have in the real world. Paralympic disciplines can be included, allowing for increased inclusion and diversity.

Enhanced Modeling and Simulation

Combining power readings with video data captured from multiple angles could allow for more accurate avatar modeling, thereby increasing the realistic nature of in-game performance. Photographs or video of riders in their preferred riding positions on the bicycle could be beneficial to those with certain body shapes or greater flexibility.

Additionally, virtual online training platforms could simulate weather simulations that a rider might normally encounter, including changes in temperature, wind direction and speed, and rain or snow. This would require cyclists to adapt their tactics to cope with the changes (such as sheltering from head and crosswinds when riding in a group or reducing visibility and avatar responsiveness when riding in rain or snow).

Virtual online platforms may offer opportunities for field-based studies related to both the training and racing aspects of cycling, and the inter-relationship between the two, as all exercises are performed using the same platform and equipment, and may offer the opportunity to recruit many participants.

In this context the so-called "ERG (short for ergometer) mode" allows the resistance of this device to be set automatically. Use of this mode requires a smart bike trainer in combination with either a compatible app or computer that makes it possible to adjust the resistance remotely and maintain constant power output during a workout.

INTEGRATING ASPECTS OF eSPORTS

Virtual athletic platforms will lead to the development of new tactics that could enhance public engagement and excitement. This may explain, at least in part, the surge in popularity of eSports, with tickets for multi-day elite competitions selling out. First events in endurance sport (Ltd, 2020; Triathlon, 2020), a fusion of real-life and virtual triathlon and cycling, immerses fans in a view of the world's best athletes and provides them with actual power, speed and heart rate data collected by Zwift. In addition to attracting new fans, this concept could provide more revenue for athletes, teams, and sponsors.

In-Game Success and Real-World Advantages

By offering discounts on specific products based on, e.g., the distance cycled, meters climbed, or points accumulated through racing, sponsors could entice users through the gamified feedback loops. To a limited extent, in-game uniforms and unlockable bikes are already offered as rewards for the completion of specific rides or challenges and this could easily be extended to benefits in real life.

This strategy to promote health would allow governments, organizations concerned with public health, and insurance companies to reward users for participation with vouchers or promotional codes.

Expansion Into New Sports

At present cycling is the leading sport in terms of virtual simulation (through platforms such as Zwift, Real Grand Tours, Rouzy, Sufferfest, Peloton, and TrainerRoad), but the use of analogous simulations for running (Zwift, NordicTrack, Peloton) is expanding. Opportunity exists for expansion into other sports for which reliable indoor training equipment is available, such as rowing and cross-country skiing. In the case of rowing, online comparison of performances has been available for some time, but without any virtual simulation.

THREATS

Cheating

Cheating remains the foremost threat to Zwift. In the case of some trainers, participants have succeeded in "hacking" and controlling the power-meter remotely (Yeager, 2019). Since some participants have been accused of "cyber-doping," a Zwift Anti-Doping Agency ("ZADA") was installed to penalize fiction wattage, misrepresentative metrics and gender swapping (Yeager, 2018).

Lack of Acceptance

Expansion of virtual competitions will require a certain level of acceptance by existing teams, athletes, coaches, sponsors, and organizers of competitions. If only a few parties accept such virtual competition, it may be viewed as less legitimate and thereby struggle to maintain interest, generate sufficient revenues, and even survive (Akenhead and Nassis, 2016).

In the end, virtual training and competition may come to be a fad (Best, 2006). Especially when the current global pandemic restrictions end and people leave their homes freely to exercise and socialize, they may prefer to return to "real-world" experiences. However, the recent worldwide survey of fitness trends for 2021 indicates clearly that virtual training is not simply a fad (Thompson, 2021).

Lack of Competition

At present Zwift is the market leader for simulated cycling competition, but without serious competitors it could become an echo chamber of sorts, reducing the drive for innovation and development that might occur if there were competitors of a similar standard.

Strength

- Affordable, can suit a wide range of budgets
- Allows for innovative team management strategies
- Realistic simulations of different environments
- Increased rider safety compared to road cycling
- Increased user and crowd engagement through gamification

Weakness

- Accuracy across platform and connected devices is highly variable
- 'Cyber-doping' has been seen
- Problems with hard- and/or software
- Technical cycling skills can be reduced

Opportunities

- Increased usage during global pandemic
- New competition formats available
- Attraction of new sponsors for events, teams, and individual athletes
- One bicycle, multiple race opportunities
- Increased talent identification
- Improved access to remote coaching
- Enhanced programming including weather simulation (wind speed/direction, temperature, rain/snow)
- Follow eSports with crowd and user engagement strategy
- In-game links to real-world rewards
- Expansion into other sports

Threats

- Cheating
- Lack of mass acceptance
- Lack of competition leading to reduced innovation
- Currently recreational and competitive, one may have to take priority
- Health risks
- Data risks

FIGURE 1 | A summary of strengths and weaknesses, as well as opportunities and threats associated with virtual training in cycling.

Competition or Recreation

Currently, Zwift offers opportunities for competitive and recreational users, but there may come a point in the future when one of these markets is more viable than others. The community aspect drives most daily users, with substantially more people using the platform for training and non-structured riding than racing and competition. Previous research has shown that exergaming supports feelings of competitiveness among those who already identify as competitive and has detrimental effects on those who identify as less competitive (Song et al., 2013). It is possible that the pattern for virtual online platforms is similar, with only those identifying as competitive feeling engaged by the racing aspect of the platform. Racing raises the profile of athletes, brands, and sponsors in a way that recreational use will not, but if user feedback on the recreational component is more positive, then racing may fade from prominence.

Health Risks

There is some risk that while competing virtually, athletes exercising at-home may push themselves beyond their own safe physical limits and experience an adverse reaction (e.g., injury, nausea, fainting, or injury) in a situation where no supervision or support is available. Cycling indoors without adequate air flow for cooling and sufficient intake of fluids can result in dehydration, thereby imposing additional physiological strain on the cyclist (Ramos-Jimenez et al., 2014).

Data Security

Finally, the large amounts of data provided by users of virtual programs for training and competition are prone to hacking (Yeager, 2019) and must be protected from inappropriate external access (Spiegel, 2018).

SUMMARY

Virtual training may offer many strengths, opportunities, weaknesses, and threats to cyclists engaging in this new technology, as summarized in **Figure 1**.

In conclusion, virtual online cycling platforms can build upon its strengths of immersive worlds, innovative drafting mechanics, and versatility by enhancing realism, improving data accuracy, and increasing the strength and reliability of its power-speed algorithms. Opportunities exist for expanding strategic partnerships with major cycling races, brands, and sponsors. User experience can be improved with the addition of video capture and "e-coaching." Threats are present in the form of cheating, a lack of acceptance and usage by a broader community, health risks and data insecurity.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Aertsens, J., De Geus, B., Vandenbulcke, G., Degraeuwe, B., Broekx, S., De Nocker, L., et al. (2010). Commuting by bike in Belgium, the costs of minor accidents. *Accid. Anal. Prev.* 42, 2149–2157. doi: 10.1016/j.aap.2010.07.008

Ag, C. (2019). *Bicycle Cities Index 2019*. Available online at: http://www.coya.com: https://www.coya.com/bike/index-2019 (accessed January 11, 2021).

Akenhead, R., and Nassis, G. P. (2016). Training load and player monitoring in high-level football: current practice and perceptions. *Int. J. Sports Physiol. Perform.* 11, 587–593. doi: 10.1123/ijspp.2015-0331

Atkinson, G., Davison, R., Jeukendrup, A., and Passfield, L. (2003). Science and cycling: current knowledge and future directions for research. *J. Sports Sci.* 21, 767–787. doi: 10.1080/0264041031000102097

Beatty, I. D. (2013). "Improving physics instruction by analyzing video games," in *AIP Conference Proceedings: American Institute of Physics*, 70–73. doi: 10.1063/1.4789654

Best, J. (2006). *Flavor of the Month: Why Smart People Fall for Fads*. Berkeley, CA: University of California Press. doi: 10.1525/9780520932357

Duking, P., Achtzehn, S., Holmberg, H. C., and Sperlich, B. (2018). Integrated framework of load monitoring by a combination of smartphone applications, wearables and point-of-care testing provides feedback that allows individual responsive adjustments to activities of daily living. *Sensors* 18:1632. doi: 10.3390/s18051632

Duking, P., Zinner, C., Reed, J. L., Holmberg, H. C., and Sperlich, B. (2020). Predefined vs data-guided training prescription based on autonomic nervous system variation: a systematic review. *Scand. J. Med. Sci. Sports* 30, 2291–2304. doi: 10.1111/sms.13802

Faric, N., Potts, H. W. W., Hon, A., Smith, L., Newby, K., Steptoe, A., et al. (2019). What players of virtual reality exercise games want: thematic analysis of web-based reviews. *J. Med. Internet Res.* 21:e13833. doi: 10.2196/13833

Farrow, M., Lutteroth, C., Rouse, P. C., and Bilzon, J. L. J. (2019). Virtual-reality exergaming improves performance during high-intensity interval training. *Eur. J. Sport Sci.* 19, 719–727. doi: 10.1080/17461391.2018.1542459

Fowler, P. M., Knez, W., Crowcroft, S., Mendham, A. E., Miller, J., Sargent, C., et al. (2017). Greater effect of east versus west travel on jet lag, sleep, and team sport performance. *Med. Sci. Sports Exerc.* 49, 2548–2561. doi: 10.1249/MSS.0000000000001374

Heesch, K. C., Sahlqvist, S., and Garrard, J. (2012). Gender differences in recreational and transport cycling: a cross-sectional mixed-methods comparison of cycling patterns, motivators, and constraints. *Int. J. Behav. Nutr. Phys. Act.* 9:106. doi: 10.1186/1479-5868-9-106

Le Meur, Y., Hausswirth, C., and Mujika, I. (2012). Tapering for competition: a review. *Sci. Sports* 27, 77–87. doi: 10.1016/j.scispo.2011.06.013

Long, M. (2020). *Zwift Raises US$450m in Series C Funding Round Led by KKR*. Available online at: https://www.sportspromedia.com/ (accessed January 11, 2021).

Ltd., W. T. R. L. (2020). *Welcome to Zwift Racing Leagues in Partnership With WTRL*. Available online at: https://www.wtrl.racing/zrl/Zwift-Racing-League.php (accessed January 11, 2021).

Mieras, M. E., Heesch, M. W., and Slivka, D. R. (2014). Physiological and psychological responses to outdoor vs. laboratory cycling. *J. Strength Cond. Res.* 28, 2324–2329. doi: 10.1519/JSC.0000000000000384

Norman, P. (2020). *Zwift Academy Returns With a Contract With Alpecin-Fenix or Canyon//SRAM on offer for winner*. Available online at: https://www.bikeradar.com/news/zwift-academy-2020/ (accessed January 11, 2021).

Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. *Comput. Educ.* 147:103778. doi: 10.1016/j.compedu.2019.103778

Ramos-Jimenez, A., Hernandez-Torres, R. P., Wall-Medrano, A., Torres-Duran, P. V., Juarez-Oropeza, M. A., Viloria, M., et al. (2014). Gender- and hydration-associated differences in the physiological response to spinning. *Nutr. Hosp.* 29, 644–651. doi: 10.3305/nh.2014.29.3.7017

Schlange, E. (2020a). *How Does Zwift Calculate Rider Speed?*. Available online at: https://zwiftinsider.com/how-does-zwift-calculate-my-speed/ (accessed January 11, 2021).

Schlange, E. (2020b). *New "PEAK ZWIFT" Achieved: 34,940*. Available online at: https://zwiftinsider.com/peak-zwift-34940/ (accessed January 11, 2021).

Song, H., Kim, J., Tenzek, K. E., and Lee, K. M. (2013). The effects of competition and competitiveness upon intrinsic motivation in exergames. *Comput. Human Behav.* 29, 1702–1708. doi: 10.1016/j.chb.2013.01.042

Spiegel, J. S. (2018). The ethics of virtual reality technology: social hazards and public policy recommendations. *Sci. Eng. Ethics* 24, 1537–1550. doi: 10.1007/s11948-017-9979-y

Staiano, A. E., Beyl, R. A., Hsia, D. S., Katzmarzyk, P. T., and Newton, R. L. Jr. (2017). Twelve weeks of dance exergaming in overweight and obese adolescent girls: transfer effects on physical activity, screen time, and self-efficacy. *J. Sport Health Sci.* 6, 4–10. doi: 10.1016/j.jshs.2016.11.005

Thompson, W. R. (2021). Worldwide Survey of Fitness Trends for 2021. *ACSM's Health Fitness J.* 25, 10–19. doi: 10.1249/FIT.0000000000000631

Torgler, B. (2007). "La Grande Boucle" determinants of success at the tour de France. *J. Sports Econ.* 8, 317–331. doi: 10.1177/1527002506287657

Triathlon, S. L. (2020). *SLT Arena Games in Rotterdam 2020*. Available online at: https://superleaguetriathlon.com/event/slt-arena-games-rotterdam/ (accessed January 11, 2021).

Van Der Kooij, K., Van Dijsseldonk, R., Van Veen, M., Steenbrink, F., De Weerd, C., and Overvliet, K. E. (2019). Gamification as a sustainable source of enjoyment during balance and gait exercises. *Front. Psychol.* 10:294. doi: 10.3389/fpsyg.2019.00294

Van Mastrigt, N. M., Smeets, J. B. J., and Van Der Kooij, K. (2020). Quantifying exploration in reward-based motor learning. *PLoS ONE* 15:e0226789. doi: 10.1371/journal.pone.0226789

Whiting, J. (2018). *How accurate is Zwift's Power Estimate for Classic Trainers?*. Available online at: https://machiine.com/2018/how-accurate-is-zwifts-power-estimate-for-classic-trainers/#:~:text=In%20the%20lower%20power%20range,or%20without%20a%20power%20meter (accessed January 11, 2021).

Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence: Teleoper. Virtual Environ.* 7, 225–240. doi: 10.1162/105474698565686

www.letour.fr (2020). *Virtual Tour de France on Zwift - Riding for Solidarity*. Available online at: https://www.letour.fr/en/virtual-tour-de-france (accessed January 11, 2021).

Yeager, S. (2018). *Thinking of Cheating on Zwift? These People Are on to You*. Available online at: https://www.bicycling.com/culture/a19723490/thinking-of-cheating-on-zwift-these-people-are-on-to-you/ (accessed January 11, 2021).

Yeager, S. (2019). *Zwift Hackers Expose the Next Generation of Cycling Doping*. Available online at: https://www.bicycling.com/news/a28912281/zwift-hacking/ (accessed January 11, 2021).

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

frontiers in Sports and Active Living

METHODS
published: 04 June 2021
doi: 10.3389/fspor.2021.678489

Decision Support System Applications for Scheduling in Professional Team Sport. The Team's Perspective

Xavier Schelling[1*], *Jose Fernández*[2], *Patrick Ward*[3], *Javier Fernández*[4,5] and *Sam Robertson*[1]

[1] Institute for Health and Sport, Victoria University, Melbourne, VIC, Australia, [2] Faculty of Health Sciences, School of Behavioral and Health Sciences, Australian Catholic University, Fitzroy, VIC, Australia, [3] Human Performance Research Centre, Faculty of Health, University of Technology Sydney (UTS), Sydney, NSW, Australia, [4] Futbol Club Barcelona, Barcelona, Spain, [5] Department of Computer Science, Universitat Politecnica de Catalunya, Barcelona, Spain

OPEN ACCESS

Edited by:
Billy Sperlich,
Julius Maximilian University of Würzburg, Germany

Reviewed by:
Julio Calleja-Gonzalez,
University of the Basque Country, Spain
Enzo Iuliano,
eCampus University, Italy

***Correspondence:**
Xavier Schelling
xschelling@gmail.com

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: 09 March 2021
Accepted: 27 April 2021
Published: 04 June 2021

Citation:
Schelling X, Fernández J, Ward P, Fernández J and Robertson S (2021) Decision Support System Applications for Scheduling in Professional Team Sport. The Team's Perspective. Front. Sports Act. Living 3:678489. doi: 10.3389/fspor.2021.678489

Background: Periodization implies the systematic planning of training and competition with the goal of reaching the best possible performance in the most important competition. In team sports, this consists of finding a flight-and-practice schedule that maximizes the opportunities to perform the periodized contents (e.g., trips, practices, games, and days off). This process is conducted whilst considering known constraints (e.g., competitive schedule, roster availability, weather, especial events, holidays, or emotional effect of days away). The way a scheduling decision support system (DSS) leads users to make a decision should allow for flexibility, whilst minimizing users' confusion and facilitating the understanding of the recommendation given by the scheduling decision support system. Traditional approaches to solving scheduling problems use either simulation models, analytical models, heuristic approaches or a combination of these methods. When it comes to evaluate how the scheduling DSS is performing, three overarching aspects need to be reviewed: context satisfaction, process efficiency, and output quality. Appropriate training periodization and scheduling of trips and training sessions are critical for teams to optimize training and recovery processes in order to maximize health and performance. This article presents a methodological framework for designing decision-support systems for scheduling in professional team sports.

Keywords: decision making, information system, sport science, optimization, scheduling

INTRODUCTION

Professional sport leagues involve millions of fans, broadcast rights, merchandizing, and advertising. Therefore, they constitute a major economic activity, where revenue maximization and logistical optimization are key factors (Kendall et al., 2010). Consequently, in popular team sports such as soccer, basketball, baseball, or ice hockey, it is common to have several games per week (i.e., $\geq$3) per team throughout a competitive season. Additionally, any professional sport requires training and traveling periodically, which should be periodized considering the competitive calendar. Periodization implies the systematic planning of training and competition with the goal of reaching the best possible performance in the most important competition of the season (Robertson and Joyce, 2018). This goal involves the development and optimization of the multiple factors that drive sport performance, which rely on psychological and physiological processes (e.g., fitness, cognition, and emotions), as well as environmental conditions (e.g., weather, equipment, rewards) (Seirul·lo, 1998).

Most professional team sports globally utilize a tournament format where each team plays against every other a fixed number of times (also known as "all-play-all" or "round robin tournament") (Ribeiro, 2012; Byl, 2013). Every team has prior knowledge of opponents along with, the date, location and time in which they will compete, which provides an opportunity to prepare for both the tournament and upcoming games (Byl, 2013). In some leagues (e.g., National Basketball Association—NBA), the exact day and time for all games is released before the season starts, in others (e.g., La Liga) they are defined throughout the season, for instance five games in advance. Each team typically has its own venue at its home city and each game is played at the venue of either one of the two teams in confrontation.

The timing of a national league season (i.e., domestic league) must be coordinated with international competitions such as World Cups, Olympic Games, Eurocup, Pan-American games, Asian games, Champions League, etc. Depending on the sport and the country, the effect of international competitions can be significant since the best players will not play in their domestic league program unless the calendar is adjusted accordingly. Some domestic leagues also include special events or tournaments such as the all-star weekend, the Challenge cups, or the Supercups.

Concerns around congested competitive schedules have been publicly shared across sports (Kloke, 2016; Holmes, 2018; Sport, 2020), with predominant reasons including a lack of training and recovery opportunities, and potential sleep deprivation, which can have a negative effect on the player's health (Teramoto et al., 2017; Lewis, 2018; Rossi et al., 2018) or teams' performance (Moskowitz and Wertheim, 2011; Mitchell et al., 2019; Esteves et al., 2020). Such effects could also lead to a lower product quality for consumers and broadcasters (Shelburne, 2017). Although the question of whether schedule density impacts injuries is complex, as it requires a multifaceted analysis, adjusting for many related factors such as prior injury, travel time, time zone difference, home vs. away, or acute vs. overuse injuries (Mack et al., 2018); sleep, training, and recovery opportunity are impaired due to the traveling schedule of team sports athletes (Sortino, 2015; Fullagar et al., 2016; Nutting and Price, 2017; Lastella et al., 2019). Additionally, in teams or leagues with lower budgets, or amateur sports, substantial differences in travel quality, particularly the presence of bus trips, non-charter flights, and the inevitable differences in hotel and restaurant accommodations should also be considered (Mitchell et al., 2019). Against this background, leagues have tried to modify schedules in the spirit of creating more non-game days and better traveling combinations (Holmes, 2018). Nevertheless, for especially congested periods of the season, some teams may still opt to rest players in order to provide them with extra recovery time, entailing a negative effect on the team's competitiveness and the game-product quality (Shelburne, 2017).

Appropriate training periodization and scheduling of trips and training sessions will be critical for teams to optimize training and recovery opportunity in order to maximize health and performance. This article presents a methodological framework to designing decision-support systems for scheduling in professional team sports. The proposal will follow a previously published decision support system framework (Schelling and Robertson, 2020) which considers the organization's needs, the efficiency of the processes, and the quality of the system's recommendation.

SCHEDULING PROBLEM DESCRIPTION

Problem Definition

Conceptually, a team's schedule problem consists of finding a flight-and-practice schedule for the pre-season and the regular season that maximizes the opportunities to perform the periodized contents (e.g., trips, practices, games, and days off). This activity is required whilst considering known constraints (e.g., competitive schedule, roster availability, weather, special events, holidays, and emotional effect of days away). Hence, designing a schedule is a combinatorial problem, consisting of a set of instances or inputs, candidate solutions for each instance, and an overall outcome for each candidate solution (Goldreich, 2008; Mahapatra et al., 2017).

Schedule-related problems have two important features (Balas, 1999): *Constraints*, a formal description of the requirements that must be satisfied by a candidate solution to the problem; for example, a team has to be at a specific date, time and location to play the upcoming game; and an *optimization indicator*, which characterizes the quality of the recommendation. The optimization indicator represents a value whose calculation is based on the recommended solution; for example, to minimize the distance traveled in a regular season.

There are two levels of planning and scheduling depending on the time scale of decision-making. The first level "predicts" the schedule, whereas the second level "reacts" to the current local situation and is often called reactive scheduling (Aytug et al., 1994). Both levels are important; predictive scheduling is useful for macro planning (i.e., season overview), utilizing invariant information available earlier, whereas reactive scheduling should allow for enhanced decision-making thanks to better and recent information, closer to the action at hand (i.e., micro planning). Reactive scheduling is more difficult to analyze and provide meaningful automated help due to the unpredictable and recency nature of the required information to make the decision. Training session scheduling is an example of reactive scheduling, where factors such as roster availability or team performance may cause disruption in the team environment requiring a different schedule from the originally planned. Coaching and performance staff are accustomed to dealing with such disruptions. However, their decisions may be crisis-oriented or biased with little attention given to the bigger picture and impact therein (Aytug et al., 1994; Cross et al., 2019). If a computer-aided method

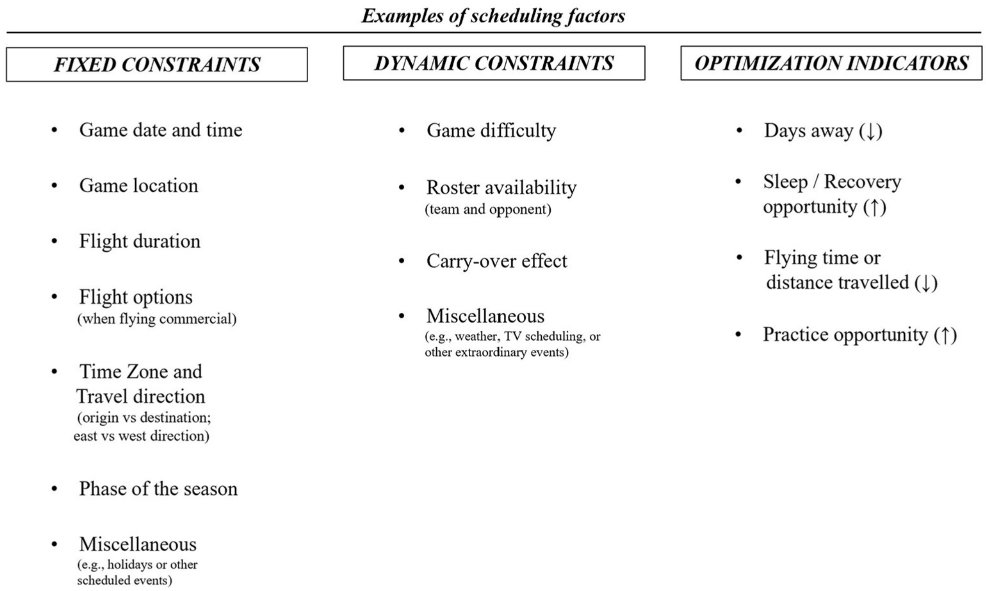

FIGURE 1 | Examples of fixed and dynamic constraints, and optimization indicators relating to scheduling in professional team sport. There are potentially an infinite number of constraints and optimization indicators that could be included. Some of them are interrelated and may change over time. Different constraints and optimization indicators can be defined among various sports.

is used for reactive scheduling it must be periodically iterated throughout the season. When new solutions require continual re-computation due to contextual changes over time the scheduling-problem is referred to as an online problem, whereas an offline problem is when information about all activities, resources, constraints and optimization indicators are known in advance, and the goal of the decision support system (DSS) is to find a single "good" solution to the problem (Wang et al., 2003).

There can be several reasons to develop a DSS for scheduling (Schelling and Robertson, 2020): the schedule simply requires application of a set of heuristic rules; the process can be automated; the current scheduling process is largely subjective or solely expertise-based; there is current disagreement among staff on how to design the schedule; new data (or criteria) allows for a re-structure of the scheduling process; team schedule has a significant impact on performance and thus warrants optimization. Additionally, when a scheduling DSS is built, the organization's knowledge about the domain becomes explicit. This enables one to study that knowledge, to critique it, to use it in training, and to preserve it over time (Fox, 1990). Last, understanding how the organization resolved scheduling-problems in the past, the available and required information-systems (hardware, software, and data workflow), the required time or deadline to solve the schedule, and the satisfaction with the implemented schedules in the past will help defining the feasibility and design of the DSS before starting its development (Schelling and Robertson, 2020).

Constraints and Optimization Indicators

A schedule is affected by several restrictions, or constraints. These can be "fixed" (those constraints set prior to the start of the season and with none or very low variability throughout the season) or "dynamic" (those which are subject to change throughout the season) (Robertson and Joyce, 2018). Some examples of fixed constraints include the competitive calendar (game date, time, and location/topography), flight duration, flight options (when flying commercial), or time zone difference. Examples of dynamic constraints include game difficulty, standings, or roster availability (**Figure 1**). Some expertise-based heuristics such as preferred arrival times or accommodation preferences must be also considered as constraints when developing any DSS.

Moreover, there are schedule-problems where the goal is to optimize (maximize or minimize) an outcome variable, for instance the numbers of days away, or the distance traveled. Some examples of schedule optimization problems are spending the least possible number of days in a city with a time zone difference larger than "*x* hours," selection of arrival time to avoid traffic in rush hours or canceling or modifying a scheduled practice session if not enough available players. In such problems the DSS will require from an optimization indicator (e.g., days away, distance traveled, recovery opportunity, practice opportunity, etc.). There are potentially an infinite number of constraints and optimization indicators that could be included, and most of them are interrelated and may change over time (Rocha, 2017) (**Figure 1**).

Data Input and Sources

When developing a decision support system, data quality, including data meaning, availability, structure, integration, accessibility, and timeliness of retrieval, are critical aspects for a successful implementation (Schelling and Robertson, 2020). When direct connections (i.e., application programming interface or API) between a team's database and the league or a website's database are not available, web harvesting or scraping techniques can be explored to automate and facilitate one-time data extraction or regular feeding from online servers (Glez-Peña et al., 2014). Considering the fixed and dynamic constraint examples shown in **Figure 1** below are listed some considerations regarding data input quality when developing decision support system for scheduling.

- Fixed constraints
  - Game location, opponent teams, dates, times, and phase of the season (pre-season, regular season, playoffs, finals, post-season) are defined by the official competitive calendar. In professional leagues the game schedule for the regular season is released several weeks before the start of the season in order to allow teams to arrange transportation and accommodation. This information is usually publicly available on each league's website (e.g., La Liga, NBA, National Football League—NFL, Major League Baseball—MLB, etc.).
  - International competition calendars are also made publicly available by the global governing body for each sport (FIBA, FIFA, IOC, etc.).
  - Geodesic distance (Karney, 2013) between cities and other travel related factors can be retrieved from public websites (e.g., www.distancecalculator.net) or automated via open source platforms.
- Dynamic constraints
  - Game difficulty, or win probability, considers factors such as game schedule, roster quality, home court advantage, team form, or game importance to provide a continuous (points spread) or discrete (win/lose) game outcome prediction for each team. Game difficulty can be developed internally as a sub-model within the scheduling decision support system, or retrieved from public sources (e.g., www.fivethirtyeight.com).
  - Daily standings and game results can be obtained from the official websites of the league, sport news websites (e.g., www.espn.com), or sport-specific sources (e.g., www.baseball-reference.com).
  - Daily roster availability can be retrieved from the team's athlete management system (AMS) or manually entered before the upcoming practice or game. Some sport news websites (e.g., www.espn.com) publish the injuries by team daily. Nevertheless, roster availability is often not accurate (i.e., low data quality) as there can be last-minute roster changes. Some leagues allow until 1 h before the start of the game to list a player as unavailable. Roster availability will also be affected by individual load-management needs (i.e., resting a player for a game or practice as a prophylactic strategy) (Drew and Finch, 2016), which is another example of reactive individual scheduling.

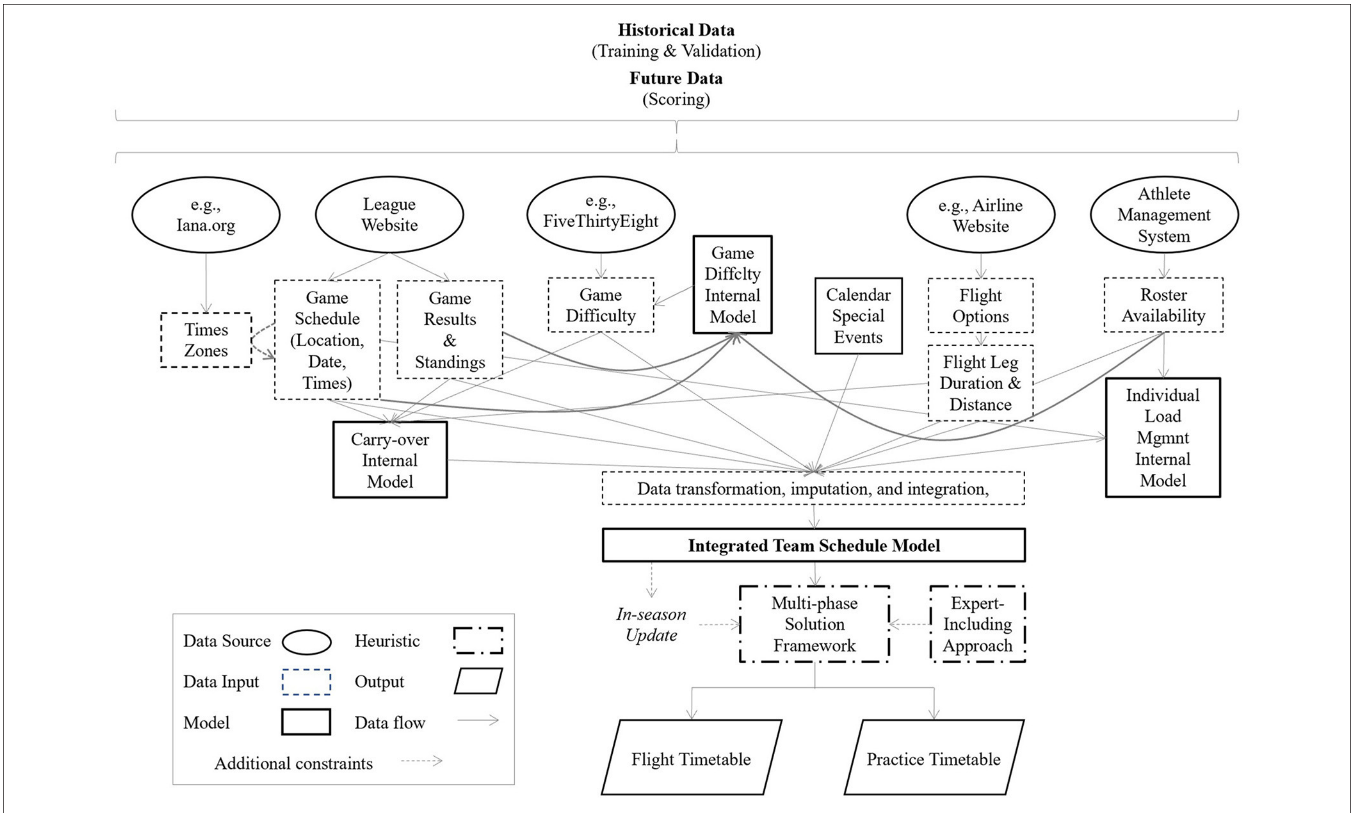

FIGURE 2 | Example of the model architecture of a scheduling decision support system.

  - Carry-over effect, or the effect of previous events on future performance (Guedes and Ribeiro, 2011; Goossens and Spieksma, 2012) will require from integrating multiple features or even having a sub-model within the scheduling DSS.

- Data input integration refers to combining multiple sources or types of data (fixed or dynamic) to create new contextual knowledge regarding the goal at hand, thereby increasing data quality (Kenett and Shmueli, 2016). Data integration could also help optimizing the decision support system's complexity and performance, for example by reducing the data dimensionality or creating richer input features (Schelling and Robertson, 2020). Some examples are:
  - Schedule congestion indicators derived from game schedule (date and time) such as number of hours between games, number of games over time (e.g., number of games in 7 days, etc.), or labeling the game congestion with arbitrary categorical indicators (e.g., back-to-back, 3-in-4, or 4-in-5).
  - Team performance indicators based on expected performance (e.g., game difficulty or win probability) and recent performance (e.g., production in attack and defense).

Figure 2 shows an example of model architecture including several data sources and sub-models. The example represents a multi-phase solution including different processes based on what needs to be scheduled, the available information, timescale, and the expert's knowledge:

- Phase 1: Initial competitive calendar analysis and exploration,
- Phase 2: Flight schedule recommendation,
- Phase 3: Flight schedule adjustment by expert,
- Phase 4: In-season input data update (this step can affect flight schedule also),
- Phase 5: Practice schedule recommendation,
- Phase 6: Practice schedule adjustment by expert.

System's Decisional Guidance

The way a scheduling DSS leads users to make a decision is referred to as decisional guidance (Morana et al., 2014; Schelling and Robertson, 2020), which considers factors such as:

- What aspect of the scheduling process the system is supporting (i.e., exploration or decision),
- How explicit the output of the scheduling system is based on its delivered knowledge (i.e., description or recommendation),
- When the scheduling system provides the outcome (i.e., real-time, prospectively, or retrospectively),
- How flexible the scheduling system is (i.e., pre-defined or interactive),
- What the users' level of knowledge on scheduling and on the DSS itself is (i.e., expert or novice),
- How the output is delivered (i.e., text, tables, graphs, or image), and
- How the scheduling system is invoked (i.e., on-demand or automatically).

Appropriate decisional guidance should allow some flexibility while minimizing users' confusion and facilitating the understanding of the recommendation given by the DSS (Silver, 1991; Montazemi et al., 1996). Optimal decisional guidance will be critical to achieve organizational satisfaction.

Table 1 shows three examples of scheduling DSS with different decisional guidance considerations. Example 1 represents a non-interactive DSS built for a one-time schedule descriptive analysis. Example 2 shows a non-interactive DSS developed to give a recommendation on flight scheduling for the entire regular season before it starts. Example 3 represents a daily DSS, automatically invoked throughout the season, which recommends daily practice schedule for the upcoming 7 days. The daily schedule can include the roster availability (**Figure 3**), the official competitive calendar, a recommendation for load distribution (**Figure 4**), and a training session load estimator (**Figure 5**).

Data visualization and user interface are powerful decisional guidance tools with tremendous potential in supporting complex

TABLE 1 | Example of various decision support systems with different decisional guidance considerations.

DSS' decisional guidance considerations	Example (1) non-interactive DSS for a one-time schedule descriptive analysis	Example (2) non-interactive DSS for one-time flight schedule before the season starts	Example (3) automatically invoked DSS for daily practice schedule for the upcoming 7 days
(1) Overall goal	One-time research	Once-a-year automation	Daily automation
(2) Influenced aspect of decision-making	Overall schedule exploration	Flight schedule selection	Daily practice schedule selection
(3) Delivered knowledge	Information	Recommendation	Information
(4) Output timing	Prospective or retrospective	Prospective	Real-time
(5) Mode	Pre-defined	Pre-defined	Interactive
(6) User's knowledge	Novice	Intermediate	Expert
(7) Communication	Table, graphs, and map	Table, graphs, and text	Table and graphs
(8) Invocation	On-demand	On-demand	Automatic

For further reading see Morana et al. (2014) and Schelling and Robertson (2020).

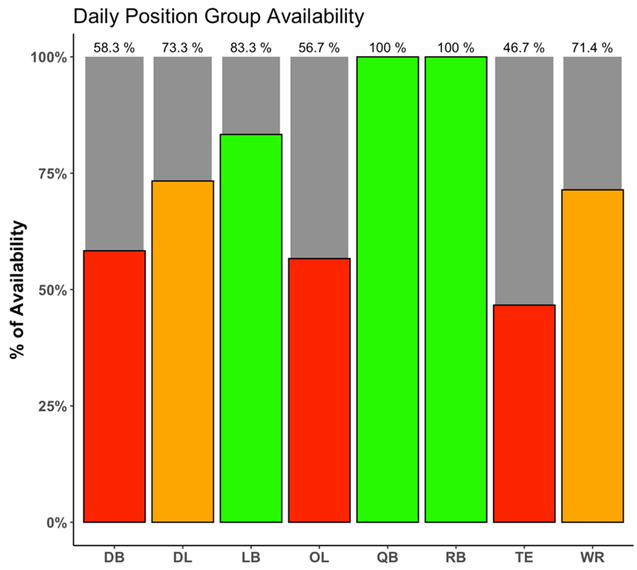

FIGURE 3 | An example of a player availability report for American football, which allows coaches and staff to quickly determine which position groups have a substantial number of players unavailable for full practice, warranting a potential change in the training plan.

decision-making (Zhang and Zhu, 1998). Excellence in statistical graphics consists of complex ideas communicated with clarity, precision, and efficiency. Graphical displays should show the data; avoid distorting what the data have to say; induce the viewer to think about the substance in the project; present many numbers in a small space; make large data sets coherent; encourage the eye to compare different pieces of data; reveal the data at several levels of detail, from a broad overview to the fine structure; serve a clear purpose: decoration, description, exploration, tabulation, or recommendation; and to be closely integrated with the statistical descriptions of a data set (Tufte, 1983). Common visualization tools include charts, diagrams, drawings, graphs, ideograms, pictograms, data plots, schematics, tables, illustrations, and maps or cartograms. In scheduling-related problems there are several recurrent visualizations.

When the goal of the DSS is calendar exploration (Example 1 in **Table 1**), one needs to contextualize the schedule and to let the expert judge if it is good or bad compared to the rest of the teams and to previous seasons. An example would be to visualize an optimization indicator such as games played per month comparing a team against the rest of the teams, showing previous seasons as well (**Figure 6**). For a non-interactive DSS recommender (Example 2 in **Table 1**), visualizing how the optimization indicator such as distance traveled or days away compares to flight schedules from previous seasons (**Figure 7**) would give context for the calendar demands and the DSS' output quality. In an interactive DSS recommender (Example 3 in **Table 1**), visualizations could show how the modifications made by the user affect the optimization indicator, which can be multiple. For instance, changing a flight date or itinerary may increase the days away, the distance traveled, or the recovery or training opportunity (**Figure 8**).

In addition to calendar exploration and optimization of travel schedules, training periodization is critically important in sports with more training opportunity, with in-season micro-cycles of typically 3–7 days in duration (Akenhead et al., 2016). Coaches and support staff must not only consider the technical and tactical objectives, but also the positive (improved fitness) and negative (increased fatigue) consequences of successive training session, including pre-defined optimization indicators, and the net result on gameday (Morton et al., 1990). As recently identified by practitioners (Cross et al., 2019), there seems to be disparity between the available scientific evidence and current industry practice (i.e., human bias) in regards scheduling of training and recovery. It is here where a DSS is useful as it can provide objective contextual information and recommendations that allow practitioners to have a load distribution overview for the upcoming micro-cycle (**Figure 4**) as well as to prescribe training sessions (**Figure 5**) considering individual needs within a team structure (i.e., reactive scheduling). This process will be mainly constrained by the competitive calendar (e.g., number of games, location, day of the week, and time) (Akenhead et al., 2016) and the players' availability (Hagglund et al., 2013). Information from the Athlete Management System (AMS) can be retrieved to determine which players are injured and will be unavailable for training in the upcoming week, which players need additional recovery time following the last game, and which players are able to participate in full. Codifying these details allows the staff to identify training loads and position groups that may be challenged to have enough players available to train on a given day. Such information can be reflected in a dashboard or web application, allowing coaches to make any necessary changes to the weekly training plan should certain positional groups be at risk due to a limited number of players being able to participate in full (**Figure 3**). Finally, once the micro-cycle structure has been designed and the available players identified, a customized session load estimator can be used to help adjust the practice and make it more appropriate considering the micro-cycle load distribution and the available players. **Figure 5** shows an example of a session load estimator that allows the support staff to build the training session with the coach and manipulate the drill duration to automatically get an estimation of player load [e.g., the sum of instantaneous rate of change of acceleration, or jerk, divided by a scaling factor (Nicolella et al., 2018)] for a given session. Tools such as this aid the decision making of the staff as drills can be removed or added from the session and training duration for a specific drill can be altered to gain an understanding of the potential training demands on a position group or individual for the upcoming session.

SCHEDULING MODELS

Traditional approaches to solving scheduling problems use either simulation models, analytical models, heuristic approaches or a combination of these methods (Aytug et al., 1994; Balas, 1999; Mahapatra et al., 2017). Simulation models are

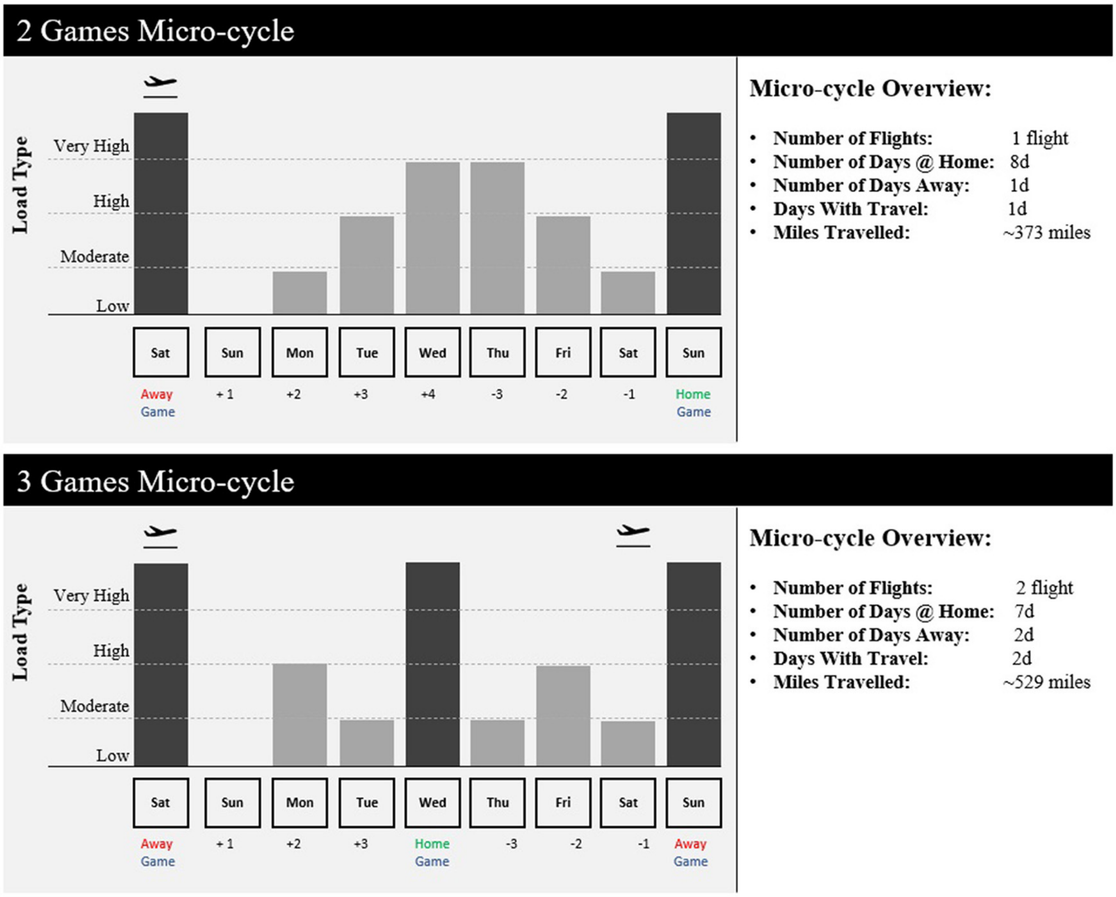

FIGURE 4 | Examples of visualization of micro-cycle load distribution in soccer with different competitive calendar constraints and outputs (number of flights, number of games, number of days off, number of practice days, etc.).

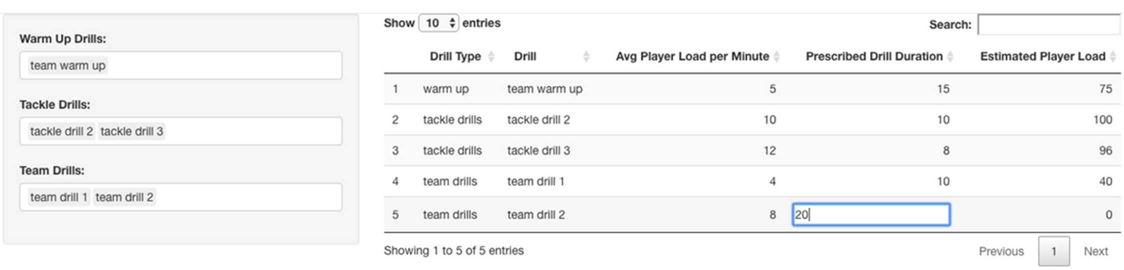

FIGURE 5 | An example of session load estimator that allows the staff to build a training plan with the coach. The staff can change the drill types and manipulate the drill duration to obtain an estimation for Player Load, allowing the coaching staff to make changes to the training session for an individual athlete or position group depending on what they are able to tolerate for a given day.

primarily used to assess schedules and are most useful for schedule exploration (e.g., initial competitive calendar analysis and tentative flight dates) (Aytug et al., 1994). Analytical models include mathematical programming models, stochastic models, and control theory approaches focusing on optimization processes. A disadvantage of these models is that the problem needs to be explicitly formulated, which is difficult for schedulers who do not have the mathematical knowledge and background (Zhou et al., 2013). Additionally, since even the most simplified scheduling problems are complex, realistically sized problems cannot be optimally solved, and real-life applications of analytical approaches are scarce (Aytug et al., 1994). Consequently, a wide body of heuristic approaches have been investigated to find near-optimal solutions in cases where finding the optimal solution is impractical (Zhou et al., 2013; Mahapatra et al., 2017). Some research has shown that human interactions with automated

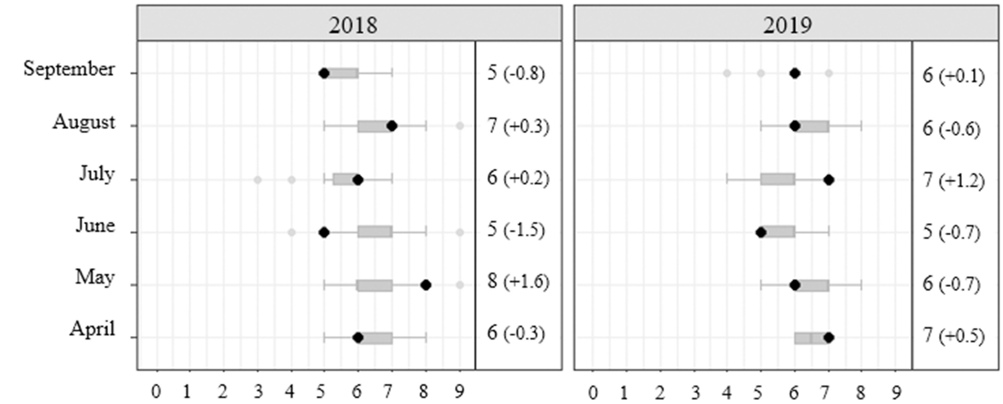

FIGURE 6 | Example of visualization to explore the number of games per month (x-axis) for a Major League Baseball team (black dot) compared to the distribution of all teams in the league (gray boxplot). The differences between the team and the league's average are shown in parentheses.

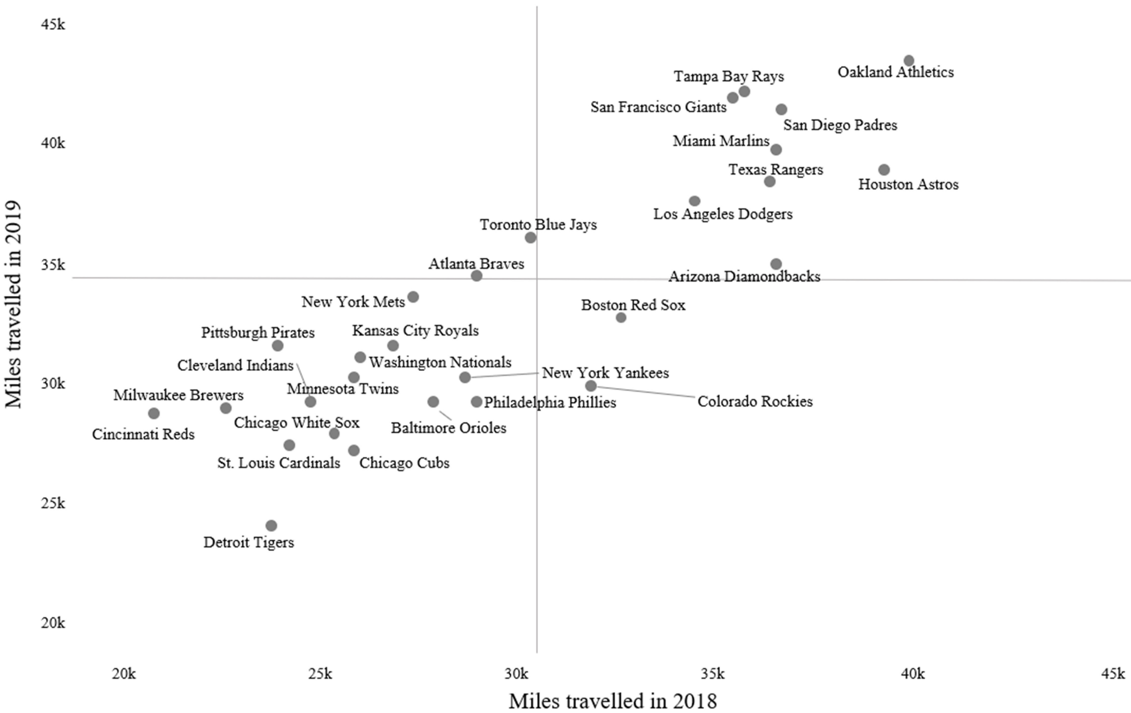

FIGURE 7 | Example of a visualization representing estimated mileage traveled by Major League Baseball teams over two consecutive seasons. Reference lines represent the average mileage traveled for 2019 (horizontal reference line) and 2018 (vertical reference line).

heuristics methods often offer improved performance (Aytug et al., 1994). Computer-based systems are better than humans at finding complex and subtle patterns in massive data sets, but humans are very effective connecting different sources of information in creative and unpredictable ways (Akata et al., 2020). DSS offers a mean to combine various types of knowledge in a manner that can be used for scheduling problems (Schelling and Robertson, 2020).

Expert systems (ES) represent a special case of knowledge-based scheduling DSS (Aytug et al., 1994). ES are developed by first acquiring the knowledge from a human expert and then codifying this knowledge into a series of algorithmic rules (**Figure 9**). Scheduling ES can recommend decisions on actual or simulated cases and do so in a way that captures the idiosyncratic nature of a specific organization. Nevertheless, many researchers (Aytug et al., 1994) believe that expert system approaches are not ideal for scheduling because most real-life environments present complex relationships that are often difficult to model with simple association rules. Two additional issues are that most environments are so dynamic that knowledge becomes obsolete too fast (Fox and Smith, 1985), and that the input of a small set of experts might

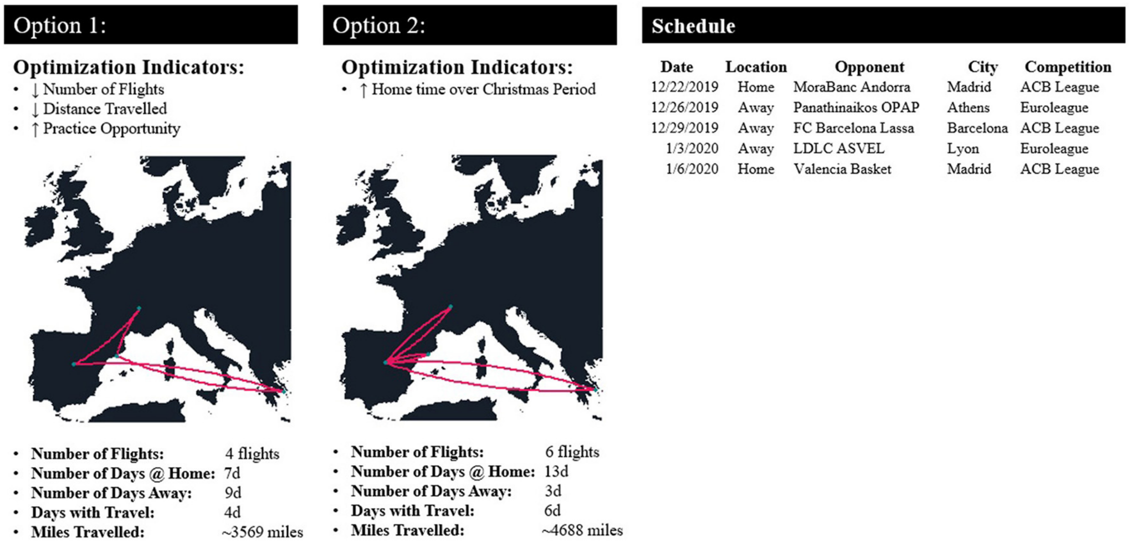

FIGURE 8 | Example representing the addition of multiple "optimization indicators" and its impact on the DSS for a professional Basketball team competing in the domestic and European league over the Christmas period.

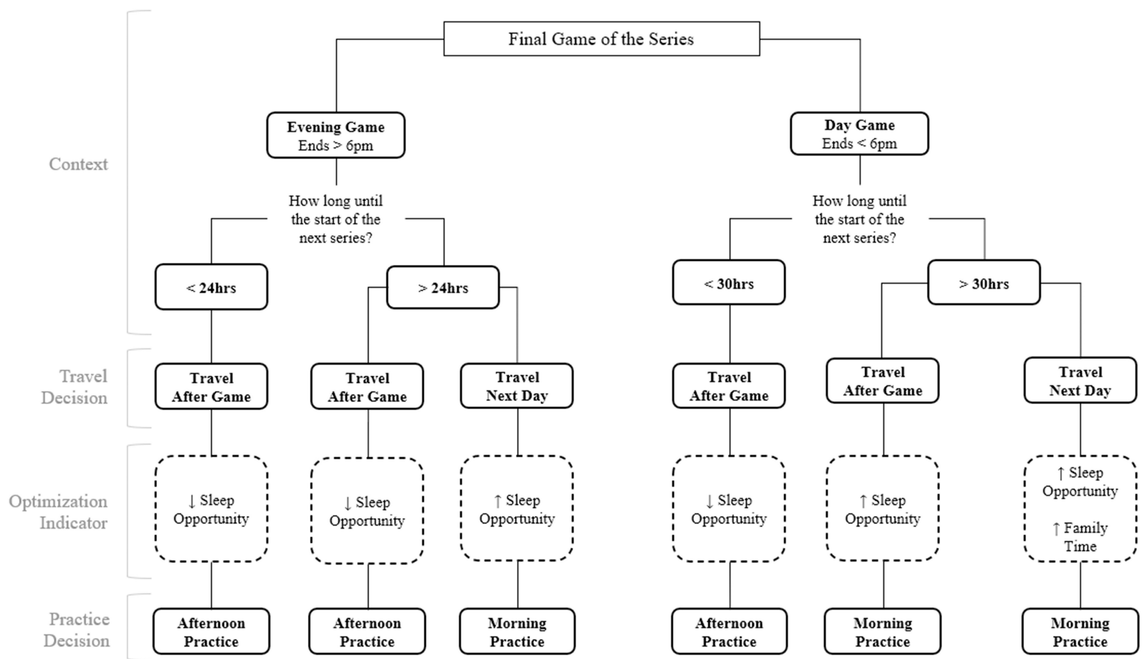

FIGURE 9 | Example of expert system (ES) applied to Major League Baseball.

focus too strongly on specific individual experience, hindering the generalization capabilities of the model. Consequently, more advanced computer-based approaches such as random search, blind search or heuristic search have been implemented for scheduling problems. Constraint-based heuristic search are methods that use knowledge about the restrictions, or constraints, of the scheduling problem to guide and limit the search of a near-optimal solution within a search space that is too large to explore entirely (Trick et al., 2012). Nevertheless, a limitation of many computer-based methods in scheduling is their inability to adapt to changing demands without human-intensive intervention. This observation has led to including learning components in scheduling DSS. Machine learning methods focus on learning from experience to provide predictions on yet-unobserved data, without requiring human intervention in the learning process, and, in many cases, being able to adapt when new data is available.

For the scheduling problem in sports, both supervised (e.g., regressions, decision trees, support vector machine, K-nearest, random forest) and unsupervised (e.g., clustering, PCA) machine learning algorithms could provide a mechanism for creating better features to be used as input for the scheduling DSS (see Song et al., 2019 for more on the interaction between machine learning and optimization processes). Some examples of richer features include the difficulty level estimation of a game, the estimation of a team's carry-over effect throughout the season or discretizing continuous variables that are difficult to model within a DSS such as player load (see the three sub-models in **Figure 2**).

Besides the computational complexities and requirements, the desired decisional guidance discussed in the previous section, requires several design considerations when choosing the analytical processes and techniques embedded in the system. The system's acceptance and its outcome interpretability will be related to the selected model architecture (Ribeiro et al., 2016). Selection of one family of algorithm over another may also change, when possible, the way in which the problem is framed for the end user (Schelling and Robertson, 2020). The scheduling DSS should aim for the most efficient and effective analytical process to solve a task while it meets the interpretability and the operational functions expected by the end-user. Developers need to design a DSS that can provide an understanding of any discrepancy between the DSS recommendation and the expert's opinion (identification of expert bias) (Kayande et al., 2009). Many standard machine learning algorithms such as logistic regression, decision trees, decision-rules learning, or K-nearest neighbors are examples of more interpretable algorithms, whereas random forest, gradient boosting, support vector machine, neural networks and deep learning fall into the less- or non-interpretable machine learning approaches (i.e., black-box algorithms) (Luo et al., 2019). When a black-box model produces significantly better recommendations than a more interpretable model, the scheduling DSS developer may consider integrating feedback within the system (Kayande et al., 2009), with tools such as partial dependence (PD) plots, individual conditional expectation (ICE), local interpretable model-agnostic explanation (LIME), or kernel Shapley values (SHAP) to help partially understand the scheduling recommendation and to ensure trust and transparency in the decision process of the model (Messalas et al., 2019). On the other hand, if there are no specific design needs of relying on the mentioned black-box methods as the main model for the DSS their capacity of exploiting non-linear relationships could still be used to derive richer features, such as the ones mentioned above. Another data-based approach that could provide a good balance between interpretability and prediction accuracy is the use of probabilistic graphical models (e.g., Bayesian networks), which would allow practitioners to obtain a clearer idea of the relationship between the different variables within the DSS and inspect the impact that one decision might have in the rest of the variables. A potential issue of probabilistic outputs and visualizations is that humans generally have more difficulty understanding these than frequency-based data with familiar units (Tversky and Kahneman, 1983).

SCHEDULING DECISION SUPPORT SYSTEM EVALUATION

When it comes to evaluate how the scheduling DSS is performing, three overarching aspects need to be reviewed: context satisfaction, process efficiency, and output quality. The first consideration refers to how satisfied the organization is with the system (e.g., is the DSS covering the organization's needs? is it technically and economically feasible?). The second aspect refers to the efficiency of the process (e.g., is the DSS user-friendly? Is the recommendation given by the DSS what the end-user expected? Is the complexity of the model adequate? Is the interpretation of the recommendation clear for the user?). The third and last criterion relates to the quality of the recommendation (e.g., is the recommended schedule been followed on its entirety by the organization? if not, how many instances have been modified? if there was an optimization indicator, did the DSS' recommendation improve historical decisions? is the DSS capable of learning based on the expert modifications?). Based on these three considerations a comprehensive DSS evaluation tool has been previously published (Schelling and Robertson, 2020), which includes feasibility, decisional guidance, data quality, system complexity, and system error as the assessment components. Nevertheless, assessing a scheduling system's error might seem cumbersome, but as discussed on the section on decisional guidance, assessing the system's output quality will require a subjective and an objective perspective. For instance, **Figure 8** shows two scheduling options based on different optimization indicators (physiological and psychological). The expert will find more suitable one option than the other for the team's context. Visualizing the degree of agreement between the scheduling DSS recommendation and the expert's decision can help evaluating the overall DSS recommendation quality, in addition to the analysis of the optimization indicators when the DSS recommendation are changed. Future research should include analyzing the efficacy of scheduling DSS on enhancing decision-making processes and key performance indicators (KPIs).

CONCLUSION

A scheduling decision support system can enhance a schedule better than a human-judgment-only approach primarily by automating certain or all processes, by objectively weighing constraints in the schedule (i.e., optimization), and allowing systematic historical comparisons, particularly if personnel changes occur. Scheduling DSS can include predictive and exploratory solutions for macroplanning (e.g., competitive calendar analysis and tentative travel schedule), and reactive solutions for microplanning (e.g., weekly session prescription and travel updates). These solutions must consider several contextual constraints (fixed and dynamic) and provide the nearest-optimal solution, since an optimal solution might not be feasible due to contextual requirements or computational complexity. Constraints and optimization indicators, as well as the advantages of the DSS adoption may differ between organizations. An integrative understanding of current scheduling practices and the organization's needs prior to the development of the DSS is warranted. Traditional approaches to solving scheduling problems use either simulation models, analytical or mathematical models, heuristic approaches, or a combination of these methods. Machine learning algorithms (supervised and unsupervised) could provide a mechanism for creating better features to be used as input (e.g., game difficulty, carry-over effect, and discretization of continuous variables) or for reducing data dimensionality (i.e., variable selection). For a better acceptance and a successful implementation, the scheduling DSS recommendation process should be as understandable as possible. Visualization techniques might be required to improve the system's interpretability. Once implemented, the system's recommendations (output) and the users' feedback (interaction) can be closely and systematically monitored for eventual improvements.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

XS: conception, design, drafting, critical revision, visuals, and final approval of the papers' version to be published. SR: critical revision and final approval of the papers' version to be published. JF, PW, and JF: critical revision, feedback, and visuals. All authors contributed to the article and approved the submitted version.

REFERENCES

Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., et al. (2020). A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer* 53, 18–28. doi: 10.1109/MC.2020.2996587

Akenhead, R., Harley, J., and Tweddle, S. (2016). Examining the external training load of an English Premier League Football Team with special reference to acceleration. *J Strength Cond Res*. 30, 2424–2432. doi: 10.1519/JSC.0000000000001343

Aytug, H., Bhattacharyya, S., Koehler, G., and Snowdon, J. (1994). A review of machine learning in scheduling. *IEEE Trans. Eng. Manag.* 41, 165–171. doi: 10.1109/17.293383

Balas, E. (1999). New classes of efficiently solvable generalized Traveling Salesman Problems. *Ann. Oper. Res*. 86, 529–558. doi: 10.1023/A:1018939709890

Byl, J. (2013). *Organizing Successful Tournaments,* 4th edn. Champaign, IL: Human Kinetics.

Cross, R., Siegler, J., Marshall, P., and Lovell, R. (2019). Scheduling of training and recovery during the in-season weekly micro-cycle: insights from team sport practitioners. *Eur J Sport Sci*. 19, 1287–1296. doi: 10.1080/17461391.2019.1595740

Drew, M., and Finch, C. (2016). The relationship between training load and injury, illness and soreness: a systematic and literature review. *Sports Med*. 46, 861–883. doi: 10.1007/s40279-015-0459-8

Esteves, P., Mikolajec, K., Schelling, X., and Sampaio, J. (2020). Basketball performance is affected by the schedule congestion: NBA back-to-backs under the microscope. *Eur. J. Sport Sci*. 21, 26–35. doi: 10.1080/17461391.2020.1736179

Fox, M. (1990). Constraint guided scheduling: a short history of scheduling research at CMU. *Comput. Ind.* 14, 79–88. doi: 10.1016/0166-3615(90)90107-Z

Fox, M., and Smith, S. (1985). "The role of intelligent reactive processing in production management," in *13th Annual CAMI Technical Conference* (Pittsburgh, PA).

Fullagar, H., Duffield, R., Skorski, S., White, D., Bloomfield, J., Kölling, S., et al. (2016). Sleep, travel, and recovery responses of national footballers during and after Long-Haul International Air Travel. *Int. J. Sports Physiol. Perform*. 11, 86–95. doi: 10.1123/ijspp.2015-0012

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., and Fdez-Riverola, F. (2014). Web scraping technologies in an API world. *Brief Bioinform*. 15, 788–797. doi: 10.1093/bib/bbt026

Goldreich, O. (2008). *Computational Complexity: A Conceptual Perspective*. Cambridge: Cambridge University Press.

Goossens, D., and Spieksma, F. (2012). Soccer schedules in Europe: an overview. *J. Scheduling*. 15, 641–651. doi: 10.1007/s10951-011-0238-9

Guedes, A., and Ribeiro, C. (2011). A heuristic for minimizing weighted carry-over effects in round robin tournaments. *J. Scheduling* 14, 655–667. doi: 10.1007/s10951-011-0244-y

Hagglund, M., Walden, M., Magnusson, H., Kristenson, K., Bengtsson, H., and Ekstrand, J. (2013). Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA Champions League injury study. *Br J Sports Med*. 47, 738–742. doi: 10.1136/bjsports-2013-092215

Holmes, B. (2018). *How Fatigue Shaped the Season, and What It Means for the Playoffs*. Available online at: https://www.espn.com/nba/story/_/id/23094298/how-fatigue-shaped-nba-season-means-playoffs (accessed May 07, 2020).

Karney, C. (2013). Algorithms for geodesics. *J. Geodesy* 87, 43–55. doi: 10.1007/s00190-012-0578-z

Kayande, U., De Bruyn, A., Lilien, G., Rangaswamy, A., and van Bruggen, G. (2009). How incorporating feedback mechanisms in a DSS affects DSS evaluations. *Inf. Syst. Res*. 20, 527–546. doi: 10.1287/isre.1080.0198

Kendall, G., Knust, S., Ribeiro, C., and Urrutia, S. (2010). Scheduling in sports: an annotated bibliography. *Comput. Oper. Res*. 37, 1–19. doi: 10.1016/j.cor.2009.05.013

Kenett, R., and Shmueli, G. (2016). "Dimensions of information quality and InfoQ assessment," in *The Potential of Data and Analytics to Generate Knowledge*, eds. R. Kenett, and G. Shmueli (Chichester, West Sussex: WILEY), 32–53.

Kloke, J. (2016). *The NHL Season Is Too Damn Long*. Available online at: https://www.vice.com/en_ca/article/wnvepb/the-nhl-season-is-too-damn-long (accessed May 07, 2020).

Lastella, M., Roach, G., and Sargent, C. (2019). Travel fatigue and sleep/wake behaviors of professional soccer players during international competition. *Sleep Health* 5, 141–147. doi: 10.1016/j.sleh.2018.10.013

Lewis, M. (2018). It's a hard-knock life: game load, fatigue, and injury risk in the National Basketball Association. *J. Athletic Train*. 53, 503–509. doi: 10.4085/1062-6050-243-17

Luo, Y., Tseng, H., Cui, S., Wei, L., Ten-Haken, R., and El-Naqa, I. (2019). Balancing accuracy and interpretability of machine learning approaches for radiation treatment outcomes modeling. *BJR Open* 1:20190021. doi: 10.1259/bjro.20190021

Mack, C., Herzog, M., DiFiori, J., Meisel, P., and Dreyer, N. (2018). A second look at NBA game schedules: response to Teramoto et al. *J. Sci. Med. Sport* 21, 228–229. doi: 10.1016/j.jsams.2017.07.023

Mahapatra, S., Dash, R., and Pradhan, S. (2017). "Heuristics techniques for scheduling problems with reducing waiting time variance," in *Heuristics and Hyper-Heuristics - Principles and Applications*, ed. J. Del Ser Lorente (London: IntechOpen Limited), 43–64.

Messalas, A., Kanellopoulos, Y., and Makris, C. (2019). "Model-agnostic interpretability with shapley values," in *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)* (Patras).

Mitchell, M., Wathen, S., and Orwig, R. (2019). Evaluating the two-game road trip in college sports: does a travel partner scheduling approach affect team competitiveness? *Sport J*. 22, 1–17. Available online at: https://bit.ly/3bhm7xk

Montazemi, A., Wang, F., Khalid, S., and Barta, C. (1996). On the effectiveness of decisional guidance. *Decis. Support Syst*. 18, 181–198. doi: 10.1016/0167-9236(96)00038-3

Morana, S., Schacht, S., Scherp, A., and Maedche, A. (2014). "Conceptualization and typology of guidance in information systems," in *Working Paper Series in Information Systems*, Vol. 7 (Mannheim: University of Mannheim), 1–13.

Morton, R., Fitz-Clarke, J., and Banister, E. (1990). Modeling human performance in running. *J. Appl. Physiol*. 69, 1171–1177. doi: 10.1152/jappl.1990.69.3.1171

Moskowitz, T., and Wertheim, L. (2011). *Scorecasting: The Hidden Influences Behind How Sports Are Played and Games are Won*. New York, NY: Three Rivers Press.

Nicolella, D., Torres-Ronda, L., Saylor, K., Schelling, X. (2018). Validity and reliability of an accelerometer-based player tracking device. *PLoS One* 13:e0191823. doi: 10.1371/journal.pone.0191823

Nutting, A., and Price, J. (2017). Time zones, game start times, and team performance:evidence from the NBA. *J. Sports Econ*. 18, 471–478. doi: 10.1177/1527002515588136

Ribeiro, C. (2012). Sports scheduling: problems and applications. *Int. Trans. Oper. Res*. 19, 201–226. doi: 10.1111/j.1475-3995.2011.00819.x

Ribeiro, M., Singh, S., and Guestrin, C. (2016). "Why should i trust you?: explaining the predictions of any classifier," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, CA).

Robertson, S., and Joyce, D. (2018). Evaluating strategic periodisation in team sport. *J. Sports Sci*. 36, 279–285. doi: 10.1080/02640414.2017.1300315

Rocha, L. (2017). Dynamics of air transport networks: a review from a complex systems perspective. *Chin. J. Aeronautics* 30, 469–478. doi: 10.1016/j.cja.2016.12.029

Rossi, A., Pappalardo, L., Cintia, P., Iaia, F., Fernandez, J., and Medina, D. (2018). Effective injury forecasting in soccer with GPS training data and machine learning. *PLoS One* 13:e0201264. doi: 10.1371/journal.pone.0201264

Schelling, X., and Robertson, S. (2020). A development framework for decision support systems in high-performance sport. *Int. J. Comp. Sci. Sport* 19, 1–23. doi: 10.2478/ijcss-2020-0001

Seirul·lo, F. (1998). *Planificación a Largo Plazo en los Deportes Colectivos. Curso sobre Entrenamiento Deportivo en la Infancia y la Adolescencia.* Available online at: http://www.motricidadhumana.com/seirul_planif_dep_colectivos.pdf (accessed May 07, 2020).

Shelburne, R. (2017). *Adam Silver: Resting Star Players 'A Significant Issue for the League'*. Available online at: https://www.espn.com/nba/story/_/id/18962901/resting-star-players-significant-issue-league (accessed May 07, 2020).

Silver, M. (1991). Decision guidance for computer based decision support. *MIS Q*. 15, 105–122. doi: 10.2307/249441

Song, H., Triguero, I., and Özcan E. (2019). A review on the self and dual interactions between machine learning and optimisation. *Prog. Artif. Intell*. 8, 143–165. doi: 10.1007/s13748-019-00185-z

Sortino, B. (2015). The impact of training and competition scheduling on the effectiveness of sleep, recovery, and performance. *J. Aust. Strength Cond*. 23, 68–75.

Sport, B. (2020). *Premier League Fixture Congestion 'Absurd' As Teams Count Christmas Cost.* Available online at: https://www.bbc.com/sport/football/50986483 (accessed May 07, 2020).

Teramoto, M., Cross, C., Cushman, D., Maak, T., Petron, D., and Willick, S. (2017). Game injuries in relation to game schedules in the National Basketball Association. *J. Sci. Med. Sport* 20, 230–235. doi: 10.1016/j.jsams.2016.08.020

Trick, M., Yildiz, H., and Yunes, T. (2012). Scheduling major league baseball umpires and the traveling umpire problem. *Interfaces* 42, 232–244. doi: 10.1287/inte.1100.0514

Tufte, E. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Tversky, A., and Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychol Rev*. 90, 293–315. doi: 10.1037/0033-295X.90.4.293

Wang, W., Mok, A., and Fohler, G. (2003). *Pre-Scheduling: Integrating Offline and Online Scheduling Techniques.* Berlin; Heidelberg: Springer.

Zhang, P., and Zhu, D. (1998). *Information Visualization in Project Management And Scheduling*. Former Departments, Centers, Institutes and Projects, 55.

Zhou, J., Love, P., Wang, X., Teo, K., and Irani, Z. (2013). A review of methods and algorithms for optimizing construction scheduling. *J. Oper. Res. Soc*. 64, 1091–1105. doi: 10.1057/jors.2012.174

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

frontiers in Sports and Active Living

ORIGINAL RESEARCH
published: 09 August 2021
doi: 10.3389/fspor.2021.715833

Unilateral Maximal Isometric Hex Bar Pull Test: Within-Session Reliability and Lower Body Force Production in Male and Female Freeski Athletes

Jonathan McPhail[1], Basílio A. M. Gonçalves[2], Jörg Spörri[3,4] and Vesa Linnamo[1]*

[1] Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland, [2] School of Allied Health Sciences, Griffith University, Brisbane, QLD, Australia, [3] Sports Medical Research Group, Department of Orthopaedics, Balgrist University Hospital, University of Zurich, Zurich, Switzerland, [4] University Centre for Prevention and Sports Medicine, Department of Orthopaedics, Balgrist University Hospital, University of Zurich, Zurich, Switzerland

OPEN ACCESS

Edited by:
Zimi Sawacha, University of Padua, Italy

Reviewed by:
Fabiola Spolaor, University of Padua, Italy
Danny Lum, Sport Singapore, Singapore

***Correspondence:**
Jonathan McPhail
jonathanmcphail1989@gmail.com

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: *27 May 2021*
Accepted: *12 July 2021*
Published: *09 August 2021*

Citation:
McPhail J, Gonçalves BAM, Spörri J and Linnamo V (2021) Unilateral Maximal Isometric Hex Bar Pull Test: Within-Session Reliability and Lower Body Force Production in Male and Female Freeski Athletes. Front. Sports Act. Living 3:715833. doi: 10.3389/fspor.2021.715833

The aim of the study was to (1) assess the within-session reliability of a unilateral isometric hex bar pull (UIHBP) maximal voluntary contraction (MVC) test and, (2) determine unilateral isometric absolute peak force (PF_{abs}) and relative peak force (PF) values in freeski athletes. Twenty-one male and eight female academy to national team freeskiers performed the novel UIHBP MVC task on a force plate and PF_{abs} and relative PF were assessed (1000 Hz). Within-session measures of PF_{abs} offered high reliability on left and right limbs for males (*ICC* = 0.91–0.94, *CV* = 2.6–2.2%) and females (*ICC* = 0.94–0.94, *CV* = 1.4–1.6%), while relative PF measures showed good to high reliability in both left and right limbs for males (*ICC* = 0.8–0.84, *CV* = 2.6–2.2%) and females (*ICC* = 0.92–0.90, *CV* = 1.4–1.7%). We observed significantly lower PF_{abs} ($p < 0.001$) and relative PF ($p < 0.001$) in females compared to males. No statistical difference was found between left and right limbs in males and females in PF_{abs} ($p = 0.98$) and relative PF measures ($p = 0.93$). The UIHBP MVC test appears to be a reliable method for assessing PF_{abs} and relative PF in male and female freeski athletes.

Keywords: freeskiing, skiing, strength testing, unilateral, isometric, maximal, voluntary, contraction

INTRODUCTION

Freeskiing is an extremely complex skill-based action sport that involves numerous technical, tactical, and psychophysical demands (Willmott and Collins, 2015). There are three freeski disciplines (slopestyle, big air and half-pipe). In freeski slopestyle, athletes perform a series of tricks using jumps, custom built rails, and other creative features such as quarter pipes. Freeski big air is performed using only one large jump and competitors perform complex tricks in the air, aiming for high amplitude, style, creative grabs, and a clean landing. During freeski half-pipe, 6–8 tricks are performed whilst skiing down a u-shaped pipe. Accordingly, there are many psychological, skill acquisition and physical factors that may influence the performance, skill execution and safety in freeskiing. To date there is no evidence-based consensus on reliable and practically meaningful physical testing protocols that could be used for screening and monitoring freeski athletes in the context of performance enhancement, injury prevention and/or rehabilitation.

Generally, periodic testing and monitoring of an athlete's neuromuscular performance at several stages during the year can be considered an effective way to provide useful information to practitioners concerning an athletes' current training state (Edwards et al., 2018). This data can be combined with an appreciation and understanding of the emotional load action sport athletes experience (Collins et al., 2018). There is, however, a paucity of data regarding what physical qualities are considered important for freeski athletes from both a supporting performance and injury risk mitigation standpoint. Nevertheless, in certain contexts, maximal strength is plausibly an important capacity to develop for potentially preventing acute and overuse injuries (Lauersen et al., 2018) and is, in certain athletic settings, known to be moderately associated with jump and sprint performance (Kirkpatrick and Comfort, 2013; Comfort et al., 2014). Furthermore, possessing greater lower body strength has been deemed advantageous in other snow sports such as snowboard cross and alpine snowboarding (Vernillo et al., 2016) and alpine skiing (Cross et al., 2021). Nonetheless, a degree of caution should be given when drawing the same conclusions to freeskiing without proper investigation of kinetics, kinematics, and individuals factors such as riding style.

With respect to testing methods, there are numerous approaches to assess athletes' maximal force capabilities. For example, the gold standard method to assess knee flexor and extensor strength is with a motor-driven isokinetic dynamometer (Knapik et al., 1983; Ly and Handelsman, 2002). Isokinetic dynamometry is recommended as it can elicit maximal efforts over a full range of motion (Caruso et al., 2012) and can be used to assess neuromuscular function through different parameters such as peak torque, total work or the peak torque ratio between agonist and antagonist muscles (Gleeson and Mercer, 1996; Bosquet et al., 2016). However, this method is expensive, time consuming and is often impractical in many instances for freeskiers, especially in-season and when testing a group.

Alternative methods to reliably assess maximal strength, are the one-repetition maximum (1 RM) test (Grgic et al., 2020) or via isometric maximal voluntary contraction (MVC) testing at specific joint angles (Drake et al., 2017). In the 1 RM test, eccentric muscles actions are often coupled with concentric actions which can be more reflective of dynamic muscle actions that occur in resistance training and sporting actions (Grgic et al., 2020). Contrasting with isokinetic dynamometry, the 1 RM test is highly cost effective, however this form of testing can also be time consuming with groups of athletes and is also often not appropriate in-season for freeskiers. During isometric contractions, the muscle-tendon unit remains at a constant length and can produce more force than a concentric muscle contraction (Abbott and Wilkie, 1953). Isometric contractions have also been shown to result in reduced structural muscle damage compared to eccentric contractions (Nosaka et al., 2003) which makes this approach of assessment popular in applied settings with athletes. However, isometric contractions performed at longer muscle lengths and for sustained durations can increase muscle soreness, damage, and fatigue (Allen et al., 2018). Finally, common isometric MVC modalities include the isometric leg press (Granacher et al., 2011; Bogdanis et al., 2019), isometric knee extension (Kubo et al., 2006; Noorkõiv et al., 2014), isometric squat (Markovic and Jaric, 2004; Eliassen et al., 2018) and isometric mid-thigh pull (West et al., 2011). These methods are often utilized in training interventions investigating neuromuscular responses to exercise (Taipale et al., 2014), exploring mechanisms of fatigue (Izquierdo et al., 2009) and can also be incorporated into rehabilitation processes (Maestroni et al., 2019; Jordan et al., 2020; Taberner et al., 2020). Despite the advantages of the isometric tests listed above, there are limitations to these methods. For example, these tests often require custom built and robust equipment fixed in place in a laboratory which can create challenges for athletes who travel extensively or train in several locations.

Although the demands of freeskiing have not been quantified, when taking into consideration the incidence and location of injuries often occurring to the knee and lower extremities (Flørenes et al., 2010; Steffen et al., 2017; Palmer et al., 2021), assessing unilateral lower body strength could be warranted. Evaluating athletes' force producing capabilities at several stages during the rehabilitation process can help identify and resolve deficits in neuromuscular performance (Maestroni et al., 2019; Taberner et al., 2020). Moreover, monitoring lower limb strength can provide objective information to help guide task progressions and support inter-disciplinary decision making on important functional milestones such as initiating running, jumping and plyometric activity (Palmieri-Smith and Lepley, 2015; Buckthorpe et al., 2020). The hex bar deadlift also referred to as a 'trap bar' has become a popular resistance training exercise to perform and is a variant of the barbell deadlift (Camara et al., 2016; Lake et al., 2017; Andersen et al., 2018). Despite this increased popularity, to the authors' knowledge the hex bar deadlift and unilateral variations have not been utilized in testing via the use of force-platforms. Using a hex bar to assess unilateral isometric MVC could provide practitioners with as an alternative testing method when other methods are not compatible or suit their setting and context. However, before using such methods in a practical setting, it is necessary to determine the level of reliability of a test (McCall et al., 2015).

Based on these considerations, the aims of the present study were to (1) evaluate the within-session reliability of absolute (PF_{abs}) and relative peak force (PF) during a novel unilateral isometric hex bar pull (UIHBP) maximal voluntary contraction (MVC) test in male and female academy to national team freeski athletes and (2) to provide sex- and level specific reference values.

MATERIALS AND METHODS

Subjects

Twenty-nine academy to national team freeski athletes gave their informed consent to participate in the study: twenty-one males aged 20 ± 2.5 years old, 176 cm ± 3.9, 70 kg ± 3.8, eight females aged 21 ± 4.6 years old, 165 cm ± 2.6, 60.3 kg ± 4.6. Only athletes without a history of knee injuries were included in the study. All participants had to have been enrolled in a freeski academy or part of a national team program. Additional eligibility criteria for the study included having had experience of at least six months of organized strength and conditioning training history and being familiarized with the testing procedures. Subjects did not take part in any physical activity in the 48 h prior to testing. The study was approved by the Ethical Committee of the University of Jyväskylä, and it was conducted according to the provisions of the Declaration of Helsinki.

Testing Procedures

Testing was conducted to assess the within-session reliability of a UIHBP MVC task. The duration of testing for each subject was 30–45 min. Subjects undertook a 15 min dynamic warm up, consisting of: 5 min of jogging and skipping, 2 min of dynamic stretching, three sets of 6–8 repetitions of bilateral and unilateral ankle pogo jumps, 2 sets of 8 linear and lateral hop and holds, 1 set of 4 squat jumps, 1 set of 4 bilateral and unilateral countermovement jumps and 3 progressive accelerations of 15–20 meters separated by 1 min rest between each acceleration. Three min of passive rest was provided after the completion of the warmup prior to starting the MVC measurements.

UIHBP MVC Test

A hex bar (27.5 kg) was loaded with 160 kg to ensure it was secured to the ground and would not move when the subjects were performing the test. The height of the hex bar handle when loaded was 335 mm, the circumference of the handle of the bar was 98 mm. The distance width between the two handles was 588 mm. The subjects were told to prepare as if they were performing a bilateral hex bar deadlift from the ground and a handheld goniometer was used to ensure a knee angle of 115° flexion. If required for taller subjects, blocks were placed under the weights to raise the bar and ensure the correct knee angle was maintained. Once subjects were in the correct starting position and were comfortable, they lifted the uninvolved non-weight bearing limb backwards, ensuring their trunk was in the same position throughout and remained still for 2–3 s before completing warm up trials at 70%, 80% and 90% of self-estimated maximal effort (**Figure 1**). A second researcher observed the test being performed and was responsible for confirming whether the participant maintained the previously described position during the test. The test was performed following 3 min of rest after the warmup was complete. The subjects were instructed to "pull against the bar and push into the ground as hard as possible" exerting maximal force for 4 s. Each limb was tested 3 times for a total of 3 trials per leg with 60 s rest between each trial. Testing was performed in a training facility and conducted using a force plate (1000 Hz, HUR labs, Finland). The force plate was calibrated before each independent test. PF_{abs} was defined as the maximum force generated during the test and relative PF as PF_{abs} divided by body mass (kg). Coachtech online measurement and feedback system (University of Jyväskylä, Finland) was used to collect the force data (Ohtonen et al., 2016) and was classified according to left and right legs due to the equilateral nature of skiing.

Statistical Analysis

We confirmed data normality using the Shapiro-Wilk test (Ghasemi and Zahediasl, 2012) and determined within-session reliability (Atkinson and Nevill, 1998) using two-way mixed-effects model ICC with a 95% confidence interval (CI), based on a single measurement (Koo and Li, 2016). Intra-individual coefficient of variation (CV), with a 95% CI, calculated as the average of the CV for each individual where MSE represents the mean squared error across trials and represents the mean of all the trials ($CV = \frac{\sqrt{MSE}}{\bar{x}} \times 100\bar{x}$) (Knutson et al., 1994). Reliability thresholds for ICC values were defined as poor (<0.50), moderate (0.50–0.75), good (0.75–0.90), and excellent (>0.90) (Koo and Li, 2016). For coefficient of variation (CV), a value of $\leq$10% was defined as reliable (Brughelli and Van Leemputte, 2013). Sex and leg differences were analyzed using a mixed model two-way analysis of variance. Bonferroni *post-hoc* tests (pairwise comparisons) were performed if significant interactions between group and time were found (VanderWeele and Mathur, 2019). All statistical analyses were conducted with custom-made scripts in MATLAB (Version R2018a, MathWorks, Natick, MA, USA), and statistical significance was set to $p < 0.05$ and confidence intervals to 95%.

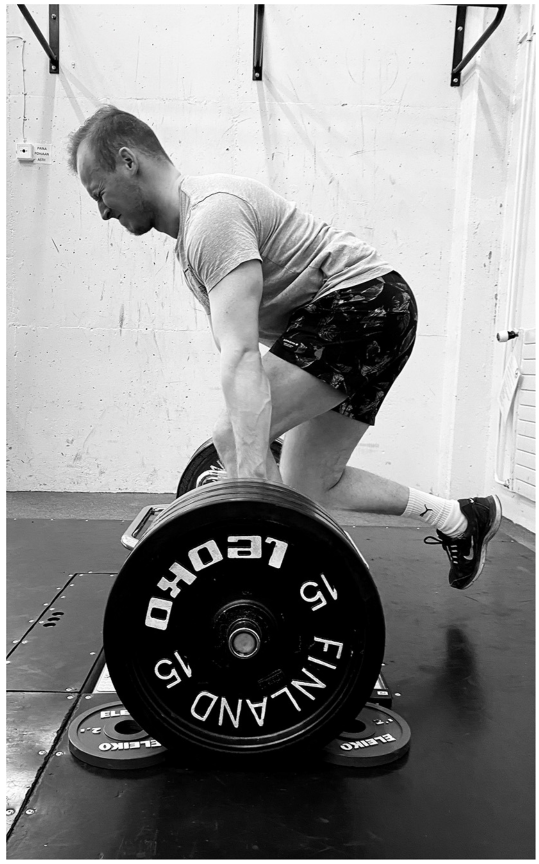

FIGURE 1 | Unilateral maximal isometric hex bar MVC task.

RESULTS

Within-session reliability variables (ICC, CV, SEM, MDC) and descriptive statistics of male and female PF_{abs} and relative PF values are presented in **Tables 1**, **2**. Within-session measures of

TABLE 1 | Within-session reliability measures of unilateral hex bar isometric pull test.

UIHBP	ICC PF_{abs}	ICC Relative PF	CV% PF_{abs}	CV% Relative PF	SEM (N)	MDC (N)
Male left	0.91 (0.80–0.96)	0.8 (0.55–0.91)	2.6% (1.9–3.3)	2.6% (1.8–3.3)	59	162
Male right	0.94 (0.87–0.98)	0.84 (0.62–0.93)	2.2% (1.8–2.7)	2.2% (1.7–2.7)	44	122
Female left	0.94 (0.74–0.99)	0.92 (0.66–0.98)	1.4% (0.8–2)	1.4% (0.78–2.2)	21	68
Female right	0.94 (0.87–0.97)	0.9 (0.55–0.98)	1.6% (0.6–2.5)	1.7% (0.6–2.8)	27	75

ICC, intraclass correlation coefficient; CV, coefficient of variation; PF_{abs}, absolute peak force; PF, peak force; SEM, standard error measurement; MDC, minimal detectable change.

TABLE 2 | Descriptive statistics for unilateral isometric hex bar pull outcome variables.

UIHBP	Mean PF_{abs} (N)	Mean relative PF (N/kg)
Male left	1708.7 ± 183	24.6 ± 1.7
Male right	1697.6 ± 195	24.5 ± 1.7
Female left	1318.9 ± 97	21.5 ± 1.3
Female right	1339.5 ± 106	21.9 ± 1.02

PF_{abs}, absolute peak force; PF, peak force. Data are shown as mean ± 1 standard deviation.

PF_{abs} offered high reliability on left and right limbs for males ($ICC = 0.91$–0.94, $CV = 2.6$–2.2%) and females ($ICC = 0.94$–0.94, $CV = 1.4$–2.2%). Relative PF measures showed good to high reliability in both left and right limbs for males ($ICC = 0.8$–0.84, $CV = 2.6$–2.2%) and females ($ICC = 0.92$–0.90, $CV = 1.4$–1.7%). No significant differences in maximal isometric force were observed between left and right legs in either PF_{abs} ($p = 0.98$) or relative PF ($p = 0.93$) measures (**Figure 2**). Significantly lower maximal isometric force was observed in females compared to males both in PF_{abs} (mean difference (95%CI) $= -376$ N (-279 to -473), $p < 0.001$) and relative PF (mean difference (95%CI) $= -3$ N/kg (-2 to -4), $p < 0.001$) (**Figure 2**). No statistical difference was found between left and right limbs in males and females in PF_{abs} ($p = 0.98$) and relative PF measures ($p = 0.93$).

DISCUSSION

The main finding of this study was that the UIHBP MVC test when performed on a force plate offered good-excellent within-session reliability in PF_{abs} and relative PF in both male and female freeski athletes. This study also provided force production data of the lower body in male and female freeski athletes. It was found that female freeski athletes produced lower PF_{abs} and relative PF when compared to male freeskiers.

Reliable testing methods and protocols are required to confidently detect meaningful changes in performance (Moeskops et al., 2018). The findings from this study offer practitioners a viable option to assess unilateral lower body strength. ICC values between 0.75 and 0.9 indicate good reliability, and values <0.90 indicate excellent reliability (Portney and Watkins, 2009; Koo and Li, 2016). The ICC calculated for the male left and right limbs for PF_{abs} were 0.91–0.94 and for relative PF were 0.8–0-84. The ICC for the female left and right limbs for PF_{abs} was 0.94–0.94 and for relative PF was 0.92–0.90. The CV is a common and robust criterion to test reliability and a CV of ≤10% is often used as the criterion to declare a variable as reliable (Brughelli and Van Leemputte, 2013). The CV calculated for the male left and right limbs for PF_{abs} were 2.6%–2.2% and for relative PF were 2.6–2.2%. The CV for female left and right limbs for PF_{abs} were 1.4%-1.6% and for relative PF were 1.4–1.7%. A force plate is the gold standard for measuring isometric muscle force (Verdera et al., 1999). Therefore, providing practitioners follow the same testing procedures as presented in the current study, they can be confident that they are collecting reliable data from their athletes. The UIHBP MVC test appears to offer equivalent reliability values when compared to the isometric mid-thigh pull and the isometric squat ($ICC = \geq 0.80$ to 0.99) (Drake et al., 2017) and therefore, offers an alternative method for assessing unilateral PF when other methods are not appropriate or feasible.

This is the first study to present data regarding unilateral lower body strength values of male and female freeskiers. As expected, significantly lower maximal isometric force in females compared to males in both PF_{abs} ($p < 0.001$) and relative PF capacities ($p < 0.001$) were observed. The main factors accounting for differences in lower body strength between men and women are likely due to muscle mass (Miller et al., 1993), greater proportion of fast type fibers (Nindl et al., 1995) and morphological characteristics such as muscle thickness, pennation angle, and fascicle length (Blazevich and Sharp, 2005; Bartolomei et al., 2019). Furthermore, although knee flexion was controlled for in this study, hip flexion was not, and this may have contributed to a certain extent to the differences between sexes. There was no statistical difference between left and right limb absolute PF_{abs} ($p = 0.98$) or relative PF ($p = 0.93$) measures. Freeskiing is an equilateral sport characterized by similar physical demands on each leg. However, certain tricks and the initiation of aerial maneuvers and landings often occur using predominantly one leg. Consequently, freeskiers may commonly utilize various forms of unilateral resistance training to seek a desired training adaptation. These findings could therefore be of potential interest for practitioners working with freeski populations. However, unilateral strength measurement values recorded in the laboratory may not be correlated to the ground reaction force kinetics in the sport (Ogrin et al., 2021). Further

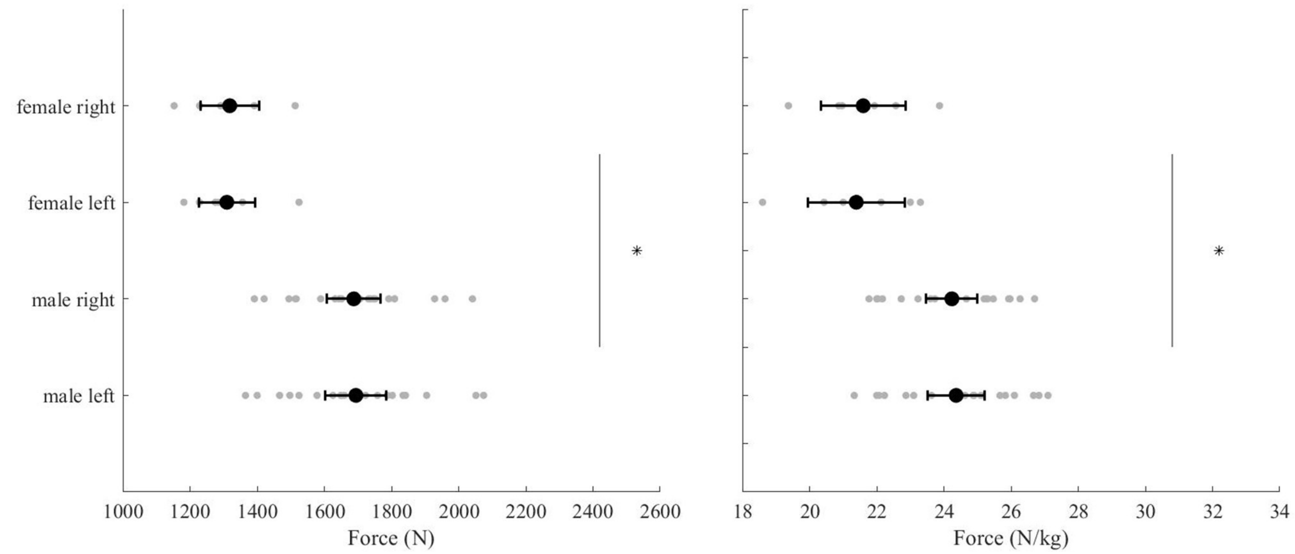

FIGURE 2 | PF_{abs} values (left) for left and right limbs in male and female freeski athletes. Relative PF values (right) for left and right limbs for male and female freeski athletes. Data are shown as mean (+95% confidence intervals) and individual data (gray dots). *Represents a significant difference between male and females.

research is required to determine a detailed physiological and neuromuscular profile of academy to elite freeski athletes. Such data combined with kinetic, kinematic, and qualitative analysis of tricks during the sport could provide meaningful information to practitioners aiming to enhance the physical preparation of these athletes and help support talent identification and long-term athlete development models. Without further investigation, it is uncertain whether the UIHBP test could be used to track both acute and chronic changes in neuromuscular performance. Nevertheless, isometric contractions have been shown to be a highly reliable means of assessing and tracking force production (Wilson and Murphy, 1996; Bazyler et al., 2015; Drake et al., 2017). However, the ability of isometric assessments to predict dynamic performance compared to alternative modalities of assessment such as isokinetic and isoinertial testing is not as well supported (Wilson and Murphy, 1996).

Recent data from the 2016 and 2020 Youth Winter Olympic Games show that the highest percentage of injuries occurred in freeski and snowboard slopestyle disciplines (Steffen et al., 2017; Palmer et al., 2021). A similar trend was also apparent at the PyeongChang 2018 Winter Olympic Games (Soligard et al., 2019). This highlights that further longitudinal/multi-season injury surveillance data of academy to elite level freeskiers and in-depth investigation of injury mechanisms are required. Previous data from Flørenes et al. (2010) showed that one quarter of all injuries encountered by freestyle ski athletes involved the knee (with 38% of these relating to the ACL). However, this study included the freestyle ski disciplines of moguls, dual-moguls, aerials, ski cross and halfpipe skiers, with no data from the freeski disciplines of slopestyle and big air. Given that it is common for freeski athletes to suffer an injury, it would appear worthwhile to monitor training load and neuromuscular status throughout the season and during appropriate stages of the return to sport process. The UIHBP MVC test outlined in the present study could potentially be incorporated into lower body rehabilitation programs. Physical properties such as maximum strength, explosive strength, and reactive strength have also been shown to influence reinjury outcomes (Kyritsis et al., 2016; King et al., 2018) and it is recommended that objective physical testing be carried out before athletes return to sport (Carolan et al., 2020). Regarding maximal strength, there is evidence highlighting that return to sport frameworks should include the assessment of unilateral quadriceps MVC during suitable phases of the rehabilitation (Buckthorpe et al., 2020; Jordan et al., 2020). Current ACL return to sport protocols recommend using a quadriceps limb symmetry index of 90% before determining readiness to return to sport (Gokeler et al., 2016; Brown et al., 2021), though it is not clear whether the same thresholds are appropriate for freeski athletes. It is important to note that although objective insight into elements of the recovery process can be useful, isometric quadriceps limb symmetries can overestimate the recovery of the injured limb (Wellsandt et al., 2017) and therefore, consideration must be given when examining the data of such measures in isolation. Practitioners are recommended to consider a holistic, multifactorial and individual approach to injury rehabilitation (Lahti et al., 2020), emphasizing movement quality (Buckthorpe, 2021), tasks that incorporate decision making and divided attention (Hughes and Dai, 2021), as well as nutritional (Shaw et al., 2019) and psychological readiness elements (Papadopoulos et al., 2018; D'Astous et al., 2020).

It must be highlighted that this data is only representative of the current cohort of freeski athletes recruited in this study and further data is required, especially from female athletes as the sample size in this study was small. Additionally, these measurements and comparisons were only taken from one specific joint angle (115° knee flexion). Further analysis from several knee and hip angles could yield different results as alternative limb arrangements can affect force production and muscle recruitment patterns (Dos' Santos et al., 2017; Goodwin and Bull, 2021). However, isometric contractions performed at longer muscle lengths can result in increased muscle damage (Allen et al., 2018). Moreover, correlation with an isokinetic dynamometer and further exploration of the current test via kinematic and EMG analysis to accurately quantify and compare muscle activation patterns could help practitioners make a more informed decision regarding test selection for their environment and specific needs. A limitation to the study is that between-session reliability was not assessed. However, it was not possible to do so without interfering with the day-to-day training of the athletes. Furthermore, rate of force development (RFD) was not analyzed in the current study. Attempting to achieve maximal force and RFD within the same contraction may result in suboptimal measures of both parameters (Maffiuletti et al., 2016). It is recommended that when assessing RFD that contractions be "fast and hard" with short durations (0.5–1.5 s) (Maffiuletti et al., 2016) and this would have interfered with the specific aim of the current study to establish the reliability of measuring peak force.

CONCLUSION

The UIHBP MVC test can be considered as a simple and quick testing approach that provides reliable measures of lower body PF_{abs} and relative PF. The study also provided unilateral lower body force production reference values for male and female freeski athletes.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Jyväskylä ethics committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

JM designed the study and concept under the supervision of VL and JS. JM conducted all data collection and measurements. BG performed the statistical analysis. JM wrote the first draft. VL, JS, and BG significantly contributed to the final version. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We are thankful to Ruka-Vuokatti Sports Academy and Freeski Team Finland Freeski athletes and coaches. With special thanks to coaches Juha Mourujärvi and Jussi Mononen.

REFERENCES

Abbott, B. C., and Wilkie, D. (1953). The relation between velocity of shortening and the tension-length curve of skeletal muscle. *J. Physiol.* 120, 214–223. doi: 10.1113/jphysiol.1953.sp004886

Allen, T. J., Jones, T., Tsay, A., Morgan, D. L., and Proske, U. (2018). Muscle damage produced by isometric contractions in human elbow flexors. *J. Appl. Physiol.* 124, 388–399. doi: 10.1152/japplphysiol.00535.2017

Andersen, V., Fimland, M. S., Mo, D. A., Iversen, V. M., Vederhus, T., Hellebø, L. R. R., et al. (2018). Electromyographic comparison of barbell deadlift, hex bar deadlift, and hip thrust exercises: a cross-over study. *J. Strength Cond. Res.* 32, 587–593. doi: 10.1519/JSC.0000000000001826

Atkinson, G., and Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 26, 217–238. doi: 10.2165/00007256-199826040-00002

Bartolomei, S., Rovai, C., Lanzoni, I. M., and di Michele, R. (2019). Relationships between muscle architecture, deadlift performance, and maximal isometric force produced at the midthigh and midshin pull in resistance-trained individuals. *J. Strength Cond. Res.* doi: 10.1519/JSC.0000000000003455

Bazyler, C. D., Beckham, G. K., and Sato, K. (2015). The use of the isometric squat as a measure of strength and explosiveness. *J. Strength Cond. Res.* 29, 1386–1392. doi: 10.1519/JSC.0000000000000751

Blazevich, A. J., and Sharp, N. C. (2005). Understanding muscle architectural adaptation: macro-and micro-level research. *Cells Tissues Organs* 181, 1–10. doi: 10.1159/000089964

Bogdanis, G. C., Tsoukos, A., Methenitis, S. K., Selima, E., Veligekas, P., and Terzis, G. (2019). Effects of low volume isometric leg press complex training at two knee angles on force-angle relationship and rate of force development. *Eur. J. Sport Sci.* 19, 345–353. doi: 10.1080/17461391.2018.1510989

Bosquet, L., Gouadec, K., Berryman, N., Duclos, C., Gremeaux, V., and Croisier, J. L. (2016). The total work measured during a high intensity isokinetic fatigue test is associated with anaerobic work capacity. *J. Sport Sci. Med.* 15:126.

Brown, C., Marinko, L., LaValley, M. P., and Kumar, D. (2021). Quadriceps strength after anterior cruciate ligament reconstruction compared with uninjured matched controls: a systematic review and meta-analysis. *Orthop. J. Sports Med.* 9:2325967121991534. doi: 10.1177/2325967121991534

Brughelli, M., and Van Leemputte, M. (2013). Reliability of power output during eccentric sprint cycling. *J. Strength Cond. Res.* 27, 76–82. doi: 10.1519/JSC.0b013e31824f2055

Buckthorpe, M. (2021). Recommendations for movement re-training after ACL reconstruction. *Sports Med.* 1–18. doi: 10.1007/s40279-021-01454-5

Buckthorpe, M., Tamisari, A., and Della Villa, F. (2020). A ten task-based progression in rehabilitation after acl reconstruction: from post-surgery to return to play–a clinical commentary. *Int. J. Sports Phys. Ther.* 15:611. doi: 10.26603/ijspt20200611

Camara, K. D., Coburn, J. W., Dunnick, D. D., Brown, L. E., Galpin, A. J., and Costa, P. B. (2016). An examination of muscle activation and power characteristics while performing the deadlift exercise with straight and hexagonal barbells. *J. Strength Cond. Res.* 30, 1183–1188. doi: 10.1519/JSC.0000000000001352

Carolan, D., King, E., Richter, C., Franklyn-Miller, A., Moran, R., and Jackson, M. (2020). Differences in strength, patient-reported outcomes, and return-to-play rates between athletes with primary vs. revision ACL reconstruction at 9 months after surgery. *Orthop. J. Sports Med.* 8:2325967120950037. doi: 10.1177/2325967120950037

Caruso, J. F., Brown, L. E., and Tufano, J. J. (2012). The reproducibility of isokinetic dynamometry data. *Isokinet. Exerc. Sci.* 20, 239–253. doi: 10.3233/IES-2012-0477

Collins, D., Willmott, T., and Collins, L. (2018). Periodization and self-regulation in action sports: coping with the emotional load. *Front. Psychol.* 9:1652. doi: 10.3389/fpsyg.2018.01652

Comfort, P., Stewart, A., Bloom, L., and Clarkson, B. (2014). Relationships between strength, sprint, and jump performance in well-trained youth soccer players. *J. Strength Cond. Res.* 28, 173–177. doi: 10.1519/JSC.0b013e318291b8c7

Cross, M. R., Rivière, J. R., Coulmy, N., Morin, J. B., and Samozino, P. (2021). Lower limb force-production capacities in alpine skiing disciplines. *Scand. J. Med. Sci. Sports* 31, 848–860. doi: 10.1111/sms.13897

D'Astous, E., Podlog, L., Burns, R., Newton, M., and Fawver, B. (2020). Perceived competence, achievement goals, and return-to-sport outcomes: a mediation analysis. *Int. J. Environ. Res. Public Health* 17:2980. doi: 10.3390/ijerph17092980

Dos' Santos, T., Thomas, C., Jones, P. A., McMahon, J. J., and Comfort, P. (2017). The effect of hip joint angle on isometric midthigh pull kinetics. *J. Strength Cond. Res.* 31, 2748–2757. doi: 10.1519/JSC.0000000000002098

Drake, D., Kennedy, R., and Wallace, E. (2017). The validity and responsiveness of isometric lower body multi-joint tests of muscular strength: a systematic review. *Sports Med. Open* 3, 1–11. doi: 10.1186/s40798-017-0091-2

Edwards, T., Spiteri, T., Piggott, B., Bonhotal, J., Haff, G. G., and Joyce, C. (2018). Monitoring and managing fatigue in basketball. *Sports* 6:19. doi: 10.3390/sports6010019

Eliassen, W., Saeterbakken, A. H., and van den Tillaar, R. (2018). Comparison of bilateral and unilateral squat exercises on barbell kinematics and muscle activation. *Int. J. Sports Phys. Ther.* 13, 871–881. doi: 10.26603/ijspt20180871

Flørenes, T. W., Heir, S., Nordsletten, L., and Bahr, R. (2010). Injuries among world cup freestyle skiers. *Br. J. Sports Med.* 44, 803–808. doi: 10.1136/bjsm.2009.071159

Ghasemi, A., and Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *Int. J. Endocrinol. Metabol.* 10:486. doi: 10.5812/ijem.3505

Gleeson, N. P., and Mercer, T. H. (1996). The utility of isokinetic dynamometry in the assessment of human muscle function. *Sports Med.* 21, 18–34. doi: 10.2165/00007256-199621010-00003

Gokeler, A., Welling, W., Zaffagnini, S., Seil, R., and Padua, D. (2016). Development of a test battery to enhance safe return to sports after anterior cruciate ligament reconstruction. *Knee Surg. Sports Traumatol. Arthrosc.* 25, 192–199. doi: 10.1007/s00167-016-4246-3

Goodwin, J. E., and Bull, A. M. (2021). Novel assessment of isometric hip extensor function: reliability, joint angle sensitivity, and concurrent validity. *J. Strength Cond. Res.* doi: 10.1519/JSC.0000000000004012

Granacher, U., Muehlbauer, T., Doerflinger, B., Strohmeier, R., and Gollhofer, A. (2011). Promoting strength and balance in adolescents during physical education: effects of a short-term resistance training. *J. Strength Cond. Res.* 25, 940–949. doi: 10.1519/JSC.0b013e3181c7bb1e

Grgic, J., Lazinica, B., Schoenfeld, B. J., and Pedisic, Z. (2020). Test–retest reliability of the one-repetition maximum (1RM) strength assessment: a systematic review. *Sports Med. Open* 6, 1–16. doi: 10.1186/s40798-020-00260-z

Hughes, G., and Dai, B. (2021). The influence of decision making and divided attention on lower limb biomechanics associated with anterior cruciate ligament injury: a narrative review. *Sports Biomech.* 1–18. doi: 10.1080/14763141.2021.1898671

Izquierdo, M., Ibanez, J., Calbet, J. A. L., González-Izal, M., Navarro-Amézqueta, I., Granados, C., et al. (2009). Neuromuscular fatigue after resistance training. *Int. J. Sports Med.* 30, 614–623. doi: 10.1055/s-0029-1214379

Jordan, M. J., Morris, N., Lane, M., Barnert, J., MacGregor, K., Heard, M., et al. (2020). Monitoring the return to sport transition after ACL injury: an alpine ski racing case study. *Front. Sports Act. living* 2:12. doi: 10.3389/fspor.2020.00012

King, E., Richter, C., Franklyn-Miller, A., Daniels, K., Wadey, R., Moran, R., et al. (2018). Whole-body biomechanical differences between limbs exist 9 months after ACL reconstruction across jump/landing tasks. *Scand. J. Med. Sci. Sports* 28, 2567–2578. doi: 10.1111/sms.13259

Kirkpatrick, J., and Comfort, P. (2013). Strength, power, and speed qualities in English junior elite rugby league players. *J. Strength Cond. Res.* 27, 2414–2419. doi: 10.1519/JSC.0b013e3182804a6d

Knapik, J. J., Wright, J. E., Mawdsley, R. H., and Braun, J. M. (1983). Isokinetic, isometric and isotonic strength relationships. *Arch. Phys. Med. Rehabil.* 64, 77–80.

Knutson, L. M., Soderberg, G. L., Ballantyne, B. T., and Clarke, W. R. (1994). A study of various normalization procedures for within day electromyographic data. *J. Electromyogr. Kinesiol.* 4, 47–59. doi: 10.1016/1050-6411(94)90026-4

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012

Kubo, K., Yata, H., Kanehisa, H., and Fukunaga, T. (2006). Effects of isometric squat training on the tendon stiffness and jump performance. *Eur. J. Appl. Physiol.* 96, 305–314. doi: 10.1007/s00421-005-0087-3

Kyritsis, P., Bahr, R., Landreau, P., Miladi, R., and Witvrouw, E. (2016). Likelihood of ACL graft rupture: not meeting six clinical discharge criteria before return to sport is associated with a four times greater risk of rupture. *Br. J. Sports Med.* 50, 946–951. doi: 10.1136/bjsports-2015-095908

Lahti, J., Mendiguchia, J., Ahtiainen, J., Anula, L., Kononen, T., Kujala, M., et al. (2020). Multifactorial individualised programme for hamstring muscle injury risk reduction in professional football: protocol for a prospective cohort study. *BMJ Open Sport Exerc. Med.* 6:000758. doi: 10.1136/bmjsem-2020-000758

Lake, J., Duncan, F., Jackson, M., and Naworynsky, D. (2017). Effect of a hexagonal barbell on the mechanical demand of deadlift performance. *Sports* 5:82. doi: 10.3390/sports5040082

Lauersen, J. B., Andersen, T. E., and Andersen, L. B. (2018). Strength training as superior, dose-dependent and safe prevention of acute and overuse sports injuries: a systematic review, qualitative analysis and meta-analysis. *Br. J. Sports Med.* 52, 1557–1563. doi: 10.1136/bjsports-2018-099078

Ly, L. P., and Handelsman, D. J. (2002). Muscle strength and ageing: methodological aspects of isokinetic dynamometry and androgen administration. *Clin. Exp. Pharmacol. Physiol.* 29, 39–47. doi: 10.1046/j.1440-1681.2002.03606.x

Maestroni, L., Read, P., Bishop, C., and Turner, A. (2019). Strength and power training in rehabilitation: underpinning principles and practical strategies to return athletes to high performance. *Sports Med.* 50, 1–14. doi: 10.1007/s40279-019-01195-6

Maffiuletti, N. A., Aagaard, P., Blazevich, A. J., Folland, J., Tillin, N., and Duchateau, J. (2016). Rate of force development: physiological and methodological considerations. *Eur. J. Appl. Physiol.* 116, 1091–1116. doi: 10.1007/s00421-016-3346-6

Markovic, G., and Jaric, S. (2004). Movement performance and body size: the relationship for different groups of tests. *Eur. J. Appl. Physiol.* 92, 139–149. doi: 10.1007/s00421-004-1076-7

McCall, A., Nedelec, M., Carling, C., Le Gall, F., Berthoin, S., and Dupont, G. (2015). Reliability and sensitivity of a simple isometric posterior lower limb muscle test in professional football players. *J. Sports Sci.* 33, 1298–1304. doi: 10.1080/02640414.2015.1022579

Miller, A. E. J., MacDougall, J. D., Tarnopolsky, M. A., and Sale, D. G. (1993). Gender differences in strength and muscle fiber characteristics. *Eur. J. Appl. Physiol. Occup. Physiol.* 66, 254–262. doi: 10.1007/BF00235103

Moeskops, S., Oliver, J. L., Read, P. J., Cronin, J. B., Myer, G. D., Haff, G. G., et al. (2018). Within-and between-session reliability of the isometric mid-thigh pull in young female athletes. *J. Strength Cond. Res.* 32:1892. doi: 10.1519/JSC.0000000000002566

Nindl, B. C., Mahar, M. T., Harman, E. A., and Patton, J. F. (1995). Lower and upper body anaerobic performance in male and female adolescent athletes. *Med. Sci. Sports Exerc.* 27, 235–241. doi: 10.1249/00005768-199502000-00013

Noorkõiv, M., Nosaka, K., and Blazevich, A. (2014). Neuromuscular adaptations associated with knee joint angle-specific force change. *Med. Sci. Sports Exerc.* 46, 1525–1537. doi: 10.1249/MSS.0000000000000269

Nosaka, K., Lavender, A., Newton, M., and Sacco, P. (2003). Muscle damage in resistance training—is muscle damage necessary for strength gain and muscle hypertrophy? *Int. J. Sport Health Sci.* 1, 1–8. doi: 10.5432/ijshs.1.1

Ogrin, J., Šarabon, N., Madsen, M. K., Kersting, U., Holmberg, H. C., and Supej, M. (2021). Asymmetries in ground reaction forces during turns by elite slalom alpine skiers are not related to asymmetries in muscular strength. *Front. Physiol.* 12:31. doi: 10.3389/fphys.2021.577698

Ohtonen, O., Ruotsalainen, K., Mikkonen, P., Heikkinen, T., Hakkarainen, A., Leppävuori, A., et al. (2016). "Online feedback system for athletes and coaches," in *Book: Science and Nordic Skiing III*, eds A. Hakkarainen, V. Linnamo, S. Lindinger (Jyväskylä: University of Jyväskylä).

Palmer, D., Engebretsen, L., Carrard, J., Grek, N., Königstein, K., Maurer, D. J., et al. (2021). Sports injuries and illnesses at the Lausanne 2020 youth olympic winter games: a prospective study of 1,783 athletes from 79 countries. *Br. J. Sports Med.* doi: 10.1136/bjsports-2020-103514

Palmieri-Smith, R. M., and Lepley, L. K. (2015). Quadriceps strength asymmetry following ACL reconstruction alters knee joint biomechanics and functional performance at time of return to activity. *Am. J. Sports Med.* 43:1662. doi: 10.1177/0363546515578252

Papadopoulos, S., Tishukov, M., Stamou, K., Totlis, T., and Natsis, K. (2018). Fear of re-injury following ACL reconstruction: an overview. *J. Res. Pract. Musculoskelet. Syst*. 2, 124–130. doi: 10.22540/JRPMS-02-124

Portney, L. G., and Watkins, M. P. (2009). *Foundations of Clinical Research: Applications to Practice (Vol. 892)*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Shaw, G., Serpell, B., and Baar, K. (2019). Rehabilitation and nutrition protocols for optimising return to play from traditional ACL reconstruction in elite rugby union players: a case study. *J. Sports Sci.* 37, 1794–1803. doi: 10.1080/02640414.2019.1594571

Soligard, T., Palmer, D., Steffen, K., Lopes, A. D., Grant, M. E., Kim, D., et al. (2019). Sports injury and illness incidence in the PyeongChang 2018 Olympic Winter Games: a prospective study of 2,914 athletes from 92 countries. *Br. J. Sports Med*. 53, 1085–1092. doi: 10.1136/bjsports-2018-100236

Steffen, K., Moseid, C. H., Engebretsen, L., Søberg, P. K., Amundsen, O., Holm, K., et al. (2017). Sports injuries and illnesses in the Lillehammer 2016 youth olympic winter games. *Br. J. Sports Med*. 51, 29–35. doi: 10.1136/bjsports-2016-096977

Taberner, M., van Dyk, N., Allen, T., Jain, N., Richter, C., Drust, B., et al. (2020). Physical preparation and return to performance of an elite female football player following ACL reconstruction: a journey to the FIFA Women's World Cup. *BMJ Open Sport Exerc. Med*. 6:e000843. doi: 10.1136/bmjsem-2020-000843

Taipale, R. S., Schumann, M., Mikkola, J., Nyman, K., Kyröläinen, H., Nummela, A., et al. (2014). Acute neuromuscular and metabolic responses to combined strength and endurance loadings: the "order effect" in recreationally endurance trained runners. *J. Sports Sci*. 32, 1155–1164. doi: 10.1080/02640414.2014.889842

VanderWeele, T. J., and Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: is the Bonferroni correction really so bad? *Am. J. Epidemiol.* 188, 617–618. doi: 10.1093/aje/kwy250

Verdera, F., Champavier, L., Schmidt, C., Bermon, S., and Marconnet, P. (1999). Reliablilty and validity of a new device to measure isometric strength in polyarticular exercises. *J. Sports Med. Phys. Fitness* 39:113.

Vernillo, G., Pisoni, C., and Thiebat, G. (2016). Physiological characteristics of elite snowboarders. *J. Sports Med. Phys. Fitness* 56, 527–533. doi: 10.3389/fphys.2018.00770

Wellsandt, E., Failla, M. J., and Snyder-Mackler, L. (2017). Limb symmetry indexes can overestimate knee function after anterior cruciate ligament injury. *J. Orthop. Sports Phys. Ther.* 47, 334–338. doi: 10.2519/jospt.2017.7285

West, D. J., Owen, N. J., Jones, M. R., Bracken, R. M., Cook, C. J., Cunningham, D. J., et al. (2011). Relationships between force–time characteristics of the isometric midthigh pull and dynamic performance in professional rugby league players. *J. Strength Cond. Res*. 25, 3070–3075. doi: 10.1519/JSC.0b013e318212dcd5

Willmott, T., and Collins, D. (2015). Challenges in the transition to mainstream: promoting progress and minimizing injury in freeskiing and snowboarding. *Sport Soc*. 18, 1245–1259. doi: 10.1080/17430437.2015.1031530

Wilson, G. J., and Murphy, A. J. (1996). The use of isometric tests of muscular function in athletic assessment. *Sports Med*. 22, 19–37. doi: 10.2165/00007256-199622010-00003

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Sports and Active Living

ORIGINAL RESEARCH
published: 10 March 2022
doi: 10.3389/fspor.2022.818556

Body Pose Estimation Integrated With Notational Analysis: A New Approach to Analyze Penalty Kicks Strategy in Elite Football

Guilherme de Sousa Pinheiro[1]*, Xing Jin[2], Varley Teoldo Da Costa[3] and Martin Lames[1]

[1] Performance Analysis and Sports Informatics, Technical University, Munich, Germany, [2] Department of Informatics, Technical University of Munich, Munich, Germany, [3] UFMG Soccer Science Center, Department of Sports Science, Federal University of Minas Gerais, Belo Horizonte, Brazil

Body orientation of football players has proven to be an informative resource related to successful penalty kicks. OpenPose is one of the most popular open-source pose estimation technologies. This study aims: (i) to verify whether OpenPose can detect relevant body orientation angles from video data of penalty kicks in elite football and (ii) to investigate the relationship between these body angles and observable behaviors analyzed *via* an observational system for penalty kick analysis in football (OSPAF) with the penalty taker and goalkeeper strategy. A total of 34 penalty videos, with standardized viewing angle, from the main European leagues (2017–2020) were analyzed. Relevant body orientation variables were selected for penalty kicks analysis and were extracted from video data through OpenPose technique. The OSPAF, previously validated by experts, was used. The mean confidence score of OpenPose measures was 0.80 ± 0.14. OpenPose Retest reliability values was 0.976 ± 0.03. Logistic regressions were performed to investigate the relationship between OpenPose investigated variables (penalty taker: shoulder, hips, and nonkicking foot orientation; goalkeeper: right and left foot, anticipation), observable behaviors (OSPAF variables), and the strategy (penalty taker: goalkeeper dependent or independent; goalkeeper: shooter dependent or independent) in penalty kicks. The selected body orientation angle (goalkeeper anticipation) measured through OpenPose correlated significantly with the goalkeeper strategy. The prediction model of the goalkeeper's strategy had its accuracy increased to 97% when the variable goalkeeper anticipation was included [$\chi^2_{(35)} = 49.648$, $p < 0.001$]. Lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and run up speed (slow) were associated with a kicker-dependent strategy. Regarding the penalty taker, the selected body angles measured through OpenPose did not associate significantly with the shooter strategy. Body orientation analysis by using OpenPose has shown sufficient reliability and provides practical applications for analyzing the strategies adopted by goalkeepers in penalty kicks in elite football.

Keywords: body orientation, performance analysis, OSPAF, OpenPose, human movement, motion capture, soccer analytics

OPEN ACCESS

Edited by:
Zimi Sawacha, University of Padua, Italy

Reviewed by:
Matteo Zago, University of Milan, Italy
Fabiola Spolaor, University of Padua, Italy

****Correspondence:***
Guilherme de Sousa Pinheiro
guilherme.pinheiro@tum.de

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: *19 November 2021*
Accepted: *26 January 2022*
Published: *10 March 2022*

Citation:
Pinheiro GS, Jin X, Costa VTD and Lames M (2022) Body Pose Estimation Integrated With Notational Analysis: A New Approach to Analyze Penalty Kicks Strategy in Elite Football. Front. Sports Act. Living 4:818556. doi: 10.3389/fspor.2022.818556

INTRODUCTION

The analysis of penalty kick performance in football has played an important role in sports analytics (Paterson et al., 2020; Noël et al., 2021; Pinheiro et al., 2021a,b). Over the past 30 years, there have been several scientific studies that identify the motivational-, strategic-, anticipatory-, attention-, and perception-based factors that can mean a successful or failed penalty kick (Memmert and Nöel, 2020). Recent research focusing on the technical dynamics of penalty kicks has also identified multiple key variables that can differentiate the players strategy (Pinheiro et al., 2021b) and enhance the overall chances of scoring a penalty kick (Jamil et al., 2020). The importance of the optimal performance of both the rival players during the penalty kick is paramount, especially since the introduction of the penalty shoot-out in major competitions to determine which team progresses after a drawn match (Fariña et al., 2013).

One prerequisite to increase the probability of successful performance is the implementation of the suitable penalty kick strategy (van der Kamp, 2006). Previous research has identified two main strategies for taking a penalty (Kuhn, 1988; van der Kamp, 2006). First, the keeper-independent strategy, where the kicker selects the target location to shoot toward before the run-up and does not attend to the actions made by the goalkeeper during the run-up. The decision of where to aim depends on the penalty taker's kicking preference (Noël et al., 2015). On the contrary, in the keeper-dependent strategy, the kicker tries to obtain information from the goalkeeper's reactions during the run-up. Nevertheless, the outcome of a penalty is determined by an interaction between the shooter's strategy (e.g., technique, speed) and the goalkeeper's strategy (Hunter et al., 2018; Pinheiro et al., 2021b). The optimal strategy depends on the keeper's behavior and the relative benefits of speed, accuracy, and unpredictability within each situation. Regarding the goalkeeper strategy, there are two approaches: the dependent and independent penalty takers. The goalkeeper who behaves according to the first group defines his movement based on the actions of the penalty taker. The second type of goalkeeper is the one who risks jumping to a corner independently of the kicker's movement (Kuhn, 1988).

The analysis of the penalty kick strategies has been investigated about numerous factors (Noël et al., 2015; Pinheiro et al., 2021b). Noël et al. (2015) developed a method for investigating penalty taker strategies, based on a controlled simulated situation. In a noncompetitive setting, youth players were instructed to take penalty kicks adopting either a keeper-independent or keeper-dependent strategy. Based on this setting, an observational system was developed to evaluate penalty kick performances by using video footage from competitive matches. Those authors identified that attention to the goalkeeper, run-up fluency, and kicking technique in combination could predict kick strategy in 92% of the penalties. However, one possible limitation is that the penalty takers followed a script denoting whether they use a keeper-independent or keeper-dependent strategy and, therefore, the design created differed very importantly from the match situation (Pinheiro et al., 2021a). Besides that, it remains unclear whether the young players disposed of a sufficient skill level to execute both the strategies with the same quality. To address the interaction process in professional football and provide a valid instrument, (Pinheiro et al., 2021b) developed an observational system for penalty kick analysis in football (OSPAF). The OSPAF met all the requirements of instrument validation.

Body orientation has been indicated as a key factor under covering the success in penalty kicks (Li et al., 2015). However, it is a yet little explored area in penalty kick analytics. There is a need within human movement sciences for a markerless motion capture system, which is easy to use and sufficiently accurate to evaluate motor performance (Nakano et al., 2020). OpenPose method adopts unique top-down position recognition by using deep learning and also the unique algorithm as affiliation recognition of body parts by Part Affinity Fields (PAFs) to detect the two-dimensional (2D) pose of multiple people in images (Nakai et al., 2019). OpenPose can recognize skeletons of multiple players in real-time, by using a simple web camera. Given a video or image, OpenPose estimates a total of 25 biometric human body parts (e.g., right knee, left knee, and right foot). The output of the algorithm is in the form of 25×3 vector for each individual, where the first two columns of the vector stand for the x-y coordinate of key points in the field domain, while the third column represents the confidence score. This method has shown high-level accuracy on multiple public benchmarks, being efficient for multiperson pose estimation (Cao et al., 2017). Zago et al. (2020) confirmed the feasibility of tracking kinematics by using OpenPose. OpenPose-based markerless motion capture can be used for human movement science with an accuracy of 30 mm or less (Nakano et al., 2020). Despite several studies in this area, key gaps remain, including a lack of research by using OpenPose to detect relevant body orientation angles in field settings and based on sports broadcasts such as penalty kicks from TV videos.

Sangüesa et al. (2019) had previously applied OpenPose to estimate the body orientation of football players from video data during match play. Those authors indicated that a time-based set of player orientations might detect specific situations where orientation is crucial in the match. Recently, Sangüesa et al. (2020a) used a player's body orientation to model pass feasibility in football. The inclusion of the orientation data estimated directly from video frames by using pose models, into a passing model, has proved to be a key feature in the decision-making process of players and is strictly correlated to the play outcome. In another study, Sangüesa et al. (2020b) mapped body pose parts (e.g., shoulders and hips) in a 2D field by combining OpenPose with a super-resolution network and merging the obtained estimation with contextual information (ball position). Results have been validated with players held electronic performance and tracking systems devices, obtaining a median error of 27° per player.

Notation analysis has been widely used to examine the technical properties of football performance through recording behavior incidence (Lames and Hansen, 2001; Hughes and Bartlett, 2004; Sarmento et al., 2014; Casal et al., 2017; Pinheiro et al., 2021b). In the recent years, there has been a vertiginous evolution in the match analysis methods, mainly motivated by the emergence of automatic registration procedures, which allows the immediate acquisition of a large amount of data related to the positioning of the players with the game (Castellano et al., 2014). The rise of sports analytics has provided a new set of metrics and statistics that can serve coaches to evaluate the player (Sangüesa et al., 2019). Nevertheless, one limitation is that one method does not entirely supply all the necessary information. There is, therefore, a need to use multimethod approach to solve sports analytics problems, analyzing variables by using different methods (Aranda et al., 2019). Methodology designs that combine different study approaches (e.g., observational and biomechanical/method that produce body angles), also known as mixed methods (Preciado et al., 2019), tend to provide a deeper understanding and reliability of the studied phenomenon (i.e., penalty kicks).

The influence variables on penalty kicks success are extensively studied (Jamil et al., 2020; Memmert and Nöel, 2020; Paterson et al., 2020; Noël et al., 2021; Pinheiro et al., 2021b). (Pinheiro et al., 2021b) recommended that future studies could use the OSPAF, applying technological methods to analyze its variables, such as computer techniques for body pose estimation and machine learning-based video analysis. To the best of our knowledge, no study has used OpenPose to detect relevant body orientation angles in penalty kicks in elite football from TV broadcast. Therefore, the aims of this study are: (i) to verify whether OpenPose can detect relevant body orientation angles from video data of penalty kicks in elite football and (ii) to investigate the relationship between these body angles and observable behaviors analyzed via OSPAF (Pinheiro et al., 2021b) with the penalty taker and goalkeeper strategy.

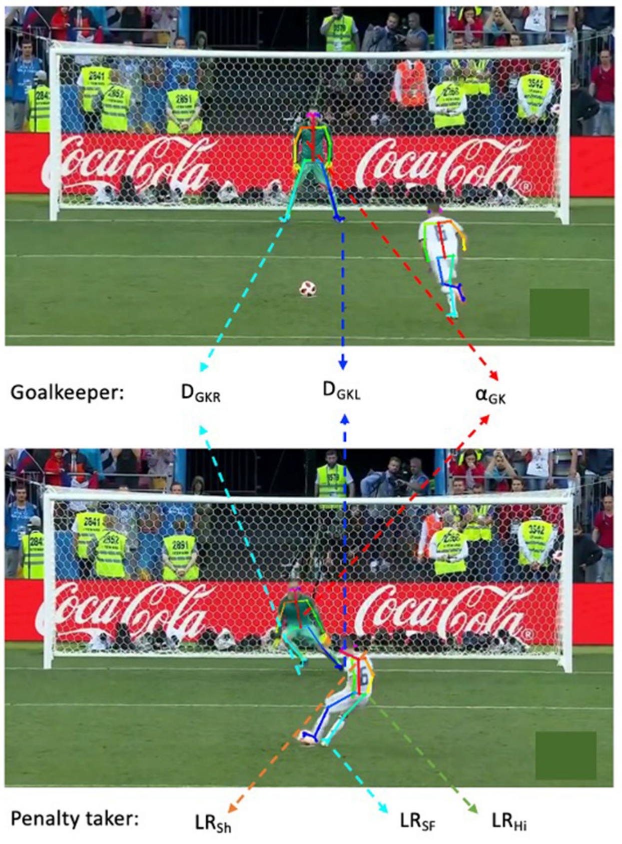

FIGURE 1 | Penalty kick viewing angle, frames analyzed, and process of pose estimation. The upper image corresponds to the moment when the penalty taker starts the run-up approaching the ball, and the down one corresponds to the moment when he touches the ball. The coordinates belonging to shoulders, hips, and non-kicking foot of penalty taker are mapped in a 2D field. LR-side Booleans (LR_{Sh}, LR_{Hi}, LR_{SF}: penalty kicker's shoulder, hips, and non-kicking foot orientation, respectively), angles (α_{Sh} α_{Hi}, α_{SF}: penalty kicker's shoulder, hip, and non-kicking foot, respectively) and confidences (C_{Sh}, C_{Hi}, C_{SF}: confidence score of the orientation of penalty kicker's shoulder, hips and non-kicking foot, respectively). Coordinates belonging to the neck and hip of the goalkeeper at these two moments are mapped to a vector and the angle is calculated (α_{GK}). D_{GKR} and D_{GKL}, the movement distance of right and left foot; C_{GK}, represent confidence score.

MATERIALS AND METHODS

Sample

The dataset consists of 34 penalty kicks from the main European football leagues (Premier League, Ligue 1, Bundesliga, LaLiga, Serie A, and Champions League; seasons 2017–2020). The videos were recorded from TV broadcasters and were registered and analyzed postevent. As the video recordings were public, confidentiality was not an issue and authorization was not required from the players observed or their representatives. The procedures performed in this study were in strict accordance with the Declaration of Helsinki as well as with the ethical standards of the Technical University of Munich.

Methodological Design

All the penalty kick data were annotated by the researchers with the OSPAF (Pinheiro et al., 2021b). Body orientation was analyzed by using OpenPose (CMU-Perceptual-Computing-Lab, 2017). The choice and analysis of the penalty kick video viewing angle was standardized (Pinheiro et al., 2021b), with a pixel resolution of 1,280 × 720. The viewing angle used in this study was the view behind the penalty taker (**Figure 1**). The confidence score, calculated by OpenPose, was used to evaluate reliability (Sangüesa et al., 2019). In order to check the stability within the observation, every penalty kick was analyzed with OpenPose twice. Retest reliability was utilized to check these repeated measurements.

Body Pose Detection and Orientation

OpenPose (version 1.4.0) was installed from GitHub (CMU-Perceptual-Computing-Lab, 2017) and run with a notebook (Apple's M1 Chip) under default settings. Orientation from pose used pretrained models and three-dimensional (3D) vision techniques to obtain a first orientation estimation of each player. Once the pose is extracted for each player, the coordinates and confidence level associated with the body parts are stored to estimate the pose orientation. As a result, in the moving skeletal pictures generated by OpenPose, the skeleton marks are shown and overlapped well with the figure of players (Nakai et al., 2019). For technical details of pose models, see Ramakrishna et al. (2014), Wei et al. (2016), and Cao et al. (2017).

In this study, the orientation of a player's body was defined as the 2D rotation of the player's upper torso around the vertical axis, which is assumed to coincide with the field projection of a normal vector placed in the center of their upper torso, involving both the shoulders and hip parts (Sangüesa et al., 2019). Especially in the case of the non-kicking foot, the hallux and the fifth toe of the support foot were used as the left-right (LR) pair to find the normal vector. Orientation was measured in degrees. For technical details of this methodological approach, see Sangüesa et al. (2019, 2020a,b).

In this study, two frames were analyzed. First, when the penalty taker starts the run-up into the ball and, second, when he touches the ball (**Figure 1**). Then, the target variables for the penalty taker (nonkick foot orientation, hips, and shoulders) and the goalkeeper [anticipation movement (explained in detail below) and right and left foot orientation] were extracted. There might be blurry frames and overlap of players. OpenPose could then fail to detect the main biometric body parts of the two players involved in this analysis; therefore, in this case, the neighboring frames, in which biometric body parts can be detected, were used.

Once the pose was extracted for the goalkeeper and penalty taker, the direct linear transformation (DLT) algorithm (Hartley and Zisserman, 2004) was used to map the coordinate information of players into a 2D field with a homography, given the 4 field corners' coordinates in the image (or its projection out of the image in the nonvisible cases). The homography was first calculated based on four 2D-to-2D point correspondences between the frames (Equation 1). From the output of OpenPose, the coordinates of the main upper-torso parts are found in the image domain; by mapping the LR pair (either shoulders or hips) in the 2D field, a first insight of the player orientation is obtained. The player can be inclined toward the right (0–90 and 270–360°) or the left (90–270°) side of the field.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \alpha H \begin{bmatrix} x \\ y \\ w \end{bmatrix} \text{, } where\ homography\ H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (1)$$

After that, the 2D field projections of the LR pair of penalty taker's shoulders, hips, and nonkicking foot (big toe and small toe) were calculated. All the body parts' orientations could point to the left or right half, based on the angle system presented by Sangüesa et al. (2019, 2020a,b). Based on the 2D projection, LR-side Booleans ($LR_{Sh,}$ $LR_{Hi,}$ and LR_{SF}: penalty kicker's shoulder, hips, and non-kicking foot orientation, respectively), angles (α_{Sh} α_{Hi}, and α_{SF}: penalty kicker's shoulder, hip, and non-kicking foot, respectively), and confidences (C_{Sh}, C_{Hi}, and C_{SF}: confidence score of the orientation of penalty kicker's shoulder, hip, and non-kicking foot, respectively) were obtained. The corresponding confidences are the average of OpenPose's player toes, shoulders, and hips confidences, respectively. **Figure 1** shows the output of OpenPose on which the key biometric body parts of an individual are detected, illustrating the estimation process of orientation.

Anticipation Movement of the Goalkeeper

The anticipation movement of the goalkeeper in the penalty kick was defined here as to how far the goalkeeper moves between: (1) the moment when the penalty taker starts the run-up approaching the ball and the (2) moment when the penalty taker first touches the ball. In detail, the line formed by the connection between the goalkeeper's neck and the middle of the hip was used to depict the position status of the goalkeeper in these two moments. Furthermore, the angle (αGK) between the two lines drawn from the two moments measures the anticipation movement of the goalkeeper. The confidence level for this measure (C_{GK}) was calculated by the average confidence scores of the neck and middle of the hip. This process is given in **Figure 1**.

The movement distance of the goalkeeper's left and right foot was also used to measure the anticipation movement. Left and right ankles were used to represent the left and right feet, respectively; moreover, coordinate information together with metric Euclidean distance was used to depict the movement distance of the goalkeeper's feet, as shown in **Figure 1**.

Ball Speed

Ball speed was determined with the open-source software program Kinovea motion analysis (version 0.8.15, Kinovea, France). This software has already been used in various studies analyzing penalty kicks (Hunter et al., 2018; Makaruk et al., 2019).

Notational Analysis

A previously developed and validated observational system (OSPAF) for penalty analysis in elite football was also used in this study (Pinheiro et al., 2021b). The protocols for the use of observational systems were adopted (Lames and Hansen, 2001; Aranda et al., 2019; Fernandes et al., 2019). All the observable behaviors recorded are shown in **Table 1**.

Data Analysis

For descriptive analysis, mean and SD were used. The Shapiro–Wilk test was performed to verify data normality. The association level between the OSPAF variables with the penalty taker and goalkeeper strategy was determined with the use of the chi-squared (χ^2) test. The effect size was determined by using the Cramer's V and classified as weak (ES $\leq$ 0.2), moderate (0.2 $<$ ES $\leq$ 0.6), and strong (ES $>$ 0.6) (Cohen, 1988). The association level between OpenPose variables with the penalty taker and goalkeeper strategy was determined with the use of the point-biserial correlation. Retest reliability was utilized to check the repeated measurements of OpenPose (Vilagut, 2014). Test-retest reliability coefficients (also called coefficients of stability) vary between 0 and 1, where 1: perfect reliability, $\geq$ 0.9: excellent reliability, $\geq$ 0.8 $<$ 0.9: good reliability, $\geq$ 0.7 $<$ 0.8: acceptable reliability, $\geq$ 0.6 $<$ 0.7: questionable reliability, $\geq$ 0.5 $<$ 0.6: poor reliability, $<$ 0.5: unacceptable reliability, and 0: no reliability (Vogt, 2005; Lindstrom, 2010). To identify which variables would be able to predict the penalty takers and goalkeeper strategy, the logistic regression (*enter method*) analyses were performed. Dimensions and categories of OSPAF were coded in Lince software (**Figure 2**; Gabin et al., 2012;

TABLE 1 | OSPAF variables.

Variables	Definition	Attribute levels
Run up speed	Running speed of the penalty kicker toward the ball	Fast or slow
Run up fluency	Characteristic of the penalty kicker's run during the approach of the ball, with or without pauses.	Continuous running or running with pauses
Run up approach angle	Penalty kicker's running angle to the ball.	Frontal or diagonal
Number of steps	Number of steps of the penalty kicker until contact with the ball	1–3; 3–5; or +5
Kicking technique	The technique used by the penalty kicker to kick the ball	Side foot kick or instep kick
Foot used to kick	Foot used by the penalty kicker to kick the ball	Right or left
Penalty taker gaze behavior	Gaze behavior of the kicker during the approach run.	Gaze at the ball or not at the ball
Goalkeeper (GK) initial posture	Position of the body segments.	Arms raised; arms down or arms extended in a position perpendicular to the goalkeeper 's trunk
Deception by the penalty taker	Indication if the kicker has done any action to distract the goalkeeper during his or her run-up	Yes or no
Goalkeeper tactical action	General evaluation of the way the goalkeeper acted during the penalty shoot-out, to the anticipatory aspect	Try to guess the location of the shot; or awaiting the penalty taker action
Goalkeeper performance	Evaluation of the goalkeeper's performance according to his movement and contact with the ball	0: GK made any final movement to the side of the goal opposite to the final ball location; 1: GK did not move from the center of the goal; 2: GK made a movement in the correct direction but did not dive and failed to make contact with the ball; 3: GK dived in the correct direction but failed to make contact with the ball; 4: GK dived in the correct direction and contacted the ball without saving it; or 5: GK successfully saved the kick
Moment of the match	Time of the match when the penalty will be taken	First half; second half or extra time or shoot out
Location of the match (kicker point of view)	Indication if the penalty kicker is from the home team, visitor, or if he plays on a neutral field.	Home, neutral or away
Momentary result (kicker point of view)	Result of the match (for the penalty kicker) at the moment the penalty was marked.	Winning, drawing or losing
Momentary result (GK point of view)	Result of the match (for the Goalkeeper) at the moment the penalty was marked.	Winning, drawing or losing
Match importance	Level of importance of the match for the team	Championship final match; decisive knockout match; group stage match; early season game; match in final stages of the season
Penalty kick direction	The direction of the ball on goal	Left; center or right
Penalty kick height	Height of the ball on goal	Upper; center or down
Penalty kick outcome	Result of the penalty kick	Goal; saved by goalkeeper or Shot misses goal (wide, over or post)
Penalty taker strategy	Overall strategy perceived by the observer (6)	Goalkeeper dependent; unclear or goalkeeper independent
Goalkeeper strategy	Overall strategy perceived by the observer (6)	Kicker independent; unclear or kicker dependent

Soto et al., 2019). Kappa levels of the OSPAF were 0.90 and 0.86—intra- and interreliability (Pinheiro et al., 2021b). The interpretation of this coefficient was adopted as follows: $\kappa > 0.8$: very good; $0.6 < \kappa < 0.8$: good; $0.4 < \kappa < 0.6$: moderate; $0.2 < \kappa < 0.4$: fair; and $\kappa < 0.2$: poor (Altman, 1991; O'Donoghue, 2010). The level of statistical significance adopted was $\alpha = 0.05$, with a 95% CI. All the data were analyzed by using JASP software (JASP Team, 2021; Computer software; JASP Version 0.14).

RESULTS

Descriptive data of all the OpenPose and OSPAF variables analyzed are presented as **Supplementary Material**.

OpenPose Confidence Score and Retest Reliability

The mean confidence score of OpenPose measures was 0.80 ± 0.14. The confidence score per variable is shown in **Table 2**.

Test-retest reliability values are shown in **Table 3**.

Influence Variables on Goalkeeper Strategy

The association between all the OpenPose and OSPAF variables with the goalkeeper's strategy was analyzed. **Table 4** presents only the variables that presented association and the respective values.

A logistic regression (*enter method*) was performed to investigate the relationship between the goalkeeper's tactical action and run-up speed on the likelihood of the goalkeeper strategy. The logistic regression model was statistically

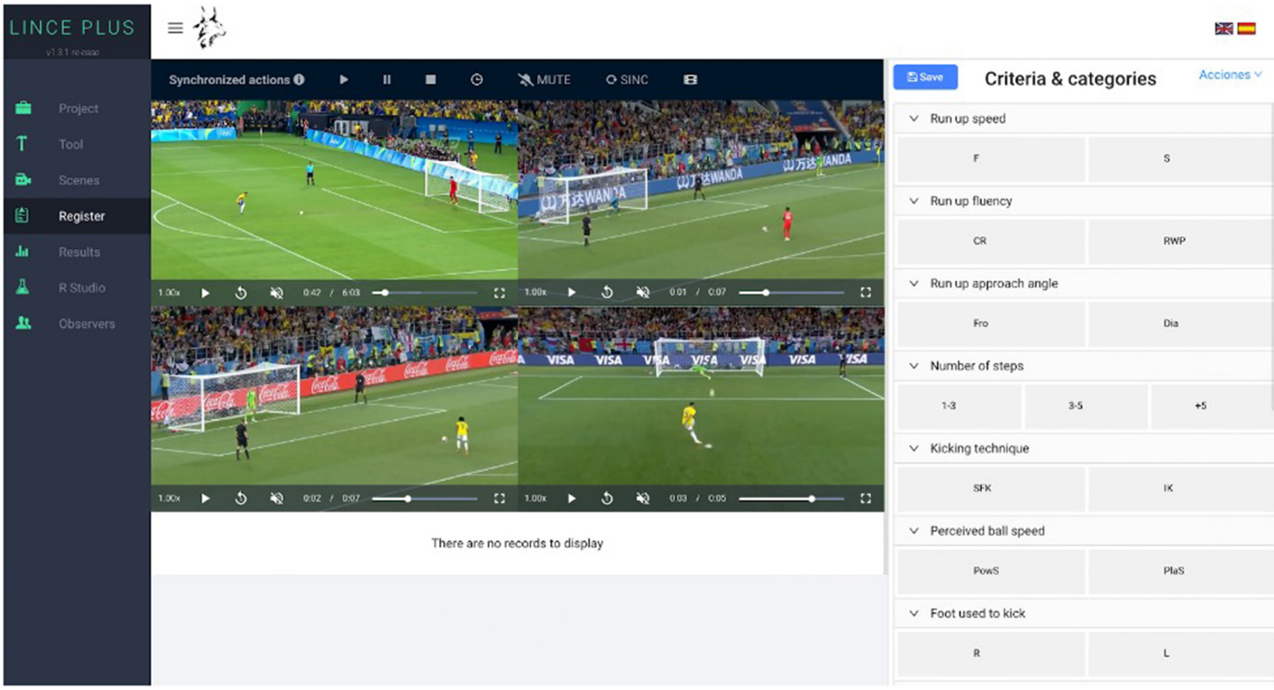

FIGURE 2 | Lince Plus interface. Example of criteria and categories coding.

significant, $\chi^2_{(36)} = 28.592$, $p < 0.001$. The model correctly classified 84.6% of cases. The goalkeeper's tactical action (awaiting) and run speed (slow) were related to a kicker-dependent strategy. While including the correlated OpenPose variable (goalkeeper anticipation) in the model [$\chi^2_{(35)} = 49.648$, $p < 0.001$], the accuracy is increased to 97.0%. Therefore, lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and run-up speed (slow) were associated with a kicker-dependent strategy.

Influence Variables on Penalty Taker Strategy

The association between all the OSPAF and OpenPose variables with the penalty taker's strategy was analyzed. **Table 5** presents only the variables that presented association and the respective values.

A logistic regression (*enter method*) was performed to investigate the relationship between the correlated OSPAF variables (run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed) on the likelihood of the goalkeeper-dependent strategy. The logistic regression model was statistically significant, $\chi^2_{(33)} = 24.819$, $p < 0.001$. The model correctly classified 97.1% of cases. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, the deception performed by the penalty taker, and lower ball speed were related to a goalkeeper-dependent strategy.

DISCUSSION

A unique method to calculate football players' orientation in in-match penalty kicks from a video has been tested. The mean confidence score of OpenPose variables was 0.80 and test-retest reliability showed an excellent reliability (Vogt, 2005; Lindstrom, 2010). The selected body orientation angle (goalkeeper anticipation) measured through OpenPose correlated significantly with the goalkeeper strategy. The prediction model of the goalkeeper's strategy had its accuracy increased when the variable goalkeeper anticipation was included. This finding corroborates the applicability of OpenPose to obtain the body orientation of professional football players during matches (Sangüesa et al., 2019).

Goalkeepers face a clear trade-off between moving early and moving in the correct direction (Hunter et al., 2018). The goalkeeper's chance of successfully saving a penalty kick is lower than that of the penalty taker to score and he must try to reverse this disadvantage by positioning himself to anticipate the direction of the kick that is about to come (Kuhn, 1988). In this study, the goalkeeper tactical action (awaiting) and run-up speed of the penalty taker (slow) were associated with a kicker-dependent strategy (84.6%). To further improve this model, the inclusion of the correlated OpenPose variable (i.e., goalkeeper anticipation) correctly classified 97.0% of cases. Corroborating previous studies (Nakai et al., 2019; Sangüesa et al., 2019, 2020a,b), the analysis of the body orientation through OpenPose has proved to be extremely useful on penalty

TABLE 2 | OpenPose confidence score per variable.

Player	Body orientation angle	Confidence score
Penalty taker	Non-kick foot orientation	0.51
	Shoulders	0.87
	Hips	0.85
Goalkeeper	Anticipation	0.87
	Left foot	0.84
	Right foot	0.83

TABLE 3 | Test-retest reliability per variable.

Player	Body orientation angle	*r*
Penalty taker	Non-kick foot orientation	0.924*
	Shoulders	0.998*
	Hips	0.991*
Goalkeeper	Anticipation	0.998*
	Left foot	0.953*
	Right foot	0.961*

*$p < 0.05$.

TABLE 4 | Association between OSPAF and OpenPose variables with the goalkeeper strategy.

	OSPAF variables	χ^2	*p*	Cramer's V
Goalkeeper strategy	Run up speed	4.875	<0.05	0.354
	GK tactical action	26.542	<0.05	0.825
	OpenPose variable	r_{pb}	*p*	
	Goalkeeper anticipation	0.959	<0.05	

TABLE 5 | Association between OSPAF and OpenPose variables with the penalty taker strategy.

	OSPAF variables	χ^2	*p*	Cramer's V
Penalty taker strategy	Run up speed	2.300	<0.05	0.243
	Run up fluency	5.512	<0.05	0.376
	Gaze behavior	22.224	<0.05	0.755
	Deception	8.770	<0.05	0.474
	OpenPose variable	r_{pb}	*p*	
	Ball speed	0.927	<0.05	

kick analytics. The improvement in the model related to the goalkeeper strategy shows the important practical application through the evaluation of the body orientation of football players by using OpenPose as a tool. These findings support previous study by Sangüesa et al. (2019, 2020a,b) and Nakai et al. (2019), which showed that skeletal data recognized by OpenPose are found to be highly applicable with sufficient accuracy. The acquisition of a set of biometric human body part orientations implies an improvement of the analysis of the penalty kick in elite football. Moreover, its integration with video allows this model to be used as a coaching resource to assess players' orientation and improve training strategies for game preparation.

Previous study has shown that the penalty outcome depends, above all, on the emerging results of the "penalty taker—goalkeeper" dyadic interaction (Lopes et al., 2012; Almeida et al., 2016; Pinheiro et al., 2021b). In this study, lower degrees of goalkeeper anticipation, the goalkeeper tactical action (awaiting), and run-up speed of the penalty taker (slow) were associated with a kicker-dependent strategy. From a behavioral perspective, the present findings corroborate this dyadic interaction between the players in a penalty kick, as results showed that the goalkeeper strategy is influenced by the run-up speed of the penalty taker. Corroborating with this finding, Noël et al. (2021) indicated that goalkeepers must consider the penalty taker's run-up for deciding when to initiate their jump to the ball. It is presumed that more successful goalkeepers wait longer to decide for a goal side because this allows them to access more reliable information from the penalty taker's kicking actions to anticipate the penalty takers' intentions (Noël et al., 2021). Analytical procedures that integrate the study of criteria related to the interactions between opponents are highly recommended in game analysis in football (Sarmento et al., 2014). In real competitions, penalty kicks are an interaction process and the observable performance is rather the emergent result of this interaction process than the display of skills and abilities of the two parties (Lames, 2006). The new approach presented in this study, combining different methods, provides a deeper understanding of the player strategy in penalty kicks, through objective identification of the anticipation of the goalkeeper (i.e., angle: αGK measured via OpenPose). To further clarify the process of interaction in the penalty kick and the goalkeeper response time, future studies could introduce a time interval before the kick or an event (exact moment of the kick) as new variables with objective parameters to be analyzed by using OpenPose.

Regarding the penalty taker, the selected body angles measured through OpenPose did not associate significantly with the shooter strategy. A possible explanation could be that the biomechanical patterns of approaching the ball during the kick may vary from player to player, regardless of the strategy adopted. Previous study has shown that kicking from an approach angle of 45 and 60° may alter aspects of kick technique, such as enhancing pelvic rotation and thigh abduction of the kicking leg at impact (Scurr and Hall, 2009). Reinforcing this, Prassas et al. (1990) reported significant differences for a substantial number of variables, related to the kicking foot, leg, the non-kicking foot, trunk, and hip segments in football kicks.

A novelty of this study is the adoption of OpenPose measurements with notational analysis (i.e., OSPAF) to analyze penalty kicks. The OSPAF is an adequate and consistent instrument for analyzing successful and non-successful penalty kick patterns (Pinheiro et al., 2021b). The analysis of observational variables in penalty shooting may provide a general description of its technical execution, which allows for detecting the shooters and the goalkeeper's strategy based on the behavioral variables studied (Pinheiro et al., 2021b). Although the variables used to detect body angles possibly relevant to the analysis of strategy of the shooter in penalty kicks in football did not correlate significantly with the penalty taker strategy, the variables measured by OSPAF (i.e., run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed) were able to correctly classify 97.1% of the penalty taker strategy. The run-up speed slow, run-up fluency running with pauses, penalty taker gaze behavior not at the ball, the deception performed by the penalty taker, and lower ball speed were related to a goalkeeper-dependent strategy. Partially corroborating these findings, Noël et al. (2015) identified three variables (attention to the goalkeeper, run-up fluency, and kicking technique) that in combination could predict kick strategy in 92% of the penalties. Previous study had also shown that run-up and spatiotemporal patterns of gaze may differ between strategies (Noël and van der Kamp, 2012; Noël et al., 2015). The difference in fluency is probably a consequence of penalty takers who use a keeper-dependent strategy to increase time at the end of the run-up by waiting for the goalkeeper to commit to one side of the goal (van der Kamp, 2006). Studies in a realistic setup pointed those penalty takers by using the keeper-dependent strategy direct their gaze more toward the goalkeeper compared to the ball and the target location (Kurz et al., 2018). In contrast, penalty takers by using the keeper-independent strategy direct their gaze more toward the ball compared to the goalkeeper and the target location (Noël and van der Kamp, 2012).

Several studies have investigated the penalty kick strategies in football (van der Kamp, 2006; Noël et al., 2015, 2021; Pinheiro et al., 2021b). However, to the best of our knowledge, this is the first study to use OpenPose to detect relevant body orientation angles in penalty kicks in elite football from TV broadcast. This study is a preliminary study in penalty kick analysis and, thus, requires further examination. This study limitation was to not use a larger sample (e.g., full season), as it could bring practical applications and be more representative. Another limitation of this study was using only one viewing angle. It was included only one standard viewing angle and video quality was standardized, as recommended by (Sangüesa et al., 2020b). Nevertheless, for comparison of penalties from different viewing angles, a 3D transformation must be adopted when using OpenPose. Camera positioning (e.g., viewing angles) could affect the accuracy and, thus, the feasibility of the systems in practical settings (Zago et al., 2020). Nonetheless, this study presents an innovative approach to the analysis of penalty kicks in football, combining notational analysis with OpenPose. Its integration with video specification allows this model to be used as a coaching tool to assess players' orientation under different penalty kicks, improving sports preparation against upcoming opponents.

Multiple practical applications can be provided, from improving and refining player strategy in penalty kicks, to producing a precise assessment of player orientation in high-level competitive scenarios. Although it is not optimal to analyze only 34 penalty kicks, results from the present preliminary data indicate that it is possible to distinguish the goalkeeper's strategy (i.e., kicker dependent vs. kicker independent) based on the degree of goalkeeper anticipation, extracted through OpenPose. The body orientation analysis gives practitioners the potential to quickly evaluate the temporal decision-making of the goalkeeper (i.e., anticipation movement of the goalkeeper) with consideration to choosing when to initiate their jump to the ball. This could help to identify which goalkeepers move early or late in the penalty kick situation. Based on the pattern of anticipation of the goalkeeper in official competitions, specific training strategies can then be developed. Besides, having a time-based set of player orientations enhances analysts' ability to evaluate the relationship of on-ball and off-ball direction with the anatomical patterns. Posture analysis by using OpenPose has been verified to be practical with our model on the goalkeeper strategy identification. Future study could train a deep learning model to provide results about pose orientation automatically and faster.

CONCLUSION

This study tested an innovative approach in applying OpenPose measures integrated with notational analysis to investigate the factors influencing the players' strategy in penalty kicks. Results showed the applicability of OpenPose for in-match penalty kick analysis and an improvement in the prediction of the goalkeeper strategy by using a body orientation variable (anticipation) extracted via OpenPose. The goalkeeper degree of anticipation, tactical action, and run-up speed of the penalty taker can be associated with the goalkeeper strategy. Observable variables such as run-up speed, run-up fluency, penalty taker gaze behavior, deception by penalty taker, and ball speed may identify the shooter strategy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article. The video recordings used were public, thus authorization was not required from the observed players.

AUTHOR CONTRIBUTIONS

GP and ML contributed to the conception, design of this study, and wrote the first draft of the manuscript. GP and XJ organized the database. GP performed the statistical analysis. GP, VC, and ML contributed to the revision of the manuscript and read and approved the presented version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Funding as a doctoral scholarship was provided by the Katholische Akademische Ausländerdienst (KAAD); the publication costs were supported by the Technical University of Munich (TUM) Publishing Fund.

ACKNOWLEDGMENTS

The authors would like to thank the Chair for Performance Analysis and Sports Informatics of the TUM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fspor.2022.818556/full#supplementary-material

REFERENCES

Almeida, C. H., Volossovitch, A., and Duarte, R. (2016). Penalty kick outcomes in UEFA club competitions (2010–2015): the roles of situational, individual and performance factors. *Int. J. Perform. Anal. Sport*. 16, 508–522. doi: 10.1080/24748668.2016.11868905

Altman, D. G. (1991). *Practical Statistics for Medical Research*. Boca Raton, FL: CRC Press.

Aranda, R., González-Ródenas, J., López-Bondia, I., Aranda-Malavés, R., Tudela-Desantes, A., and Anguera, M. T. (2019). "REOFUT" as an observation tool for tactical analysis on offensive performance in soccer: mixed method perspective. *Front. Psychol*. 10:1476. doi: 10.3389/fpsyg.2019.01476

Cao, Z., Simon, T., Wei, S., and Sheikh, Y. (2017). "Realtime multi-person 2D pose estimation using part affinity fields," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Honolulu, HI: IEEE), 1302–1310. doi: 10.1109/CVPR.2017.143

Casal, C. A., Maneiro, R., Ardá, T., Marí, F. J., and Losada, J. L. (2017). Possession zone as a performance indicator in football. The game of the best teams. *Front. Psychol*. 8:1176. doi: 10.3389/fpsyg.2017.01176

Castellano, J., Alvarez, D., and Bradley, P. S. (2014). Evaluation of research using computerised tracking systems (Amisco® and Prozone®) to analyse physical performance in elite soccer: a systematic review. *Sports Med*. 44, 701–712. doi: 10.1007/s40279-014-0144-3

CMU-Perceptual-Computing-Lab (2017). *OpenPose: Real-Time Multi-Person Keypoint Detection Library for Body, Face, Hands, and Foot Estimation*. Available online at: https://github.com/CMU-Perceptual-Computing-Lab/openpose (accessed November 1, 2021).

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2). *Auflage. Hillsdale, NJ: Erlbaum.*

Fariña, A. R., Fábrica, G., Tambusso, P. S., and Alonso, R. (2013). Taking the goalkeeper's side in association football penalty kicks. *Int. J. Perform. Anal. Sport*. 13, 96–109. doi: 10.1080/24748668.2013.11868634

Fernandes, T., Camerino, O., Garganta, J., Pereira, R., and Barreira, D. (2019). Design and validation of an observational instrument for defence in soccer based on the Dynamical Systems Theory. *Int. J. Sports Sci. Coach*. 14, 138–152. doi: 10.1177/1747954119827283

Gabin, B., Camerino, O., Anguera, M. T., and Castañer, M. (2012). Lince: multiplatform sport analysis software. *Proc. Soc. Behav. Sci*. 46, 4692–4694. doi: 10.1016/j.sbspro.2012.06.320

Hartley, R., and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, *2nd Edn*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511811685

Hughes, M., and Bartlett, R. (2004). "The use of performance indicators in performance analysis," in *Notational Analysis of Sport, 2nd Edn*, eds M. Hughes and I. Franks (London: Routledge), 166–188.

Hunter, A. H., Murphy, S. C., Angilletta, M. J. Jr., and Wilson, R. S. (2018). Anticipating the direction of soccer penalty shots depends on the speed and technique of the kick. *Sports* 6:73. doi: 10.3390/sports6030073

Jamil, M., Littman, P., and Beato, M. (2020). Investigating inter-league and inter-nation variations of key determinants for penalty success across European football. *Int. J. Perform. Anal. Sport*. 20, 892–907. doi: 10.1080/24748668.2020.1794720

JASP Team (2021). *JASP (Version 0.16)[Computer software]*.

Kuhn, W. (1988). "Penalty-kick strategies for shooters and goalkeepers," in *Science and Football*, eds T. Reilly, K. Lees, and K. Davids (London: E and FN Spon), 489–492.

Kurz, J., Hegele, M., and Munzert, J. (2018). Gaze behaviour in a natural environment with a task-relevant distractor: how the presence of a goalkeeper distracts the penalty taker. *Front. Psychol*. 9:19. doi: 10.3389/fpsyg.2018.00019

Lames, M. (2006). Modelling the interaction in game sports - relative phase and moving correlations. *J. Sports Sci. Med*. 5, 556–560.

Lames, M., and Hansen, G. (2001). Designing observational systems to support top-level teams in game sports. *Int. J. Perform. Anal. Sport* 1, 85–91. doi: 10.1080/24748668.2001.11868251

Li, Y., Alexander, M., Glazebrook, C., and Leiter, J. (2015). Prediction of kick direction from kinematics during the soccer penalty kick. *Int. J. Kin. Sports Sci*. 3, 1–7. doi: 10.7575/aiac.ijkss.v.3n.4p.1

Lindstrom, D. (2010). *Schaum's Easy Outline of Statistics, 2nd Edn*. New York, NY: McGraw-Hill Education.

Lopes, J. E., Araújo, D., Duarte, R., Davids, K., and Fernandes, O. (2012). Instructional constraints on movement and performance of players in the penalty kick. *Int. J. Perform. Anal. Sport* 12, 311–345. doi: 10.1080/24748668.2012.11868602

Makaruk, H., Porter, J. M., Sadowski, J., Bodasińska, A., Zieliński, J., Niznikowski, T., et al. (2019). The effects of combining focus of attention and autonomy support on shot accuracy in the penalty kick. *PLoS ONE* 14:e0213487. doi: 10.1371/journal.pone.0213487

Memmert, D., and Nöel, B. (2020). *The Penalty Kick: The Psychology of Success*. Aachen: Meyer and Meyer Sport.

Nakai, M., Tsunodam, Y., Hayashi, H., and Murakoshi, H. (2019). "Prediction of basketball free throw shooting by OpenPose," in *New Frontiers in Artificial Intelligence. JSAI-isAI 2018. Lecture Notes in Computer Science*, *vol 11717*, eds K. Kojima, M. Sakamoto, K. Mineshima, and K. Satoh (Cham: Springer), 31. doi: 10.1007/978-3-030-31605-1_31

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., et al. (2020). Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras. *Front. Sports Act. Living*. 2:50. doi: 10.3389/fspor.2020.00050

Noël, B., Furley, P., van der Kamp, J., Dicks, M., and Memmert, D. (2015). The development of a method for identifying penalty kick strategies in association football. *J. Sports Sci*. 33, 1–10. doi: 10.1080/02640414.2014.926383

Noël, B., and van der Kamp, J. (2012). Gaze behaviour during the soccer penalty kick: an investigation of the effects of strategy and anxiety. *Int. J. Sport Psychol*. 43, 326–345.

Noël, B., van der Kamp, J., and Klatt, S. (2021). The interplay of goalkeepers and penalty takers affects their chances of success. *Front. Psychol*. 12:645312. doi: 10.3389/fpsyg.2021.645312

O'Donoghue, P. (2010). *Research Methods for Sports Performance Analysis*. London: Routledge.

Paterson, G., van der Kamp, J., and Savelsbergh, G. (2020). Moving advertisements systematically affect gaze behavior and performance in the soccer penalty kick. *Front. Sports Act. Living*. 1:69. doi: 10.3389/fspor.2019.00069

Pinheiro, G. S., Costa, V. T., and Lames, M. (2021a). "Penalty kicks in elite football: identifying factors related to the player strategy," in *International Society of Sport Psychology (ISSP) 15th World Congress Proceeding*, Taipei, 228–229, Pinheiro, G. S., Nascimento, V. B., Dicks, M., Costa, V. T., and Lames, M. (2021b). Design and validation of an observational system for penalty kick analysis in football (OSPAF). *Front. Psychol*. 12:661179. doi: 10.3389/fpsyg.2021.661179

Prassas, S. G., Terauds, J., and Nathan, T. (1990). "Three-dimensional kinematic analysis of high and low trajectory kicks in soccer," in *Proceedings of the VIIIth Symposium of the International Society of Biomechanics in Sports*, eds N. Nosek, D. Sojka, W. Morrison, and P. Susanka (Prague: Conex), 145–149.

Preciado, M., Anguera, M. T., Olarte, M., and Lapresa, D. (2019). Observational studies in male elite football: a systematic mixed study review. *Front. Psychol*. 10:2077. doi: 10.3389/fpsyg.2019.02077

Ramakrishna, V., Munoz, D., Hebert, M., Bagnell, J. A., and Sheikh, Y. (2014). "Pose machines: articulated pose estimation *via* inference machines," in *European Conference on Computer Vision* (Berlin: Springer), 33–47. doi: 10.1007/978-3-319-10605-2_3

Sangüesa, A. A., Haro, G., Ballester, C., and Martín, A. (2019). *Head, Shoulders, Hip and Ball... Hip and Ball! Using Pose Data to Leverage Football Player Orientation. Barça sports analytics summit*. Barcelona: Barça Innovation Hub.

Sangüesa, A. A., Martín, A., Fernández, J., Ballester, C., and Haro, G. (2020a). "Using player's body-orientation to model pass feasibility in soccer," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (Nashville, TN), 3875–3884.

Sangüesa, A. A., Martín, A., Fernandez, J., Rodríguez, C., Haro, G., and Ballester, C. (2020b). Always look on the bright side of the field: merging pose and contextual data to estimate orientation of soccer players. *arXiv e-prints* arXiv:2003.00943.

Sarmento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N., and Leitão, J. C. (2014). Match analysis in football: a systematic review. *J. Sports Sci*. 32, 1831–1843. doi: 10.1080/02640414.2014.898852

Scurr, J., and Hall, B. (2009). The effects of approach angle on penalty kicking accuracy and kick kinematics with recreational soccer players. *J. Sports Sci. Med*. 8, 230–234.

Soto, A., Camerino, O., Iglesias, X., Anguera, M. T., and Castañer, M. (2019). LINCE PLUS: research software for behavior video analysis. *Apunts Educ. Fís. Deportes* 137, 149–153. doi: 10.5672/apunts.2014-0983.es.(2019/3).137.11

van der Kamp, J. (2006). A field simulation study of the effectiveness of penalty kick strategies in soccer: late alterations of kick direction increase errors and reduce accuracy. *J. Sports Sci*. 24, 467–477. doi: 10.1080/02640410500190841

Vilagut, G. (2014). "Test-retest reliability," in *Encyclopedia of Quality of Life and Well-Being Research*, eds A. C. Michalos (Dordrecht: Springer), 3001. doi: 10.1007/978-94-007-0753-5_3001

Vogt, W. P. (2005). *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*. Newcastle upon Tyne: Sage.

Wei, S., Ramakrishna, V., Kanade, T., and Sheikh, Y. (2016). "Convolutional pose machines," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (San Juan, PR: IEEE), 4724–4732. doi: 10.1109/CVPR.2016.511

Zago, M., Luzzago, M., Marangoni, T., De Cecco, M., Tarabini, M., and Galli, M. (2020). 3D tracking of human motion using visual skeletonization and stereoscopic vision. *Front. Bioeng. Biotechnol*. 8:181. doi: 10.3389/fbioe.2020.00181

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Sports and Active Living

BRIEF RESEARCH REPORT
published: 16 June 2022
doi: 10.3389/fspor.2022.921418

Automated Urinal-Based Specific Gravity Measurement Device for Real-Time Hydration Monitoring in Male Athletes

Brian F. Bender, Nick J. Johnson, Jasmine A. Berry, Kelvin M. Frazier and Michael B. Bender*

Intake Health, Raleigh, NC, United States

OPEN ACCESS

Edited by:
Cain Craig Truman Clark, Coventry University, United Kingdom

Reviewed by:
Marijn M. Speeckaert, Ghent University Hospital, Belgium
Jakub Chycki, Jerzy Kukuczka Academy of Physical Education in Katowice, Poland

***Correspondence:**
Brian F. Bender
brian@intake.health

Specialty section:
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

Received: 15 April 2022
Accepted: 24 May 2022
Published: 16 June 2022

Citation:
Bender BF, Johnson NJ, Berry JA, Frazier KM and Bender MB (2022) Automated Urinal-Based Specific Gravity Measurement Device for Real-Time Hydration Monitoring in Male Athletes. Front. Sports Act. Living 4:921418. doi: 10.3389/fspor.2022.921418

Acute and chronic hydration status is important for athlete safety and performance and is frequently measured by sports scientists and performance staff in team environments via urinalysis. However, the time required for urine collection, staff testing, and reporting often delays immediate reporting and personalized nutrition insight in situations of acute hydration management before training or competition. Furthermore, the burdensome urine collection and testing process often renders chronic hydration monitoring sporadic or non-existent in real-world settings. An automated urinalysis device (InFlow) was developed to measure specific gravity, an index of hydration status, in real-time during urination. The device was strongly correlated to optical refractometry with a mean absolute error of 0.0029 (±0.0021). Our results show this device provides a novel and useful approach for real-time hydration status via urinalysis for male athletes in team environments with high testing frequency demands.

Keywords: urinalysis, hydration, sports science, sports technology, wellness, safety, athletic performance

INTRODUCTION

Water is essential for life, playing such vital physiological roles as a cellular and tissue building material, a solvent and reaction medium, a carrier of nutrients and waste, and a medium for thermoregulation and shock absorption (Jéquier and Constant, 2010). As dehydration ensues and leads to a state of hypohydration, negative impacts on blood flow, skeletal muscle metabolism, cardiovascular strain, and thermoregulation often lead to impaired physiological function and athletic performance such as a shorter time to exhaustion and lower exercise intensity (Cheuvront and Kenefick, 2014). This is particularly true for athletes and other highly active individuals, where sweat output is high and performance optimization is a top priority (Sawka et al., 2007b). It is well-known that dehydration impairs aerobic performance and is increasingly demonstrating impairment in areas of strength and power, cognitive function, mood, and sleep (Cheuvront and Kenefick, 2014; Harris et al., 2019; Deshayes et al., 2020). Maintenance of euhydration has been stressed for endurance athletes. However, a state of hypohydration has been shown to negatively affect skill-based performance metrics in sports such as soccer (McGregor et al., 1999; Edwards et al., 2007) and basketball (Baker et al., 2007a,b).

The extreme importance of euhydration on preserving organ function and health has resulted in the evolution of sensitive and precise homeostatic mechanisms to maintain fluid and electrolyte balance and results in physiological changes that have been used as biomarkers of hydration status (Jéquier and Constant, 2010). One regulatory mechanism is related to thirst, generated via a neuroendocrine response to the osmotically driven shrinking of cells when water deficits result in intracellular water leaving the cell to dilute an overly ionic extracellular fluid space (Cheuvront and Kenefick, 2014; Leib et al., 2016). Another physiological mechanism triggered during intracellular volume contraction is signaling from the antidiuretic hormone vasopressin, triggering the kidneys to produce a smaller volume of more concentrated urine (Popkin et al., 2010). This unique role of the kidneys to regulate blood osmolality is what has led to the use of several urine indices as biomarkers of hydration status, including urine osmolality, urine specific gravity (USG), 24-h urine volume, urine color, and urine conductivity (Armstrong et al., 1994).

A more concentrated urine sample, as indicated by a higher urine osmolality, higher urine specific gravity, lower 24-h urine volume, darker urine color, or higher urine conductivity, correlates to other commonly used biomarkers of hypohydration status such as blood plasma osmolality and body mass decrease. Urine osmolality is typically measured via freezing point depression and represents the concentration of all solutes in solution. Urine specific gravity measures the density of the urine solution relative to water and thus heavier solutes, such as glucose and creatinine, can bias the results. Urine color is often measured via comparison to color charts and can be a quick and easy method but is subject to user error and some potential confounding physiological conditions or presence of supplements. Urine conductivity is a function of conductive species in solution, largely sodium, and correlates to total solute concentration. All these techniques trend together. However, no individual measurement can provide a complete picture of hydration status, nor can each be reliable in all individuals and for all use cases. For example, a rugby player with exceptionally high lean body mass typically excretes higher rates of larger molecules like creatinine that bias urine specific gravity toward the higher end of the scale, suggesting a more hypohydrated state when compared to leaner runners, despite similar blood plasma osmolality measurements (Hamouti et al., 2010). Furthermore, low Index of Individuality (II) for several of these biomarkers (Cheuvront et al., 2010) leaves the need for repeated testing and individual baselining important for better assessing dynamic hydration status (Cheuvront et al., 2011).

Measuring athlete hydration status in real-world settings is often difficult, necessitating a balance between accuracy, cost, and ease-of-use (Belval et al., 2019). Methods for measuring hydration have been reviewed elsewhere, including their benefits and limitations (Barley et al., 2020). For example, body mass change and bioelectrical impedance analysis (BIA) are non-invasive and relatively simple (players only need to stand on the device for a few moments). However, as with all hydration assessment techniques (Armstrong, 2007), these methods possess limitations. Confounding activities include recent food ingestion, fluid ingestion, urination, defecation, and intensive physical activity (Mialich et al., 2014). In addition, due to both logistical challenges and reliability, USG is generally recommended over BIA in athletic settings for serial hydration assessment (Barley et al., 2020). The current processes most often used for urine-testing are manual, requiring players to urinate into cups which are later collected by staff that perform dipstick or optical refractometer testing. This is both labor and time-intensive and results in infrequent testing and/or delayed reporting. Optimal solutions are often dubbed "invisible monitoring," which require no athlete burden and facilitate buy-in (Windt et al., 2020). In addition, some of the best results for player optimization of health and safety comes from player empowerment that drives self-regulation (Kim and Cruz, 2021). An automated, accurate USG measurement device that allows players to self-monitor hydration status could provide high-compliance testing and improved hydration awareness.

MATERIALS AND METHODS

Study Design

The study was carried out in February 2022 at three public U.S. University athletic training facilities. All measurements were performed using surplus human urine samples to requirement ($\geq$50 mL) from routine testing. Anonymized samples were used for all experiments. Samples were stored at room temperature, did not undergo any processing or centrifugation, and were analyzed within 2 h of sample collection.

The use of patient samples complied with all relevant national regulations and institutional policies. The study does not conform to NIH definition of a Clinical Trail per NOT-OD-15-015. In addition, the study does not conform to the definition of human subjects research per 45 CFR 46, as only unidentifiable surplus samples from routine testing were used in the study.

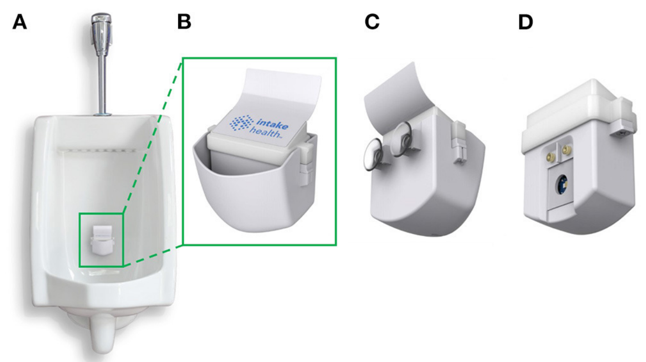

FIGURE 1 | The InFlow Urinalysis System. **(A)** Fully assembled system installed in a urinal using rear suction cups capable of catching a stream of urine in real-time. Urine fills the cup at a rate faster than it can drain through a hole in the bottom. **(B)** The front, fully assembled InFlow system magnified. LED lights under the top of the insert housing flash red, yellow, or green to categorize results for the user as dehydrated, mildly dehydrated, and hydrated, respectively. **(C)** Rear view of the full assembly showing suction cups for installation. **(D)** Rear view of the removed insert sub-assembly showing the sensors that interface with the urine sample during use. As urine fills the cup, fluid engulfs the insert from the bottom. The metal electrodes register the presence of fluid by shorting the pin voltage and initiate testing via the LED and photodiodes residing behind a glass window.

TABLE 1 | Imprecision of InFlow system measured in USG units using artificial urine control.

Material	Mean USG	Total imprecision (SD)	Total imprecision (%CV)	Within unit (SD)	Between unit (SD)
Artificial urine control	1.0325	0.0009	0.09%	0.0001	0.0009

Participants included 151 NCAA male football athletes from three collegiate institutions. Urine samples were collected into plastic cups by each player and brought to performance staff and dietitians during the normal course of their activities for USG testing using manual (Teckoplus) and digital optical refractometers (MISCO Palm Abbe and Atago 3741 PEN) and dipsticks (Diagnox Urinox-10) for a subset of the tests. After normal testing, the surplus urine samples were poured through the InFlow system.

InFlow System

The InFlow system is designed to capture urine in real-time during a urination event from a urinal (**Figure 1A**). The system has a cup to easily catch and fill with urine (**Figure 1B**). The system is installed by pressing the unit against the wall of a urinal using the suction cups on the back of the device (**Figure 1C**). During urination, urine quickly fills the cup volume faster than it can drain through a small hole in the bottom of the cup. A removable insert housed within the cup (**Figure 1D**) holds the electronics, sensors, and power. As the cup fills, the fluid covers the testing chamber, turning on the system and performing a test in <2 s.

Analytical Imprecision and Bias

For analytical testing and quantitative analysis of the InFlow system, the mean (μ) and standard deviation (SD) were calculated. Method comparison results for the InFlow system were assessed using Bland-Altman difference plots and regression analysis (including Pearson's r correlation coefficients) for quantitative parameters (USG). Confidence intervals and prediction intervals at 95% were calculated for InFlow performance against the manual optical refractometer.

Analytical system performance was assessed using artificial urine control (Aldon Life Sciences, IS5070). The SD and coefficient of variation (CV%) (SD/μ × 100) of total imprecision were calculated by testing artificial urine control across 10 units in triplicate. The within-unit SD was calculated as the average SD across triplicate back-to-back runs from the same unit across 10 units. The between-unit SD was calculated using artificial urine control across 10 units. For each pool, the "observed" reference USG value was established for each specimen using a manual optical refractometer and taking the mean USG. The InFlow system mean and SD are derived from measurements through the urinalysis device. Data was analyzed using the Westgard model, using Total Error (TE) and TE (%) defined by Equations (1) and (2), respectively. The threshold used for acceptable percent Total Allowable Error (TAE%) for USG was ±0.6% (Ricós et al., 1999). Results provided from analytical sensitivity experiments were rounded to 4 decimal places, except TE (%) which was rounded to 2 decimal places.

$$Total\ Error\ (TE) = |Bias| + 2SD \tag{1}$$

$$TE\,(\%) = \left(TE \div reference\ mean\right) \times 100 \tag{2}$$

Diagnostic Performance

Assessment of classification of dehydration was performed at a criterion value (USG ≥ 1.020) designated by the American College of Sports Medicine (ACSM) and the National Athletic Training Association (NATA) (Casa et al., 2000; Sawka et al., 2007a). A positive result was assigned to a dehydrated sample and a negative result was assigned to a euhydrated sample. True Positives (TP) were assigned to samples the InFlow system classified as a positive result when the manual optical refractometer reported a positive result, and True Negatives (TN) were assigned to samples the InFlow system classified as a negative result when the manual optical refractometer reported a negative result. In contrast, False Positives (FP) were assigned to samples the InFlow system classified as a positive result when the manual optical refractometer reported a negative result, and False Negatives (FN) were assigned to samples the InFlow system classified as a negative result when the manual optical refractometer reported a positive result. Diagnostic accuracy, sensitivity, specificity, and precision were calculated (Zweig and Campbell, 1993).

Receiver operator characteristic (ROC) analysis was performed to assess diagnostic accuracy represented by the area under the ROC curve (AUC) (Zweig and Campbell, 1993). There is no established analytical goal for dehydration; a recommended minimum of 80% for sensitivity and specificity was used (Zweig and Campbell, 1993; Cheuvront et al., 2010), which would represent odds of 4 to 1 in favor of a correct classification.

RESULTS

The distribution of USG values among the sample population ranged from 1.003 to 1.036 (**Supplementary Figure 1**). The mean USG was 1.018 (±0.009) and approximately 45% of samples tested were hypohydrated (USG ≥1.020). This distribution USG mean is similar to population USG means of other athletes (1.018 ± 0.009) prior to exercise (Stover et al., 2006). The percentage of hypohydrated samples is less but similar to (66%) other NCAA athlete samples (Volpe et al., 2009) and the range and distribution (SD) provide an adequate and representative array of urine samples for system performance evaluation.

The InFlow system's design was chosen to minimize testing burden on the user while maintaining adequate accuracy for hydration reporting consistent with existing protocols and testing equipment. Results from precision studies are shown in **Table 1**. Very low total imprecision was demonstrated using artificial urine control material. Imprecision estimates (within unit, between unit) were estimated in USG "units." The within-unit imprecision of 0.0001 is comparable to digital optical refractometer resolution (0.0001) (Atago, 2022; MISCO, 2022).

Analytical performance studies (**Figure 2A**) demonstrated strong correlation to manual optical refractometry ($r = 0.90$; $n =$ 247 specimens, USG range 1.003–1.036). The mean population error (**Figure 2B**) was not significantly different than zero at any USG range with root mean squared error (RMSE) of 0.0036. The mean absolute error (±SD) as a function of USG (0.0029 ± 0.0022) tended to trend in a positive direction but this trend was not significant (**Supplementary Figure 2**). The error between the InFlow system and a manual optical refractometer was compared against the error between digital optical refractometers and the manual optical refractometer via Bland-Altman plot of agreement (**Figure 2C**). All points representing error between InFlow results and the manual optical refractometer fell between the limits of agreement established by clinically relevant USG thresholds based on inter- and intraindividual variability for hydration assessment (Cheuvront et al., 2010, 2011).

Analytical imprecision and bias testing demonstrated acceptable TE (%) [defined as <0.6% TAE(%)] at all USG levels tested (**Table 2**). The analytical imprecision, CV_A, of 0.09% was below half the intraindividual variation, CV_I, for USG (0.4%), as recommended in clinical chemistry best practices (Fraser and Harris, 1989; White et al., 2004; Cheuvront et al., 2010).

Diagnostic performance evaluation yielded an accuracy of 87%, a sensitivity of 87%, a specificity of 88%, and a precision of 85% (TP = 96; TN = 120; FP = 17; FN = 14). These values exceed cutoff values for sensitivity and sensitivity of 70% used elsewhere for urine-based hydration classification at 1.020 (Hooper et al., 2016) and our own analytical goal of 80% (Cheuvront et al., 2010). These assessments demonstrate the InFlow system performs adequately for USG testing for hydration assessment given USG's inter- and intraindividual variability. ROC analysis produced an AUC of 0.94, providing evidence of generally acceptable diagnostic accuracy (**Supplementary Figure 3**).

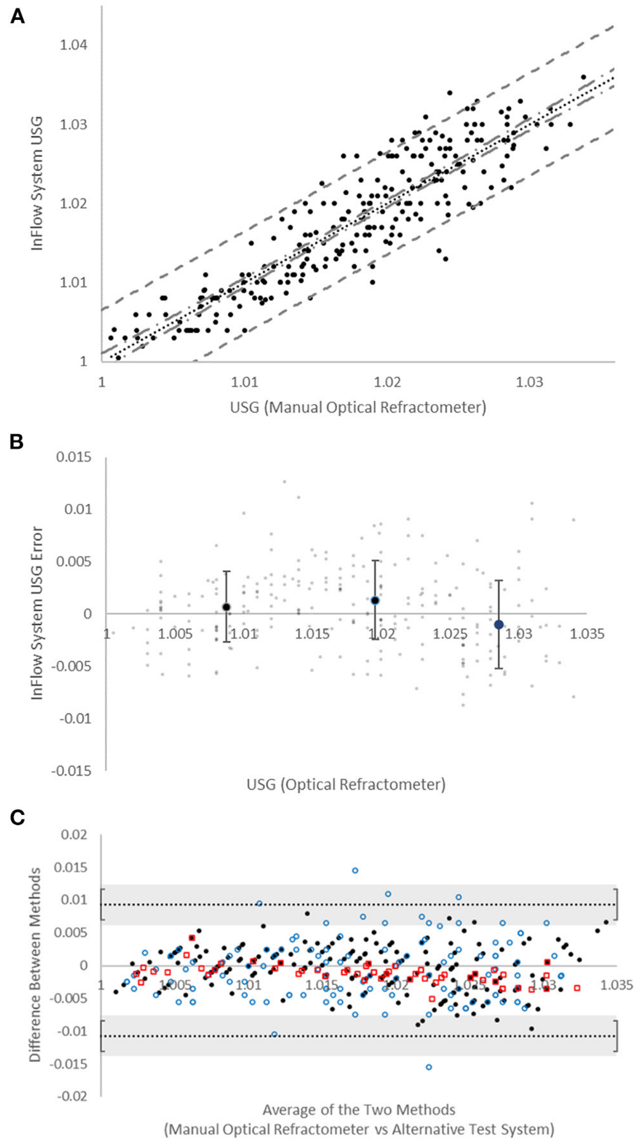

FIGURE 2 | Device accuracy. **(A)** Comparison of manual optical refractometer (x-axis) versus InFlow system (y-axis) results. Dotted line (···) is linear regression ($r = 0.90$). Dot-dash line (- · -) represents 95% confidence interval. Dashed line (- -) represents 95% prediction interval. **(B)** Bias (in USG "units") of InFlow system vs. manual optical refractometry. Large dots represent averages at each USG range (<1.015, 1.015–1.025, >1.025). Error bars represent SD. Small dots represent individual test results. **(C)** Bland-Altman plots of agreement between the manual optical refractometer with the InFlow system (●), the MISCO digital optical refractometer (○), and the Atago digital optical refractometer (□). All systems fall with the agreement limits set at the reference change value (0.010) for USG established via CV_I and CV_G (Cheuvront et al., 2010, 2011). Error bars represent the SD of the agreement limits.

DISCUSSION

Experiments demonstrated acceptable analytical and diagnostic performance of the automated InFlow system for USG, including imprecision and accuracy. The instrument worked reliably with sample flowing through the system, which is analogous to real-world use in a urinal. We did not encounter ambient lighting-related difficulties (which others have reported with digital optical refractometers) (Minton et al., 2015), because the InFlow system measurement chamber is internally housed and shielded from external lighting conditions. No staining of the glass window occurred during the study or during prolonged benchtop testing of over 45 days of testing. Power analyses demonstrated a low sleep current while not in use of around 10 μA and a test current output of approximately 0.41 mAh (**Supplementary Figure 4**). The InFlow system performed over 5,000 tests per charge, similar to the digital optical refractometers

TABLE 2 | Analytical sensitivity of Inflow system by USG range as measured against manual optical refractometry.

USG range	Number of samples in set	Observed mean of sample set	InFlow mean of samples	InFlow SD	Bias	TE (%)
<1.015	96	1.0088	1.0095	0.0023	0.0007	0.53
1.015–1.025	85	1.0196	1.0209	0.0023	0.0013	0.59
>1.025	66	1.0286	1.0276	0.0025	–0.0010	0.59

tests. The InFlow system includes a wireless Qi charging system for simple battery recharging.

The InFlow system was compared to manual optical refractometry for error analysis. The mean absolute error (±SD) between manual optical refractometry and the InFlow system was 0.0029 (±0.0021). To compare these results to other available tools for measuring USG, two digital optical refractometers were compared to manual optical refractometry (**Supplementary Figure 5**). The mean absolute error of the MISCO digital optical refractometer (0.0038 ± 0.0047) was higher than the InFlow system, but the mean absolute error of the Atago digital optical refractometer (0.0016 ± 0.0011) was lower than the InFlow system. All three systems were deemed interchangeable for use in hydration assessment via USG in a sports environment based on Bland-Altman analysis (**Figure 2C**). Based on recommendations for setting Bland-Altman agreement limits on biologically and analytically relevant criteria (Giavarina, 2015), limits of ±0.010 were used as defined by the reference change value for USG-based hydration assessment given its intraindividual ($CV_I = 0.4$) and interindividual ($CV_G = 1.0$) variability (Cheuvront et al., 2010, 2011). These results generally point to interchangeability between the digital optical refractometers and the InFlow system for use in USG reporting (Giavarina, 2015). All datapoints fell within the limits with the exception of 4% ($n = 5$) of MISCO digital optical refractometer readings.

Urine dipstick testing has known error associated with manual color comparison, lighting variation, sample size variation, timing variation, and the inherent variability associated with USG binning by 0.005 USG unit increments (de Buys Roessingh et al., 2001; Smith et al., 2017). A subset of urine samples ($n = 119$) was randomly selected for urine dipstick analysis. USG error (±SD) for dipstick testing compared against manual optical refractometry was 0.0051 (±0.0047) with $r = 0.76$ (**Supplementary Figure 6**). This error was significantly higher than the InFlow system ($p < 0.001$; $\alpha = 0.05$).

Although there are a number of commercially available handheld digital optical refractometers including Palm Abbe (MISCO); PEN, UG-α, PAL-10S (Atago); Clinic-Chek, USG-Check, and TS METER D (Reichert; Depew, NY), this system represents the first automated device designed to measure USG from a urinal in real-time as the individual urinates into the system. This represents a significant improvement to USG testing in high frequency testing environments such as those found within collegiate and professional athletic programs. Testing time is a significant hurdle to large-scale, frequent hydration testing in team settings. This leads to delayed action. It is not uncommon for players to already leave the locker room and begin training or competition before hydration results have been measured, assessed, and reported. Similarly, the high burden on performance staff, necessitating tracking down players, handling urine cups, labeling, testing, and reporting, leads to a sporadic testing schedule. The InFlow system significantly reduces, and at times eliminates, the testing and reporting time by analyzing results in real-time during the act of urination and reporting those results directly to the player instantaneously.

The InFlow system provides significant improvement over manual quality control (QC) errors common to clinical testing procedures. Pre-analytical errors typically account for most QC errors and include mislabeling of sample containers and sorting errors (Delanghe and Speeckaert, 2014). Similarly, post-analytical errors such as mistakes in data transcription are common (Hammerling, 2012). InFlow's automated testing framework eliminates the need for sample collection, labeling, and data transcription and thus reduces, or eliminates, these common QC errors.

Future areas of research may include assessment of varying physiological and environmental conditions that present in altered urine color, such as conditions like rhabdomyolysis or medication/supplement use. Future research will also compare the InFlow system to other markers of hydration, such as blood and urine osmolality, alongside mechanistic studies in urine composition to improve accuracy and error reporting such as albumin (known to present during intense physical exercise) and creatinine (known to exist in higher concentrations in individuals with high lean muscle mass). In addition, broadening and diversifying the sample of users may improve the test statistics. Similarly, this system may be assessed for useability and accuracy in other environments that may provide benefit such as within industrial settings, military settings, and general consumer-facing health and wellness settings. Finally, altered designs for use among female athletes is another area of ongoing research.

In conclusion, the automated InFlow system was demonstrated to be a fast, simple, and accurate way to measure USG. The InFlow system met accuracy requirements for reliably monitoring USG for hydration assessment given its biological variation within an automated testing platform for male users from a urinal measured during the act of urination. The results from this report may prove valuable for those interested in evaluating use of the InFlow system in a variety of settings and applications in long-term, longitudinal hydration monitoring and behavior change studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

BB conceived the analytical approach, derived the algorithms, analyzed the data, and wrote and edited the manuscript. BB, NJ, and MB conceived the overall system design, built and tested prototype hardware and firmware, and performed sample testing. MB assisted in field test data collection. JB assisted in statistical data analysis and manuscript editing. KF assisted with manuscript editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by funding from the National Science Foundation under grant 2026127. The views and opinions of the authors do not reflect those of the National Science Foundation.

ACKNOWLEDGMENTS

The authors would like to acknowledge Mighty Studios, Inc. for their work on product design efforts for manufacturing and associated images.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fspor.2022.921418/full#supplementary-material

REFERENCES

Armstrong, L. E. (2007). Assessing hydration status: the elusive gold standard. *J. Am. Coll. Nutr.* 26, 575S−584S. doi: 10.1080/07315724.2007.10719661

Armstrong, L. E., Maresh, C. M., Castellani, J. W., Bergeron, M. F., Kenefick, R. W., LaGasse, K. E., et al. (1994). Urinary indices of hydration status. *Int. J. Sport Nutr.* 4, 265–279. doi: 10.1123/ijsn.4.3.265

Atago (2022). *Clinical Refractometers*. Available online at: https://www.atago.net/USA/products_clinical.php

Baker, L. B., Conroy, D. E., and Kenney, W. L. (2007a). Dehydration impairs vigilance-related attention in male basketball players. *Med. Sci. Sports Exerc.* 39, 976–983. doi: 10.1097/mss.0b013e3180471ff2

Baker, L. B., Dougherty, K. A., Chow, M., and Kenney, W. L. (2007b). Progressive dehydration causes a progressive decline in basketball skill performance. *Med. Sci. Sports Exerc.* 39, 1114–1123. doi: 10.1249/mss.0b013e3180574b02

Barley, O. R., Chapman, D. W., and Abbiss, C. R. (2020). Reviewing the current methods of assessing hydration in athletes. *J. Int. Soc. Sports Nutr.* 17, 52. doi: 10.1186/s12970-020-00381-6

Belval, L. N., Hosokawa, Y., Casa, D. J., Adams, W. M., Armstrong, L. E., Baker, L. B., et al. (2019). Practical hydration solutions for sports. *Nutrients* 11, 1550. doi: 10.3390/nu11071550

Casa, D. J., Armstrong, L. E., Hillman, S. K., Montain, S. J., Reiff, R., Rich, B. S., et al. (2000). National athletic trainers' association position statement: fluid replacement for athletes. *J. Athletic Train.* 35, 212–224.

Cheuvront, S. N., Ely, B. R., Kenefick, R. W., and Sawka, M. N. (2010). Biological variation and diagnostic accuracy of dehydration assessment markers. *Am. J. Clin. Nutr.* 92, 565–573. doi: 10.3945/ajcn.2010.29490

Cheuvront, S. N., Fraser, C. G., Kenefick, R. W., Ely, B. R., and Sawka, M. N. (2011). Reference change values for monitoring dehydration. *Clin. Chem. La. Med.* 49, 1033–1037. doi: 10.1515/CCLM.2011.170

Cheuvront, S. N., and Kenefick, R. W. (2014). "Dehydration: physiology, assessment, and performance effects," in *Comprehensive Physiology*, ed R. Terjung (Hoboken, NJ: John Wiley & Sons, Inc.), 257–285. doi: 10.1002/cphy.c130017

de Buys Roessingh, A. S., Drukker, A., and Guignard, J. P. (2001). Dipstick measurements of urine specific gravity are unreliable. *Arch. Dis. Childhood* 85, 155–157. doi: 10.1136/adc.85.2.155

Delanghe, J., and Speeckaert, M. (2014). Preanalytical requirements of urinalysis. *Biochem. Med.* 24, 89–104. doi: 10.11613/BM.2014.011

Deshayes, T. A., Jeker, D., and Goulet, E. D. B. (2020). Impact of pre-exercise hypohydration on aerobic exercise performance, peak oxygen consumption and oxygen consumption at lactate threshold: a systematic review with meta-analysis. *Sports Med.* 50, 581–596. doi: 10.1007/s40279-019-01223-5

Edwards, A. M., Mann, M. E., Marfell-Jones, M. J., Rankin, D. M., Noakes, T. D., and Shillington, D. P. (2007). Influence of moderate dehydration on soccer performance: physiological responses to 45 min of outdoor match-play and the immediate subsequent performance of sport-specific and mental concentration tests. *Brit. J. Sports Med.* 41, 385–391. doi: 10.1136/bjsm.2006.033860

Fraser, C. G., and Harris, E. K. (1989). Generation and application of data on biological variation in clinical chemistry. *Crit. Rev. Clin. Lab. Sci.* 27, 409–437. doi: 10.3109/10408368909106595

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochem. Med.* 25, 141–151. doi: 10.11613/BM.2015.015

Hammerling, J. A. (2012). A review of medical errors in laboratory diagnostics and where we are today. *Lab. Med.* 43, 41–44. doi: 10.1309/LM6ER9WJR1IHQAUY

Hamouti, N., del Coso, J., Ávila, A., and Mora-Rodriguez, R. (2010). Effects of athletes' muscle mass on urinary markers of hydration status. *Eur. J. Appl. Physiol.* 109, 213–219. doi: 10.1007/s00421-009-1333-x

Harris, P. R., Keen, D. A., Constantopoulos, E., Weninger, S. N., Hines, E., Koppinger, M. P., et al. (2019). Fluid type influences acute hydration and muscle performance recovery in human subjects. *J. Int. Soc. Sports Nutr.* 16, 15. doi: 10.1186/s12970-019-0282-y

Hooper, L., Bunn, D. K., Abdelhamid, A., Gillings, R., Jennings, A., Maas, K., et al. (2016). Water-loss (intracellular) dehydration assessed using urinary tests: how well do they work? Diagnostic accuracy in older people. *Am. J. Clin. Nutr.* 104, 121–131. doi: 10.3945/ajcn.115.119925

Jéquier, E., and Constant, F. (2010). Water as an essential nutrient: the physiological basis of hydration. *Eur. J. Clin. Nutr.* 64, 115–123. doi: 10.1038/ejcn.2009.111

Kim, H.-D., and Cruz, A. B. (2021). Psychological influence of self-management on exercise self-confidence, satisfaction, and commitment of martial arts practitioners in Korea: a meta-analytic approach. *Front. Psychol.* 12, 691974. doi: 10.3389/fpsyg.2021.691974

Leib, D. E., Zimmerman, C. A., and Knight, Z. A. (2016). Thirst. *Curr. Biol.* 26, R1260–R1265. doi: 10.1016/j.cub.2016.11.019

McGregor, S. J., Nicholas, C. W., and Lakomy, H. K. (1999). The influence of intermittent high-intensity shuttle running and fluid ingestion on the performance of a soccer skill. *J. Sports Sci.* 17, 895–903. doi: 10.1080/026404199365452

Mialich, M. S., Maria, J., Sicchieri, F., Afonso, A., and Junior, J. (2014). Analysis of body composition : a critical review of the use of bioelectrical impedance analysis. *Int. J. Clin. Nutr.* 2, 1–10. doi: 10.12691/ijcn-2-1-1

Minton, D. M., O'Neal, E. K., and Torres-McGehee, T. M. (2015). Agreement of urine specific gravity measurements between manual and digital refractometers. *J. Athletic Train.* 50, 59–64. doi: 10.4085/1062-6050-49.3.47

MISCO (2022). *Human Urine Refractometer Specifications.* Available online at: https://www.misco.com/product/human-urine-refractometer-specific-gravity-total-solids-pa202x-093-095/

Popkin, B. M., D'Anci, K. E., and Rosenberg, I. H. (2010). Water, hydration, and health. *Nutr. Rev.* 68, 439–458. doi: 10.1111/j.1753-4887.2010.00304.x

Ricós, C., Alvarez, V., Cava, F., García-Lario, J., v., Hernández, A., et al. (1999). Current databases on biological variation: pros, cons and progress. *Scand. J. Clin. Lab. Investig.* 59, 491–500. doi: 10.1080/00365519950185229

Sawka, M. N., Burke, L. M., Eichner, E. R., Maughan, R. J., Montain, S. J., and Stachenfeld, N. S. (2007a). Exercise and fluid replacement. *Med. Sci. Sports Exerc.* 39, 377–390. doi: 10.1249/mss.0b013e31802ca597

Sawka, M. N., Burke, L. M., Eichner, E. R., Maughan, R. J., Montain, S. J., and Stachenfeld, N. S. (2007b). American College of Sports Medicine position stand. Exercise and fluid replacement. *Med. Sci. Sports Exerc.* 39, 377–390.

Smith, G. T., Dwork, N., Khan, S. A., Millet, M., Magar, K., Javanmard, M., et al. (2017). "Reducing user error in dipstick urinalysis with a low-cost slipping manifold and mobile phone platform (Conference Presentation)," in *Proc. SPIE*. (San Francisco, CA). doi: 10.1117/12.2256593

Stover, E. A., Petrie, H. J., Passe, D., Horswill, C. A., Murray, B., and Wildman, R. (2006). Urine specific gravity in exercisers prior to physical training. *Appl. Physiol. Nutr. Metab.* 31, 320–327. doi: 10.1139/h06-004

Volpe, S. L., Poule, K. A., and Bland, E. G. (2009). Estimation of prepractice hydration status of national collegiate athletic association division I athletes. *J. Athlet. Train.* 44, 624–629. doi: 10.4085/1062-6050-44.6.624

White, G. H., Farrance, I., and AACB Uncertainty of Measurement Working Group (2004). Uncertainty of measurement in quantitative medical testing: a laboratory implementation guide. *Clin. Biochem. Rev.* 25, S1–S24.

Windt, J., MacDonald, K., Taylor, D., Zumbo, B. D., Sporer, B. C., and Martin, D. T. (2020). "To tech or not to tech?" A critical decision-making framework for implementing technology in sport. *J. Athlet. Train.* 55, 902–910. doi: 10.4085/1062-6050-0540.19

Zweig, M. H., and Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* 39, 561–577. doi: 10.1093/clinchem/39.4.561

Conflict of Interest: The authors declare that they are employees of Bender Tech, LLC (dba Intake Health) and receive salaries through their positions with the company.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Sports and Active Living

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fspor.2022.937525

Check for updates

OPEN ACCESS

EDITED BY
Cain Craig Truman Clark,
Coventry University, United Kingdom

REVIEWED BY
Hajnalka Vago,
Semmelweis University, Hungary
Ivan Tomasic,
Mälardalen University, Sweden

*CORRESPONDENCE
Ådne Ausland
adne_ausland@outlook.com

SPECIALTY SECTION
This article was submitted to Sports Science, Technology and Engineering, a section of the journal Frontiers in Sports and Active Living

RECEIVED 06 May 2022
ACCEPTED 06 July 2022
PUBLISHED 25 July 2022

CITATION
Ausland Å, Sandberg EL, Jortveit J and Seiler S (2022) Heart rhythm assessment in elite endurance athletes: A better method? *Front. Sports Act. Living* 4:937525. doi: 10.3389/fspor.2022.937525

Heart rhythm assessment in elite endurance athletes: A better method?

Ådne Ausland[1]*, Edvard Liljedahl Sandberg[2], Jarle Jortveit[2] and Stephen Seiler[1]

[1]Department of Sport Science and Physical Education, Faculty of Health and Sport Sciences, University of Agder, Kristiansand, Norway, [2]Department of Cardiology, Sorlandet Hospital, Arendal, Norway

Introduction: Arrhythmias also occur among elite endurance athletes. Conventional diagnostic tools for assessment of arrhythmias suffer from limited availability and usability challenges, particularly under the demanding training conditions of an elite athlete. Among endurance athletes, there is a need for out-of-hospital monitoring to enhance detection of arrhythmias under conditions that are relevant and potentially provocative of underlying pathology. The Norwegian patch ECG247 Smart Heart Sensor has been developed to simplify the assessment of heart rhythm disorders. The current study aimed to evaluate the ECG247 Smart Heart Sensor function and usability in an elite athlete environment.

Methods: A total of 13 professional cyclists from the UNO-X Pro Cycling Team were examined with the ECG247 Smart Heart Sensor during training camp in Spain, December 2021. All ECG data were analyzed by cardiologists at Sorlandet Hospital Arendal, Norway. The athletes also completed a brief questionnaire registering their training (from on-bike monitoring units) and provided self-assessment of usability parameters after the test.

Results: In 8 of 13 athletes (69% male, age 23 ± 4 years), two test periods were performed with different ECG patches, resulting in a total of 21 tests with continuous ECG monitoring. Average total ECG test duration per athlete was 144 ± 47 h (89 ± 24 h/patch). Athletes performed an average of 15 ± 5 training h during each test. The ECG quality from all tests was considered satisfactory for rhythm analysis—also during exercise. The reported usability of the ECG247 Smart Heart Sensor was high, and no athletes reported trouble sleeping or training with the sensor. The automatic arrhythmia algorithm reported episodes of possible arrhythmias in 5 (24%) tests; 2 atrial flutter, 2 supraventricular tachycardia and 1 bradycardia (heart rate < 30/min). Manual assessment by physicians verified the episode of bradycardia but revealed normal sinus rhythm in all other tests. No false negative events were identified in over 1,800 h of ECG collection.

Conclusion: The ECG247 Smart Heart Sensor allowed for high quality ECG monitoring with high usability during intensive exercise in athletes.

KEYWORDS

atrial fibrillation, endurance athletes, elite athletes, cardiac screening, cardiac arrhythmia

Introduction

The importance of large volumes of training to perform at a high level in endurance sports is well documented among elite athletes (Seiler, 2010; Tønnessen et al., 2014; Stöggl and Sperlich, 2015). Elite endurance athletes' annual training volume typically ranges from 500 to well above 1,000 h (Billat et al., 2001; Tønnessen et al., 2014; Metcalfe et al., 2017). Endurance exercise is also established as an efficacious method of reducing the risk of developing cardiovascular diseases (CVD). However, there are multiple studies suggesting that "excessive" long-lasting and high-volume endurance training may paradoxically increase the risk of developing certain types of heart disease, particularly arrhythmias (Madias, 2008; Goodman et al., 2015). Atrial Fibrillation (AF) is one of the most common cardiac arrhythmias reported among endurance athletes and AF incidence in athletes has been a theme of considerable research interest (Grimsmo et al., 2011; Andersen et al., 2013; Sanchis-Gomar et al., 2016; Lippi et al., 2021; Newman et al., 2021).

Today's gold-standard for diagnosing cardiac arrhythmias is a 12-lead electrocardiogram (ECG). An ECG test is performed by healthcare personnel in a clinical setting and provides a time-limited snapshot of the heart's electrical function (Quer et al., 2020). Some specific cardiac arrhythmias may be transient, such as AF, and a 12-lead ECG recording period lasting only a few minutes may fail to detect intermittent cardiac arrhythmias. Continuous ECG-recordings are needed to detect and diagnose specific cardiac arrhythmias and the equipment used for long-term ECG recordings is often referred to as "Holter monitoring" (Kułach et al., 2020). A Holter monitor system typically requires a recording device worn on the hip and coupled to at least three cables attached to electrodes on the chest. The system is applied to the patient by specialized healthcare professionals, and is usually worn for ∼24–72 h (Lutfullin et al., 2013). Most Holter systems are not water repellent. Consequently, the Holter monitor system may limit movements and can loosen or detach with physical activity and hard exercise.

For an elite athlete training daily, a Holter monitor prescription will prevent the athlete from training normally, thereby decreasing the validity of the ECG monitoring process. Arrhythmias among elite athletes often occur during exercise (Madias, 2008). A Holter monitor may limit the intensity or continuity of the exercise (Lutfullin et al., 2013). 12-lead ECG and Holter monitoring are dependent on assistance from healthcare personnel, and therefore are subject to limited availability, limited test duration time, and usability challenges. In the context of a high-performance endurance sport team, cardiac screening with today's clinical tools becomes so time consuming that it may be avoided by athletes and coaches despite the appearance of symptoms of concern.

Several new systems purporting to provide long-term ECG monitoring are available. "Smart" watches and training accessories can provide identification of arterial pressure waves. However, international guidelines require ECG documentation for the diagnosis of arrhythmias. Self-applied, single lead ECG patches are currently available on the market for home-based use. However, there is a lack of research evidence regarding their validity and utility in a high-performance endurance athlete population. TheECG247™ Smart Heart Sensor is a new, mobile, long-term patch ECG monitoring device that has undergone extensive testing in a home health care setting (Sandberg et al., 2021; Jortveit and Fensli, 2022; Jortveit et al., 2022) and is approved by European directives for medical devices (93/42/EEC). It provides continuous monitoring of the heart rhythm for up to 7 days and can be used during exercise. The device is small, wireless, and easy to apply and use without any clinical expertise or assistance (Appsens, 2021). All the data acquired by the sensor is uploaded to cloud storage through a smartphone application and can be easily accessed by health care professionals. The user also has access to real-time ECG feedback during testing. The ECG sensor patch is applied over the sternum and remains attached through the whole monitoring period. Monitoring duration is limited by the integrity of the fixation of the sensor patch to the skin over time (up to 7 days). ECG247 has not been systematically tested on athletes. If this ECG patch technology withstands the use characteristics of elite athletes (vigorous movement, sweat, showers, etc.), it can potentially become a viable alternative for screening and cardiac rhythm monitoring in athletes.

The aims of this study were: (1) to evaluate how the ECG247 Smart Heart Sensor technical solution performs in a setting representative of the demands of high-performance endurance athletes during daily training, (2) to investigate the perception of comfort and usability among elite athletes training in demanding field conditions, and (3) to evaluate the ECG quality and automatic arrhythmia detection during high endurance training.

Materials and methods

Study design

This study was designed as a descriptive field test of the ECG247 Smart Heart Sensor (Appsens AS, Lillesand, Norway) on elite endurance athletes from the Uno-X Pro Cycling Team while performing a high volume of endurance training. Methods were designed to accommodate the practical demands of the athletes while assessing both technical and practical aspects of using the ECG device in a sports medicine context. All data collection and testing were performed in December 2021.

Study subjects and procedure

The field test was completed during a 14-day training camp for the Uno-X Pro Cycling Team in Spain, December 2021, and a total of 13 athletes (9 male) were monitored with the ECG247 sensor. These athletes were selected from the entire team (∼50 athletes) by the Uno-X team leadership. They participated in an information meeting and provided signed informed consent prior to the start of ECG data collection. The athletes agreed to wear the sensor for 3–6 days (depending on quality of the ECG recording). The research project leader was present at the training camp during the test period and answered questions from athletes. During the training camp, collected ECG recordings were simultaneously reviewed by physicians at Sorlandet Hospital Arendal in Norway. Cardiological support was provided during the field-testing period to ensure rapid communication with athletes in the event of detected arrhythmias or if false positive events arose. After completion of the field-testing period, the physicians performed a manual review of the complete ECG recordings from every athlete and provided a detailed report for each athlete. Acceptable ECG quality was defined as the ability to determine rhythm (sinus rhythm or specific arrhythmia) based on the physician's assessment. In case of doubt, additional physicians (cardiologists) were consulted.

ECG247 smart heart sensor

The ECG247 is a single-lead patch ECG-monitoring device. The monitoring system consists of a one-time "multi-day use" electrode patch that is attached over the sternum, a re-usable ECG sensor, a smartphone application, a back-end cloud service, and a web portal (Figure 1). The ECG247 sensor continuously monitors the hearth rhythm and automatically detects and categorizes arrhythmias in real-time by using algorithms based on artificial intelligence in the sensor and at the back-end service. The ECG247 Smart Heart Sensor system and the arrhythmia detection algorithms are described previously (Sandberg et al., 2021; Jortveit and Fensli, 2022). The ECG-recordings are transferred *via* Bluetooth to the ECG247 application on the smartphone, and simultaneously uploaded to the back-end cloud service (Figure 2). In cases when the Bluetooth communication between the sensor and the smartphone is interrupted, the ECG247 application will send a notice, with reestablishment of the connection made automatically. In addition, the sensor has an internal memory for ECG storage in case communication error with the phone. The user has ownership and access to the results in the web portal and can provide permissions for sharing of ECG data with their physician or other healthcare professionals. User authentication is provided using the Firebase Service (Google, Mountain View, CA, USA), which generates a two-factor authentication required for access to sensitive health information. All information stored in the web portal is coded as Fast Health Interoperability Resources (FHIR).

All detected arrhythmias are uploaded and saved in the back-end cloud service and sorted by severity in the web portal. The user can also manually highlight up to 1 min of ECG recording by activating this function on the sensor. This allows the user to "tag" ECG measurements when they subjectively experience what they perceive to be a disturbance in heart rhythm.

Laboratory ECG smart heart sensor pilot test

Prior to the field testing of ECG247, preliminary pilot testing in the laboratory was conducted on 6 (4 male, 2 female) physically active sport science students. The primary purpose of the pilot test was to investigate how different movements (cycling, double-poling, running) affected the ECG recordings, as well as evaluate the tolerance of the single-use electrode for repeated bouts of exercise and showering. The positive results of this preliminary test were also deemed a necessary pre-condition for further testing with UNO-X Pro Cycling Team. The test protocol in the laboratory consisted of 15 min efforts on each exercise modality. These efforts were divided into 5 min segments with small successive increases in work intensity. A 5 min rest period was provided between each 15 min exercise bout (Figure 3). Double-poling (Figure 4A) was performed on a Concept2 Skierg (Concept2, Morrisville, VT, USA), cycling (Figure 4B) on a Wattbike AtomX (Wattbike, Nottingham, England), and running (Figure 4C) on a motorized treadmill (Lode Katana Sport, Lode B. V., Groningen, Netherlands).

Ethical considerations

The study was carried out according to the Declaration of Helsinki and data collection methods were approved from a data security perspective by the Norwegian Center for Research Data and was approved by the Ethics Committee of the Faculty for Health and Sport Science, University of Agder.

Athlete participants were not randomly selected by Uno-X team leadership. Athletes with history of reporting possible arrhythmic symptoms were selected to be among the test participants to participate in the test. Consequently, a cardiologist was brought in early to provide additional information to the athletes. In this process, the cardiologist informed the participants that the current algorithms of the ECG247 were not specifically designed for athletes exercising at high heart rates. This increased the likelihood of false positive detection of supraventricular tachycardia (SVT) and Atrial Flutter (AFLU) when heart rate (HR) was normally elevated during training sessions. Therefore, false positive events related to these tachycardias were anticipated and discussed with the athletes.

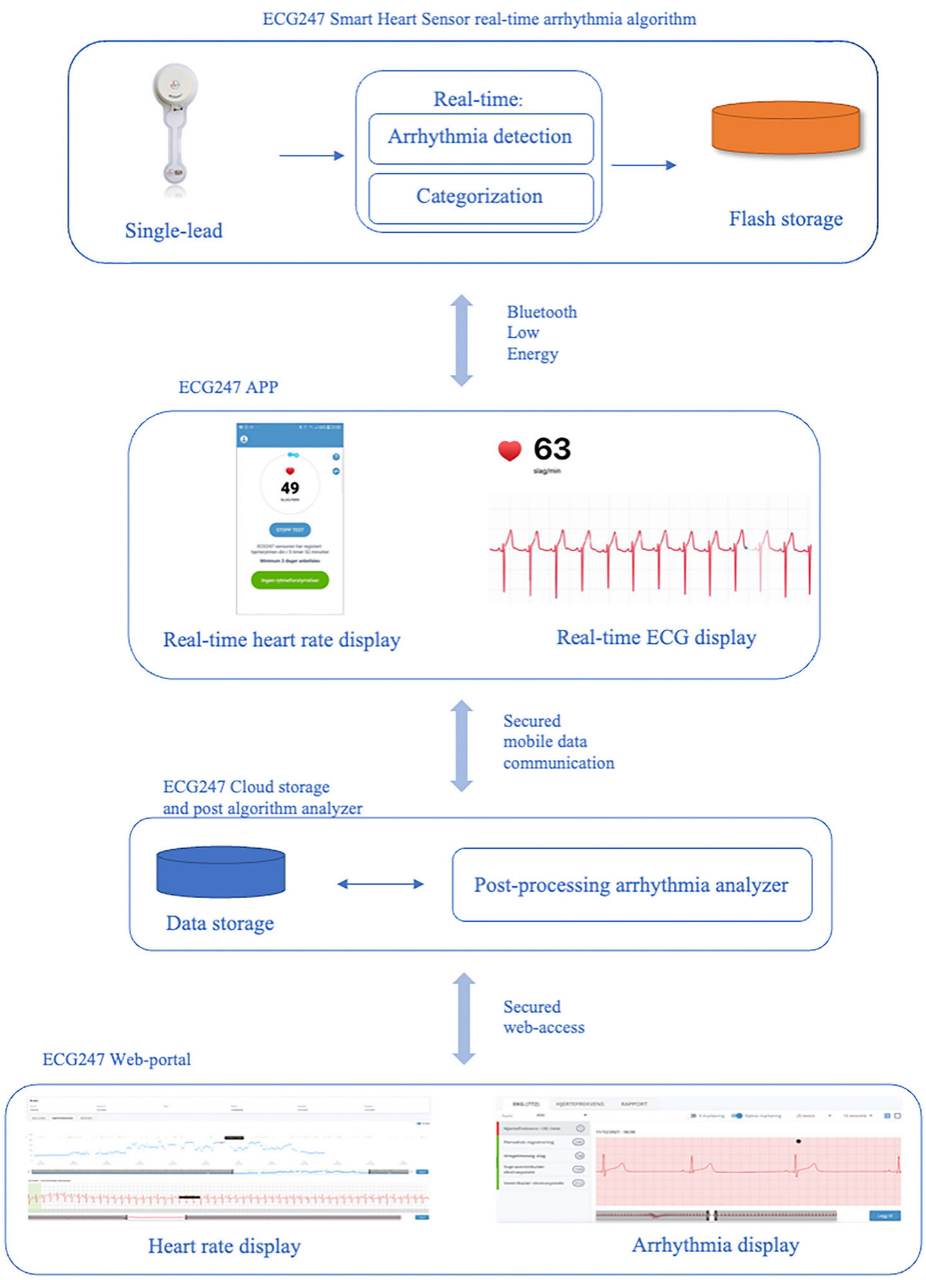

FIGURE 1
The ECG Smart Heart Sensor system: sensor with real-time arrhythmia detection, smartphone application, back-end cloud service with post-processing arrhythmia analyzer, and web portal.

Results

Laboratory ECG247 smart heart sensor pilot test

A total of 6 (4 male) subjects completed the preliminary pilot testing in the laboratory. Figure 5 demonstrates the ECG recordings of the different modalities of one of the subjects. Running (Figure 5A) showed more disturbance in the ECG recordings among all subjects compared with cycling (Figure 5B) and XC-ski double-poling (Figure 5C). All ECG recordings were evaluated by a cooperating physician within 48 h of the test period and considered satisfactory for rhythm analysis in all the tests.

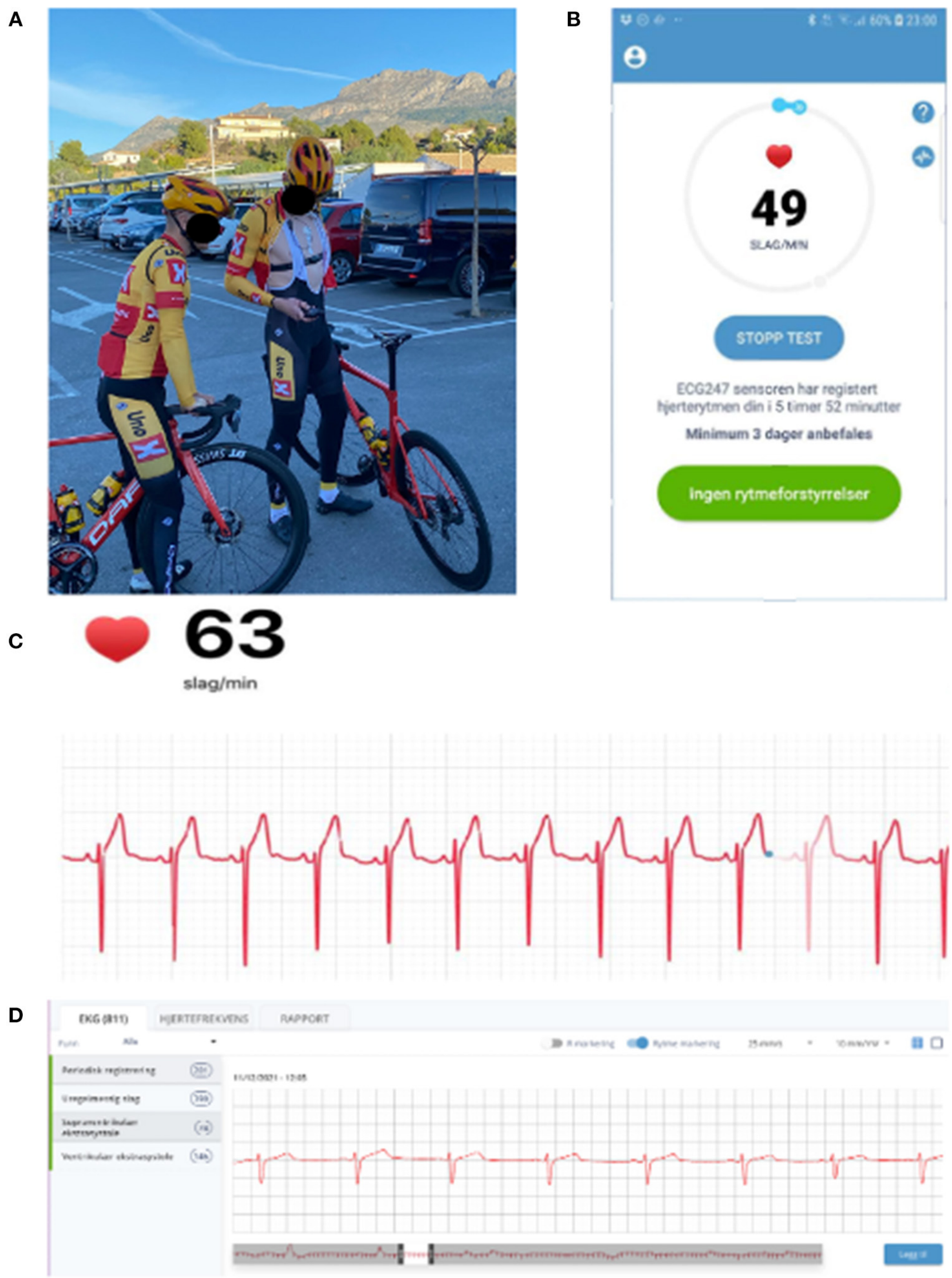

FIGURE 2
(A) The ECG247 sensor placed over the sternum, screenshots from **(B,C)** the ECG247 mobile application and **(D)** the web portal.

Field test of the ECG247 smart heart sensor

Continuous ECG monitoring was successfully performed on a total of 13 athletes. The average age of the participants was 23 ± 4 years (69% males). In 8 of 13 athletes, 2 test periods were performed, resulting in 21 continuous ECG monitoring periods of at least 43 h. New tests, with new single use electrode patches were started due to partial detachment of the electrode from the skin ($n = 1$), ECG signal degradation was identified remotely by the physician ($n = 4$), and by request from athletes ($n = 3$). The mean athlete test duration time was 144 ± 47 h, with an average functional duration of 89 ± 24 h for each ECG patch/test period. During the test

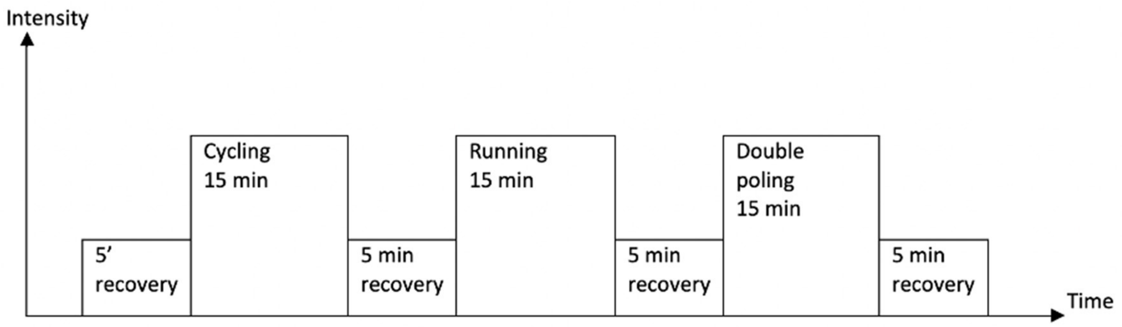

FIGURE 3
Test protocol for the pilot test of the ECG247. Started with applying the sensor and connect to the participants phones. Recovery consisted of walking and sitting. The intensity increased slightly every 5 min during the efforts.

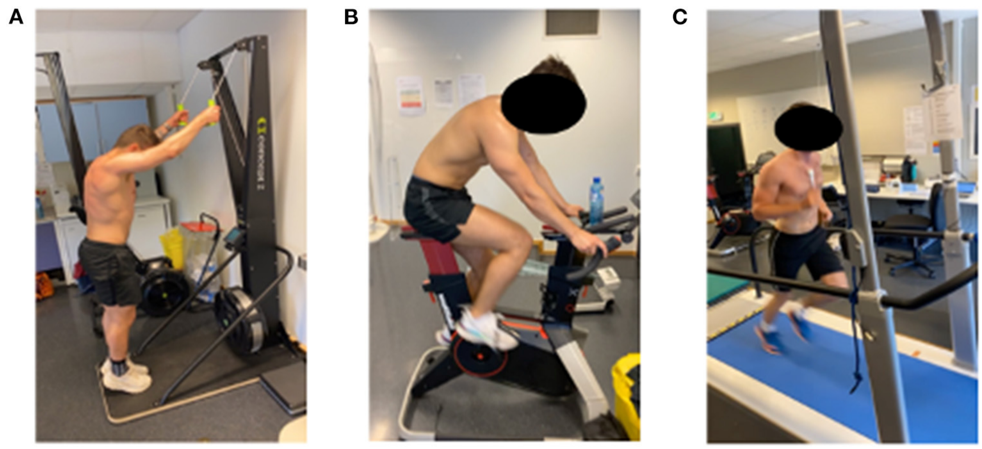

FIGURE 4
Laboratory exercise modalities evaluated during preliminary testing of ECG24 sensor: **(A)** Ski double-poling, **(B)** Cycling, **(C)** Running.

period, an average of 24 ± 6 h of training was performed by each athlete, with an averaging 15 ± 5 training h for each electrode patch.

Self-reported usability of ECG247 smart heart sensor

Four participants reported some discomfort (itching) underneath the sensor patch on the chest. Three of these four athletes reported that the itching stopped after the first 24 h of the test. Nine athletes reported forgetting that they were wearing the sensor from time to time. No athletes reported trouble sleeping or training with the ECG247 sensor. None of the 13 tested athletes reported problems with the connection or Smart phone application. However, 7 athletes reported concerns and questions around results during the test period (Table 1).

ECG quality and automatic arrhythmia detection

The ECG quality from all tests was considered satisfactory by the physicians for rhythm analysis—also during exercise. One short nocturnal episode of bradycardia (heart rate <30/min) was detected by the ECG247 automatic algorithms and verified by the physician. Two short episodes of SVT and 2 short episodes of AFLU in four different athletes were marked by the ECG247 system, but all of these were refuted by the manual assessment of the physicians. User-initiated recordings were performed five times without any pathological ECG findings (Table 2).

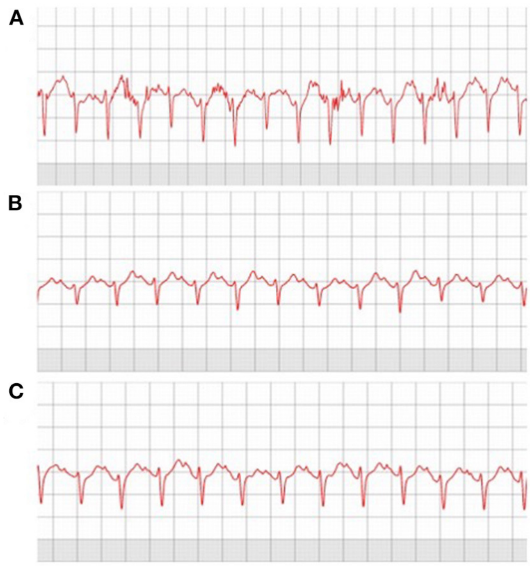

FIGURE 5
Laboratory exercise modalities evaluated during preliminary testing of ECG247 sensor: **(A)** Running, **(B)** Cycling, **(C)** Double poling XC.

TABLE 1 Usability of ECG247 smart heart sensor.

	All ($n = 13$)
Itching	4
No reported discomforts	9
Disturbed sleep	0
Disturbed training	0
Disturbed phone connection	0
Concerns during the test	7

Values are presented as prevalence.
n, number of participants.

TABLE 2 Characteristics and diagnostic evaluation for the field tests.

	Athletes ($n = 13$)	Tests ($n = 21$)
Age (y)	23 ± 4	
Test duration (hours)	144 ± 47	89 ± 24
Training volume (hours)	24 ± 6	15 ± 5
Showers (times)	6 ± 1	4 ± 1
Recording periods < 72 h	3	3
ECG247 algorithm detection		
AF and severe arrhythmia	0	0
Bradycardia	1	1
False positive SVT	2	2
False positive AF	2	2
False negative	0	0
Patient-initiated recordings		
Recordings	3	5
Physician review detection of arrythmia	1	1

Values are presented as mean ± standard deviation and prevalence.
SVT, supraventricular tachycardia; AFlu, atrial flutter; n, number of participants.

Discussion

This study of the ECG247 Smart Heart Sensor technical performance in 13 endurance athletes from the Uno-X Pro Cycling Team during extensive training verified technical quality, usability, and ECG quality satisfactory for heart rhythm assessment, also during exercise.

The findings of the present study suggest that the ECG247 Smart Heart Sensor provides an easy and technical acceptable method of monitoring cardiac health in athletes with minor negative side effects or annoyances. The system overcomes the limitations of conventional diagnostic tools for assessment of rhythm disorders like limited availability, limited test duration time, and usability challenges, particularly under the demanding training conditions of an elite athlete.

The reported usability of the ECG247 Smart Heart Sensor was high, and no athletes reported trouble sleeping or training while wearing the sensor. The project leader present at the training camp during the field testing received athlete concerns during the test period. These concerns arose mainly from reports from the application saying that there was a possible arrhythmia. Most of these events were determined to be false positive. The patch sensor showed promising usability also in a team training camp context. The sensor enables transition of the assessment of arrhythmias from the hospitals to the athlete's training and competition environment. Professional sports teams are often composed of multinational athletes, with different healthcare service providers. An out-of-hospital, reusable cardiac rhythm device could make assessment of heart rhythm disorders and heart symptoms cheaper and less time-consuming compared with conventional hospital methods. In addition, ECG247 Smart Heart Sensor did not limit exercise in any way, which is a crucial detail when monitoring elite athletes.

A purpose of the pilot test was to investigate how different movements (cycling, double-poling, running) affected the ECG recordings, as well as initially evaluate the tolerance of the single-use electrode for repeated bouts of exercise and showering. The pilot testing completed as a prelude to the present study illustrates that there are some differences in ECG quality across exercise modalities. There was one incident of a false positive test (AFLU) during preliminary lab testing, which provided perspectives about the need to ensure the safety and psychological wellbeing of the athletes during the field trial. A cardiologist was brought in to analyze potential arrhythmias simultaneously during the test period. The quality of the

ECG recordings was considered satisfactory for hearth rhythm assessment in the pilot test. However, more work is needed on the different exercise modalities and their potential influence on the quality of the ECG recordings.

The findings from the present field testing will inform algorithm adaptation for sport medicine applications. This athlete population represented a severe test of the technical solution given the high training volumes performed. The capacity of the solution to deliver continuous, interpretable ECG recordings for at least 48 h was deemed as a cutoff for minimum viability in a sports medicine context. The arrythmia detection algorithms employed were originally based on a sedentary, primarily elderly population. SVTs and AFLU may be near identically to the ECG of an athlete exercising with abrupt changes in HR.

Therefore, the physician on the research team anticipated a risk of false positive findings associated with the high heart rates achieved during normal training in this elite athlete group. Prior to the field test, athlete volunteers were informed that the integrated arrhythmia analyzing algorithm was sensitive to abrupt HR elevation and might falsely detect events of AFLU and SVTs However, the proportion of false positive arrhythmia events detected by the automatic algorithm was relatively low (19%). Importantly, no actual ECG arrythmias went undetected (false negative) by the algorithmic solution in over 1,800 h of ECG monitoring.

Strength and limitations

The main strengths of the present study were: (a) cooperation with a professional cycling team, which provided an excellent field-testing environment, and (b) extensive preliminary pilot testing. A training camp, with a professional cycling team was an appropriate environment for testing whether ECG247 Smart Heart Sensor withstands the typical patterns of athlete training several hours daily, showering, etc. In addition, testing the sensor during a training camp was a good simulation for investigating how it works in a team context. On-time access to a cardiologist was crucial for this initial study because it provided both reassurance for the athletes and ensured an optimal analysis process. However, this was not an interventional study, and there was no comparison with today's best practice (Holter monitoring). Cycling is also one of the endurance sports with the least amount of movement in the upper body. Therefore, the present findings should not be generalized to all sports movements.

A single-lead ECG may be more difficult to interpret by a physician compared to a 3-lead ECG from a Holter system. However, the number of leads is less important for the interpretation of narrow QRS complex arrhythmias like AF and SVT. The position of the single-lead ECG patch is essential for high signal quality on the ECG recordings during physical exercise. The ECG247 Smart Heart Sensor is placed directly over the sternum. This is an anatomical placement with little multi-directional skin stretch and muscular movement under the electrode and therefore presumably less electrical signal disturbance compared with placing the sensor left on the chest, over skin and muscle that introduces significant resonant movement artifact during exercise. For sport-use, sternal ECG electrode placement of single-lead patch electrodes may be optimal.

Practical applications

The importance of large volumes of training to perform at a high level in endurance sports is well documented among elite athletes (Seiler, 2010; Tønnessen et al., 2014; Stöggl and Sperlich, 2015). These findings highlight the need for cardiac screening methods which are easily accessible and do not interfere with the everyday training of an elite athlete. The field test of ECG247 Smart Heart Sensor illustrates how assessment of possible heart rhythm disorders can be performed in an elite team environment, without any interference of training and sleeping rhythms.

As mentioned, cycling is one of the endurance sports with the least amount of movement in the upper body. Therefore, additional field testing of this device in other athlete groups, such as runners, is warranted.

Conclusion

The study demonstrates that the ECG247 Smart Heart Sensor allowed high quality ECG monitoring with high usability during intensive exercise in athletes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Faculty for Health and Sport Science, University of Agder. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ÅA and SS was responsible for the data collection. ES was responsible for manual assessment of the ECG data. ÅA drafted and SS, ES, and JJ critically revised the manuscript. All authors gave final approval and agreed to be accountable for all aspects of this work.

Acknowledgments

The present study was conducted in collaboration with the University of Agder, Uno-X Pro Cycling Team, and Sorlandet Hospital, Arendal, Norway.

Conflict of interest

ES has received speaking fees from Pfizer. JJ has received speaking fees from Amgen, AstraZeneca, BMS, Boehringer Ingelheim, Novartis, Pfizer, and Sanofi. He is a shareholder in AppSens AS and is employed in the company.

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Andersen, K., Farahmand, B., Ahlbom, A., Held, C., Ljunghall, S., Michaelsson, K., et al. (2013). Risk of arrhythmias in 52 755 long-distance cross-country skiers: a cohort study. *Eur. Heart J.* 34, 3624. doi: 10.1093/eurheartj/eht188

Appsens (2021). *ECG247 Smart Sensor System: User Guide.* https://ecg247.com2021; Available online at: https://ecg247.com/media/k2gld5fc/ecg247_user_guide_p2_26_eng.pdf (accessed December 01, 2021).

Billat, V. L., Demarle, A., Slawinski, J., Paiva, M., and Koralsztein, J. P. (2001). Physical and training characteristics of top-class marathon runners. *Med. Sci. Sports Exerc.* 33, 2089–2097. doi: 10.1097/00005768-200112000-00018

Goodman, J. M. P., Burr, J. F. P., Banks, L. P., and Thomas, S. G. P. (2015). The Acute Risks of Exercise in Apparently Healthy Adults and Relevance for Prevention of Cardiovascular Events. *Can J Cardiol.* 32, 523–532. doi: 10.1016/j.cjca.2016.01.019

Grimsmo, J., Maehlum, S., Moelstad, P., and Arnesen, H. (2011). Mortality and cardiovascular morbidity among long-term endurance male cross country skiers followed for 28-30 years. *Scand J Med Sci Sports.* 21, e351–e8. doi: 10.1111/j.1600-0838.2011.01307.x

Jortveit, J., and Fensli, R. (2022). Remote ECG Monitoring by ECG247 Smart Heart Sensor. *International Journal of Telemedicine and Applications.* 2022, 6812889. doi: 10.1155/2022/6812889

Jortveit, J., Früh, A., and Odland, H. H. (2022). Paroxysmal tachycardia diagnosed by ECG247 smart heart sensor in a previously healthy child. *Case Reports in Pediatrics.* 2022, 9027255. doi: 10.1155/2022/9027255

Kułach, A., Dewerenda, M., Majewski, M., Lasek-Bal, A., and Gasior, Z. (2020). 72 hour Holter monitoring, 7 day Holter monitoring, and 30 day intermittent patient-activated heart rhythm recording in detecting arrhythmias in cryptogenic stroke patients free from arrhythmia in a screening 24 h Holter. *Open Med (Wars).* 15, 697–701. doi: 10.1515/med-2020-0203

Lippi, G., Sanchis-Gomar, F., and Cervellin, G. (2021). Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int J Stroke.* 16, 217–221. doi: 10.1177/1747493019897870

Lutfullin, I. Y., Kim, Z. F., Bilalova, R. R., Tsibulkin, N. A., Almetova, R. R., Mudarisova, R. R., et al. (2013). A 24-hour ambulatory ECG monitoring in assessment of QT interval duration and dispersion in rowers with physiological myocardial hypertrophy. *Biol Sport.* 30, 237–241. doi: 10.5604/20831862.1077547

Madias, J. E. (2008). Athletes, Ventricular Arrhythmias, Electrophysiological Testing, Microvolt T-Wave Alternans, and a Follow-Up of 30 21 Months: A Need for Follow-Up Updates. *Ann Noninvasive Electrocardiol.* 13, 319–320. doi: 10.1111/j.1542-474X.2008.00239.x

Metcalfe, A. J., Menasp,à, P., Villerius, V., Quod, M., Peiffer, J. J., Govus, A. D., et al. (2017). Within-season distribution of external training and racing workload in professional male road cyclists. *Int J Sports Physiol Perform.* 12, 142–146. doi: 10.1123/ijspp.2016-0396

Newman, W., Parry-Williams, G., Wiles, J., Edwards, J., Hulbert, S., Kipourou, K., et al. (2021). Risk of atrial fibrillation in athletes: a systematic review and meta-analysis. *British journal of sports medicine.* 55, 1233–1238. doi: 10.1136/bjsports-2021-103994

Quer, G., Freedman, B., Steinhubl SR (2020). Screening for atrial fibrillation: predicted sensitivity of short, intermittent electrocardiogram recordings in an asymptomatic at-risk population. *Europace.* 22, 1781–1787. doi: 10.1093/europace/euaa186

Sanchis-Gomar, F., Perez-Quilis, C., Lippi, G., Cervellin, G., Leischik, R., Löllgen, H., et al. (2016). Atrial fibrillation in highly trained endurance athletes—description of a syndrome. *Int. J. Cardiol.* 226, 11–20. doi: 10.1016/j.ijcard.2016.10.047

Sandberg, E. L., Grenne, B. L., Berge, T., Grimsmo, J., Atar, D., Halvorsen, S., et al. (2021). Diagnostic accuracy and usability of the ECG247 smart heart sensor compared to conventional holter technology. *J. Healthc Eng.* 2021, 5230947–5230948. doi: 10.1155/2021/5230947

Seiler, S. (2010). What is best practice for training intensity and duration distribution in endurance athletes? *Int. J. Sports Physiol. Perform.* 5, 276–291. doi: 10.1123/ijspp.5.3.276

Stöggl, T. L., and Sperlich, B. (2015). The training intensity distribution among well-trained and elite endurance athletes. *Front. Physiol.* 6, 295. doi: 10.3389/fphys.2015.00295

Tønnessen, E., Sylta, Ø., Haugen, T. A., Hem, E., Svendsen, I. S., Seiler, S., et al. (2014). The road to gold: training and peaking characteristics in the year prior to a gold medal endurance performance. *PLoS ONE.* 9, e101796. doi: 10.1371/journal.pone.0101796

frontiers | Frontiers in Sports and Active Living

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fspor.2022.939980

Check for updates

OPEN ACCESS

EDITED BY
Laura Gastaldi,
Politecnico di Torino, Italy

REVIEWED BY
Hiroshi Takemura,
Tokyo University of Science, Japan
Maria António Castro,
Escola Superior de Saúde-Politécnico de Leiria ESSLEI, Portugal

*CORRESPONDENCE
Thomas Robert
thomas.robert@univ-eiffel.fr

SPECIALTY SECTION
This article was submitted to Sports Science, Technology and Engineering,
a section of the journal
Frontiers in Sports and Active Living

RECEIVED 09 May 2022
ACCEPTED 30 June 2022
PUBLISHED 25 July 2022

CITATION
Lahkar BK, Muller A, Dumas R, Reveret L and Robert T (2022) Accuracy of a markerless motion capture system in estimating upper extremity kinematics during boxing. *Front. Sports Act. Living* 4:939980. doi: 10.3389/fspor.2022.939980

Accuracy of a markerless motion capture system in estimating upper extremity kinematics during boxing

Bhrigu K. Lahkar[1], Antoine Muller[1], Raphaël Dumas[1], Lionel Reveret[2] and Thomas Robert[1]*

[1]Univ Lyon, Univ Gustave Eiffel, Univ Claude Bernard Lyon 1, LBMC UMR_T9406, Lyon, France,
[2]INRIA Grenoble Rhone-Alpes, LJK, UMR 5224, Grenoble, France

Kinematic analysis of the upper extremity can be useful to assess the performance and skill levels of athletes during combat sports such as boxing. Although marker-based approach is widely used to obtain kinematic data, it is not suitable for "in the field" activities, i.e., when performed outside the laboratory environment. Markerless video-based systems along with deep learning-based pose estimation algorithms show great potential for estimating skeletal kinematics. However, applicability of these systems in assessing upper-limb kinematics remains unexplored in highly dynamic activities. This study aimed to assess kinematics of the upper limb estimated with a markerless motion capture system (2D video cameras along with commercially available pose estimation software Theia3D) compared to those measured with marker-based system during "in the field" boxing. A total of three elite boxers equipped with retroreflective markers were instructed to perform specific sequences of shadow boxing trials. Their movements were simultaneously recorded with 12 optoelectronic and 10 video cameras, providing synchronized data to be processed further for comparison. Comparative assessment showed higher differences in 3D joint center positions at the elbow (more than 3 cm) compared to the shoulder and wrist (<2.5 cm). In the case of joint angles, relatively weaker agreement was observed along internal/external rotation. The shoulder joint revealed better performance across all the joints. Segment velocities displayed good-to-excellent agreement across all the segments. Overall, segment velocities exhibited better performance compared to joint angles. The findings indicate that, given the practicality of markerless motion capture system, it can be a promising alternative to analyze sports-performance.

KEYWORDS

markerless vs. marker-based, kinematic analysis, evaluation, elite sport, upper-limb, sports-performance

Introduction

Boxing is an intensive combat sport, involving highly dynamic and non-symmetrical movements of the front and rear arms with the role of attack or defense as situation demands. In such sports, high-performance athletes are often characterized by their agility, i.e., the ability to punch or evade swiftly by maintaining fluidity of motion. To achieve a powerful punch during offensive action and quick retraction during defense, coordination of the body segments plays a vital role (Dinu and Louis, 2020). As body segments' coordination is often a consequence of how the adjacent segments are oriented to each other (Zajac and Winters, 1990; Putnam, 1993), estimating segment pose (positions and orientations) during boxing may be helpful to analyze the performance athletes. Furthermore, the velocities at which body segments move and coordinate with each other have been reported to vary across athletes based on their skills (Putnam, 1993). Therefore, estimating velocities of the body segments seems essential to analyze sports-performance during boxing.

To quantify body segment kinematic variables, marker-based motion capture has been most widely used. In such systems, skin markers are placed on the specific anatomical landmarks, based on which body segment coordinate systems are defined to estimate 3D pose of the segments. While marker-based methods are traditionally referred to as standard, they are commonly performed in a laboratory environment and require adequate skills in physical palpation of landmarks. Even with necessary skills, such palpation is examiner-dependent and at times tends to produce systemic bias for an examiner (Johnson et al., 2018). Furthermore, joint kinematics are also largely affected by soft tissue artifact (Camomilla et al., 2017; Lahkar et al., 2021). Alternatively, measurements based on wearable sensors such as inertial measurement units have been recently shown effective in natural environment in estimating joint angles of the lower extremity with moderate to strong accuracy (Al Borno et al., 2022). Studies also presented the use of inertial measurement units in estimating hand velocity (Kimm and Thiel, 2015; Punchihewa et al., 2020) and other body segments (Dinu and Louis, 2020) during a sport activity. While such studies are useful for understanding differences in skills between athlete groups, placing sensors or markers on the body surface may be inconvenient and potentially distracting to an athlete and practically impossible during a live combat.

With the rapid advancement of computer vision research, human movement study has received a significant stride allowing unobtrusive capture of data using video-based markerless motion capture (Colyer et al., 2018; Armitano-Lago et al., 2022). These methods rely on 2D video data combining with generative or discriminative algorithms to estimate human pose in 3D (Colyer et al., 2018). Generative approach often involves fitting a predefined model of the subject to 2D visual cues such as image features from detectors or to 3D cues such as a visual hull reconstruction with the help of silhouette matching algorithms (Corazza et al., 2006, 2007; El-Sallam et al., 2013). On the other hand, learning-based discriminative algorithms, particularly deep neural network, involves detecting sparse set of learned features such as joint key points describing a subject's pose in 2D. In this family, openly accessible pose estimator like OpenPose (Cao et al., 2021) has received significant attention in human movement analysis and similarly DeepLabCut (Mathis et al., 2018) for both human and non-human activities. As these tools are primarily intended for 2D pose estimation, some studies leveraged its potential in estimating 2D kinematics of the lower limb (hip, knee, and ankle joint) during gait (Stenum et al., 2021), vertical jump (Drazan et al., 2021), and under water running (Cronin et al., 2019). Progressing further, others focused on estimating 3D poses from 2D images of multiple calibrated cameras using triangulation during walking (Nakano et al., 2020; Needham et al., 2021; Pagnon et al., 2022), jumping (Nakano et al., 2020; Needham et al., 2021), running (Needham et al., 2021; Pagnon et al., 2022), cycling (Pagnon et al., 2022), and throwing (Nakano et al., 2020). While these studies demonstrated the potential of openly accessible pose estimation tools in estimating 3D joint kinematics, most of them primarily evaluated the lower extremity. As far as we are aware of Nakano et al. (2020) estimated 3D joint positions of the shoulder, elbow, and wrist and evaluated against traditional marker-based approach during walking, jumping, and ball throwing activity. A mean absolute error up to 4 cm was observed at the wrist, 4.7 cm at the elbow, and 2.2 cm at the shoulder during throwing activity.

In a recent development in markerless video-based systems, Theia3D (Theia Markerless, Inc., Kingston, Ontario) has emerged as a rapidly evolving commercial pose estimation software. The software implements deep convolutional neural network combining with standard biomechanical pose estimation approaches (inverse kinematics) to estimate 3D pose of human body segments. Using this tool, studies showed decent kinematic accuracies compared to marker-based method while maintaining good repeatability both in laboratory environment (Kanko et al., 2021a,b) and in community settings (Mcguirk et al., 2022; Riazati et al., 2022). These studies, however, mainly provide the assessment of the lower

TABLE 1 Demographic details of the athletes.

Athlete	Gender	Age (years)	Height (m)	Body mass (kg)
1	Male	20	1.72	54
2	Male	18	1.90	78
3	Female	19	1.63	59

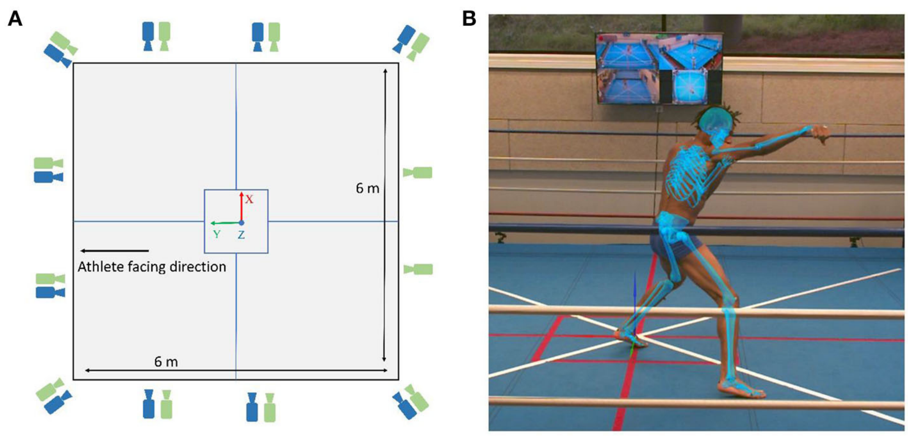

FIGURE 1
(A) Layout of the boxing ring with green and blue cameras depicting optoelectronic and video cameras, respectively. **(B)** An example of the boxing punch with Theia3D multibody model overlaid on the 2D video image.

TABLE 2 Boxing trials and their specific characteristics.

Trial	Characteristics
1	Direct from the front arm to the face
2	Direct from the front arm to the body
3	Double of the front arm to the face
4	Rear arm jab + front arm hook
5	Front arm (uppercut to the body + hook to the body + hook to the face)

extremity kinematics during either treadmill or over ground walking activity.

While it is relevant to evaluate the usability of markerless systems in a highly dynamic and non-symmetrical sport such as boxing, it still remains unknown how accurate these systems are in estimating upper-limb kinematics as compared to marker-based approach. This study aimed to assess whether a markerless approach (use of video cameras + commercial pose estimation software Theia3D) can be used to estimate upper-limb kinematics as an alternative to the state-of-the-art marker-based approach for sports-performance analysis during "in the field" boxing.

Materials and methods

Participants

A total of three elite boxers volunteered in the study at the boxing arena of National Institute of Sport, Expertise, and Performance (INSEP, Paris, France). Out of the three boxers, one is competing at the national level and two others at the international level. All of them are undergoing regular training at INSEP for Paris Olympics, 2024. Demographic details of the athletes are presented in Table 1. The athletes, after being fully informed about the objectives and protocol of the study, signed an informed consent form. The study and the procedures were approved by an institutional review board.

Data acquisition setup and protocol

Boxing data were collected synchronously using an optoelectronic marker-based system (12 Qualisys Miqus and Arqus cameras; 2–5 megapixel) at 300 Hz, and using a markerless 2D video-based system (10 Qualisys Miqus video cameras; 2 megapixel) at 60 Hz. Both the types of cameras, optoelectronic and video, were placed next to each other as a pair around the boxing ring, except two optoelectronic cameras placed separately to the posterior aspect of the athlete (Figure 1A). All the cameras were connected to Qualisys Track Manager for allowing them to be synchronized and calibrated in space and time, giving a single global reference frame nearly at the center of the boxing platform. Camera setup and placement was performed by the team members with expertise in both optoelectronic and video-based motion analysis. Specific attention was provided to the 2D video-based cameras to comply with recommended specifications for resolution, focus, and exposure time.

Prior to the sessions, the boxers were outfitted with 44 retro-reflective skin markers placed by a single operator with adequate palpation skills on the relevant landmarks of the whole body (Wu et al., 2002, 2005). The details

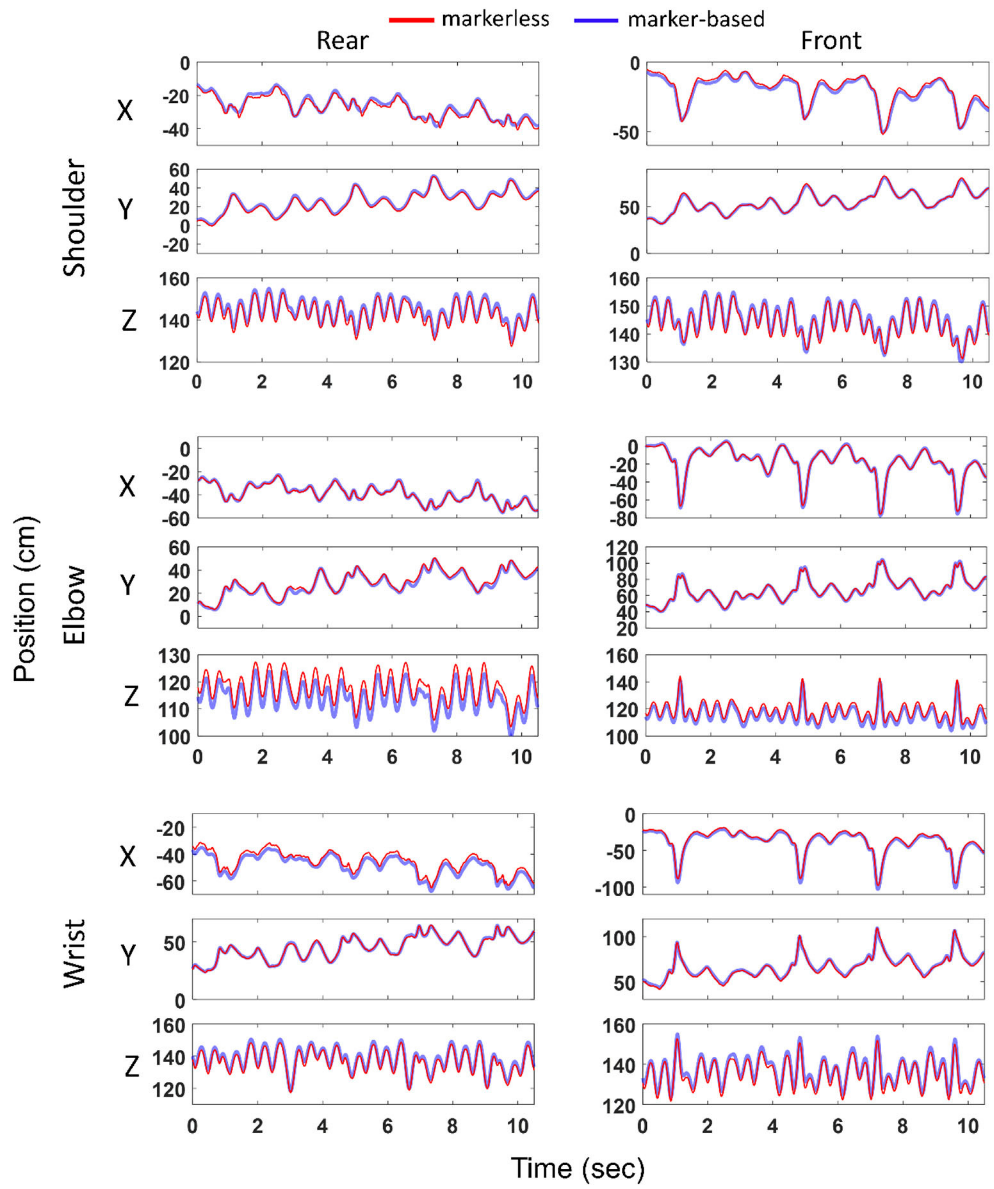

FIGURE 2
Joint center positions (ordinate) at the shoulder, elbow and wrist in the global reference frame (*X, Y, Z*) computed with marker-based and markerless methods and represented over time (abscissa). Left and right columns represent joints of the rear and front limbs, respectively. Example shown for the second athlete and first boxing trial. Blue and red colors represent marker-based and markerless joint center positions, respectively.

of the marker-set and their anatomical locations are provided in the Supplementary Material. A professional coach instructed each boxer to perform specific five shadow boxing trials of different characteristics, with 4–5 repetitions in each trial (Table 2). In between repetitions within a trial, boxers were instructed to perform footwork and remain in defensive pose with elbow flexed guarding their body and face, as classically performed during a contest.

Data processing and analysis

Multibody models

Theia3D embedded multibody kinematic model consists of two separate kinematic chains: one for the lower extremity and one for the upper extremity, and a separate head segment with six degrees of freedom (DoFs) (https://www.theiamarkerless.ca/docs/model.html). In this study, we will only adhere to the upper extremity model in the following descriptions. The upper

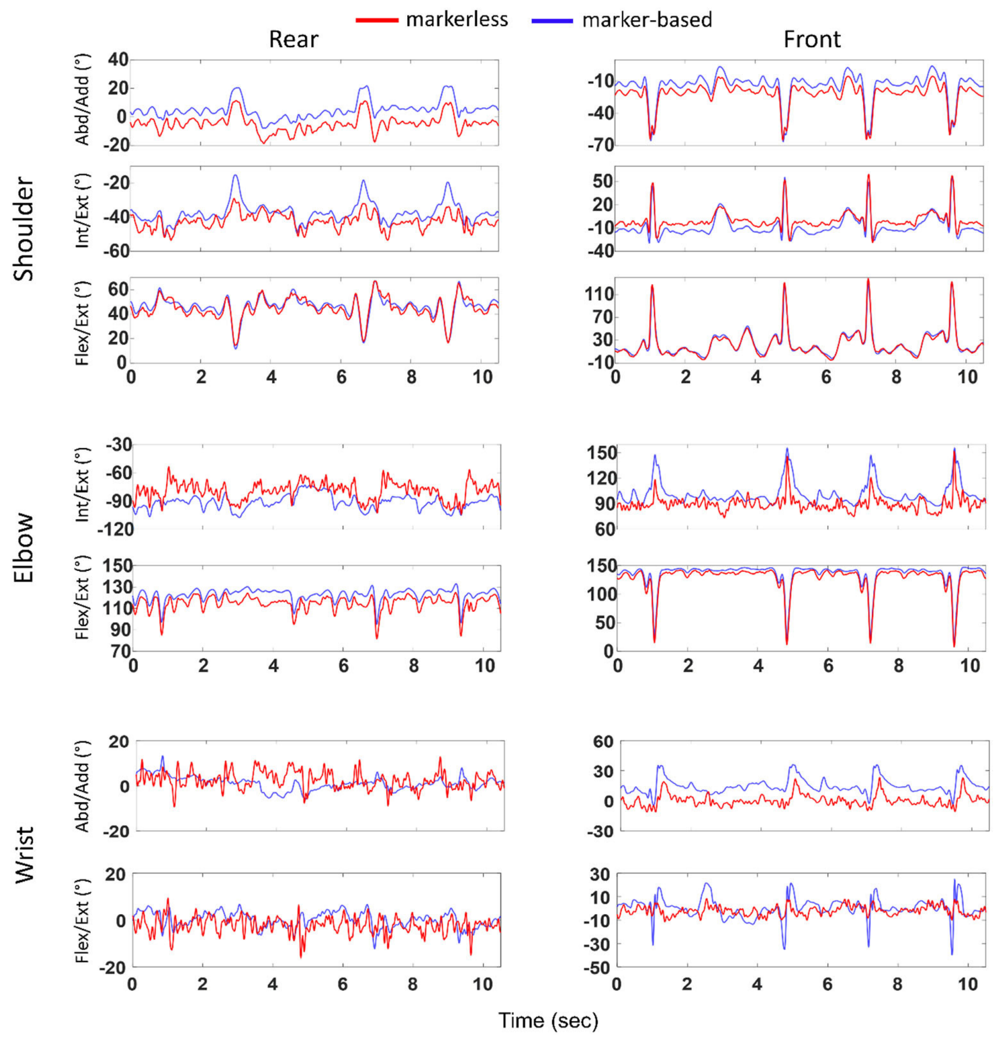

FIGURE 3
Joint angles (ordinate) at the shoulder, elbow, and wrist computed with marker-based and markerless methods and represented over time (abscissa). Left and right columns represent joints of the rear and front limbs, respectively. Example shown for the second athlete and first boxing trial. Blue and red colors represent marker-based and markerless joint angles, respectively.

extremity chain comprises the thorax as root segment with six DoFs with respect to the ground, followed by the clavicle, upper arm, forearm, and hand segments bilaterally. The clavicle, at its proximal end, is connected to the thorax with a two rotational DoFs constraint, while distally connected to the upper arm with a three rotational DoFs at the shoulder joint. The elbow and the wrist joints are constrained to have two DoFs, restricting abduction/adduction (Abd/Add) at the elbow and internal/external (Int/Ext) rotation at the wrist.

For the marker-based multibody model, a gender-specific generic template was created in Visual3D (C-motion, Germantown, USA, v2021.11.3) to have identical body segments and joint constraints as that of the Theia3D model. The shoulder, cervical, lumbar, and thoracic joint centers were defined based on the regression equations adopted from the study of Dumas and Wojtusch (2018). The midpoint between the medial and lateral humeral epicondyles was defined as the elbow joint center and the midpoint between the ulnar and radius styloid processes as the wrist joint center. Segment reference frames were defined following the methodology reported in the study of Dumas and Wojtusch (2018).

For both the models, the center of mass position for each segment was defined according to the study of Dumas and Wojtusch (2018).

Kinematic estimation

Markerless motion capture data were processed with Theia3D (v2021.2), a deep learning-based software. The underlying principle of the software is detailed elsewhere

TABLE 3 Bland–Altman bias (*b*), confidence interval (*CI*) along with coefficient of determination (R^2) and root mean square difference (*RMSD*) between markerless and marker-based methods for joint angles at the shoulder, elbow, and wrist.

Joints	Side	*b*	*CI*	R^2	*RMSD*
Abduction/Adduction (°)					
Shoulder	Front	3.7	13	0.90	6.6
	Rear	−0.1	15	0.37	6.3
Wrist	Front	7.2	17	0.31	11
	Rear	0.2	21	0.39	9.1
Internal/External (°)					
Shoulder	Front	−8.7	9.4	0.83	12
	Rear	2.5	19	0.41	8.1
Elbow	Front	13	30	0.17	23
	Rear	−12	18	0.21	18
Flexion/Extension (°)					
Shoulder	Front	2.4	13	0.88	10
	Rear	0.3	8	0.77	7.3
Elbow	Front	−6.2	5.3	0.99	7.4
	Rear	−5.4	8.5	0.87	7
Wrist	Front	7.4	22	0.41	14
	Rear	−0.7	59	0.27	20

Units for all parameters are in degrees except R^2 (no unit).

in the study of Kanko et al. (2021a) and briefly delineated hereafter. The software relies on synchronized and calibrated videos as input that uses pre-trained deep convolutional neural networks to estimate 2D positions of learned key features (e.g., joint locations and surface landmarks) within the frames of video data, thus enabling to obtain the features in 3D space. The embedded multibody kinematic model is adapted to fit 3D subject-specific features, and a multibody kinematic optimization scheme (Begon et al., 2018) allows to perform 3D pose estimation during an activity. In this study, estimated 3D poses (4-by-4 matrices) of the body segments were exported to Visual3D to compute joint kinematics and segment velocities. Figure 1B illustrates an example of the Theia3D model obtained with multibody kinematic optimization during boxing.

Regarding the marker-based data, the generic multibody template was adapted to obtain subject-specific scaled models, and segment's pose estimation throughout all motion frames was obtained using multibody kinematic optimization (Begon et al., 2018) within Visual3D. Proper segment-specific marker weights were implemented and tuned based on residual analysis, with highest weight at the thorax, followed by the upper arm, forearm, and hand.

The markerless vs. marker-based method was assessed by the following kinematic variables: joint center positions, joint angles, and linear segment velocities. The joint center positions at the shoulder, elbow, and wrist were retrieved from the pose matrices resulting from multibody kinematic optimization. Then, 3D Euclidean distances between corresponding joint centers across all the trials and subjects were computed. The joint angles at the shoulder (between thorax and upper arm), elbow (between upper arm and forearm), and wrist (between forearm and hand) were computed with cardan sequences of rotation adopted from the study of Wu et al. (2002). Linear segment velocity magnitudes were derived from the center of mass positions of each segment in the global reference frame. The kinematic variables were exported to MATLAB (MathWorks, USA), and a 4th-order low-pass Butterworth filter was implemented to filter both the joint angles and segment velocities with cutoff frequency of 8 Hz.

Statistical analysis

The deviation between corresponding joint centers estimated with marker-based and markerless system was assessed as mean (standard deviation) or median (interquartile range) based on normality outcomes across all the trials and subjects. The degree of agreement between joint angles resulting from both the methods was assessed using Bland–Altman analysis (Bland and Altman, 1986). Bias (*b*), confidence interval (*CI*; 1.96 times standard deviation or 1.45 times interquartile range for non-normal distributions), coefficient of determination (R^2), and root mean square difference (*RMSD*) were calculated for comparison. The same statistical parameters were used for comparing segment velocity magnitudes. All the analyses were performed for the front and rear limbs separately using customized MATLAB routines.

Results

Joint center positions

An example (second athlete and first trial) of the joint center positions in the global reference frame estimated with markerless vis-à-vis marker-based approach is presented in Figure 2. Overall across all the subjects and trials, the joints of the front and rear limbs followed similar trajectories measured with both the systems. Differences [median (interquartile range)] between markerless and marker-based joint centers for the front shoulder, elbow, and wrist were found as 2.3 (0.8), 3.1 (0.8), and 1.8 (1.2) cm, respectively. These values for its rear counterparts were 2.3 (1.3), 3.1 (0.9), and 2.2 (2.5) cm, respectively.

Joint angles

Figure 3 illustrates an example (second athlete and first trial) of the joint angles at the shoulder, elbow, and wrist obtained with markerless and marker-based systems. Overall, the kinematic

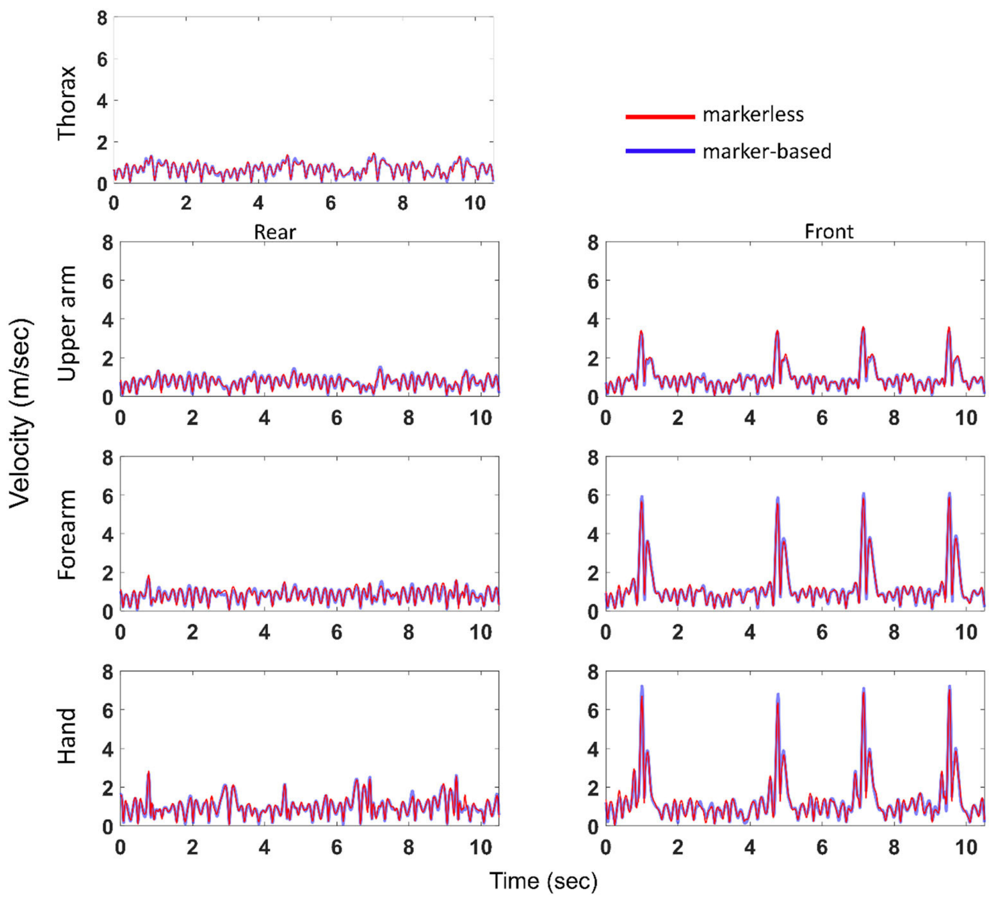

FIGURE 4
Segment velocity magnitudes (ordinate) at the thorax, and at the front and rear upper arm, forearm, and hand computed with marker-based and markerless methods and represented over time (abscissa). Example shown for the second athlete and first boxing trial. Blue and red colors represent marker-based and markerless segment velocities, respectively.

profiles estimated with both the methods exhibited qualitatively similar pattern, with some noticeable offsets.

Table 3 represents the statistical parameters *b*, *CI*, R^2, and *RMSD* when comparing markerless joint angles with marker-based ones. No values are reported for the elbow and wrist joint along Abd/Add and Int/Ext rotation, respectively, as these DoFs were restricted in the multibody kinematic optimization.

Along Abd/Add axis, higher *bias*, *CI,* and *RMSD*, and lower R^2 values were found at the wrist as compared to the shoulder. Similar outcomes were observed for Int/Ext rotation, with higher *bias*, *CI,* and *RMSD*, and lower R^2 at the elbow joint. As for the flexion/extension (Flex/Ext) axis, overall, lower *bias* was noticed at the shoulder joint, whereas lower *CI* and *RMSD* were observed at the elbow joint. When comparing across DoFs, highest *bias* and *RMSD* were seen along Int/Ext axis (*bias*: 2.5–13°; *RMSD*: 8.1–23°), followed by Flex/Ext (*bias*: 0.3–7.4°; *RMSD*: 7.3–20°) and Abd/Add (*bias*: −0.1 to 7.2°; *RMSD*: 6.3–11°).

When comparing across joints, lowest *bias* (−0.1°) and lowest *RMSD* (6.3°) were noticed at the shoulder joint, while revealing largest values at the elbow (*bias* up to 13° and *RMSD* up to 23°). Interestingly, between joints on both the sides, lower *bias*, R^2, and *RMSD* were observed at all the rear-side joints compared to its front counterparts with few exceptions.

Segment velocities

Figure 4 demonstrates linear segment velocity magnitudes at the thorax, upper arm, forearm, and hand for the second athlete and first trial. The velocity profiles estimated by both the systems displayed qualitatively similar patterns. The median peak velocities across all the athletes and trials measured by the marker-based system were different among segments, with highest velocity of 7.5 m/s at the front hand, followed by the forearm with 5.5 m/s, upper arm with 3.2 m/s, and thorax with 1.6 m/s. These peak velocities were observed while the boxers

TABLE 4 Bland–Altman bias (*b*), confidence interval (*CI*) along with coefficient of determination (R^2) and root mean square difference (*RMSD*) between markerless and marker-based methods for segment velocity magnitudes at the thorax, upper arm, forearm, and hand.

Velocity magnitudes (m/s)					
Segments	**Side**	***b***	***CI***	R^2	***RMSD***
Thorax		0.00	0.12	0.96	0.07
Upper arm	Front	−0.01	0.14	0.98	0.09
	Rear	−0.02	0.13	0.97	0.08
Forearm	Front	−0.03	0.20	0.98	0.14
	Rear	−0.02	0.14	0.98	0.09
Hand	Front	−0.01	0.23	0.98	0.17
	Rear	−0.01	0.17	0.97	0.11

Units for all parameters are in m/s except R^2 (no unit).

were throwing punches, and some small velocities (∼0–1.5 m/s) were noticed in between punches when they were performing footwork. The markerless system estimated similar results, with 7.0 m/s at the hand, 5.5 m/s at the forearm, 3.5 m/s at the upper arm, and 1.6 m/s at the thorax.

The results of Bland–Altman analysis showed a relatively good agreement between both the systems for the segment velocity magnitudes (Table 4). Very small bias was observed for every segment. Confidence intervals were slightly higher (between 0.10 and 0.25 m/s), but remained small compared to the peak velocity observed during the punch. Overall, the segment with lowest velocity (i.e., thorax) performed the highest level of agreement between both the systems. When comparing between the sides, the rear-side segments showed better agreement as compared to the front-side ones.

Discussion

The purpose of the study was to assess whether markerless motion capture system can be exploited to estimate upper-limb kinematics as a substitute to maker-based approach for analyzing sports-performance during "in the field" boxing. We assessed joint center positions, joint angles, and segment velocities obtained with a commercially available markerless motion data processing software (Theia3D) compared to those estimated with classical marker-based method. Multibody models and optimization methods were designed to match at best between the two approaches.

Across all the subjects and trials, the median 3D distances between corresponding joint centers were noticed in the range ∼1.5–2.5 cm for all the joints, except the elbow exceeding 3 cm. Our findings were comparable to those, who reported an average difference in the range 1.1–2.4 cm for the upper extremity joints during a treadmill walking activity (Kanko et al., 2021a) and in the range ∼2.0–4.7 cm during a throwing activity (Nakano et al., 2020).

The upper limb joint angles captured a varying agreement across all the joints and DoFs. Highest *bias* and *RMSD* were observed along Int/Ext rotation axis and lowest along Abd/Add axis, confirming the remarks reported for the lower extremity joints during gait (Kanko et al., 2021a). However, the values obtained for the upper extremity joints were higher than those obtained for the lower extremity. For instance, *RMSD* along Int/Ext rotation axis was found in the range 6.9–13.2° for the lower extremity (Kanko et al., 2021a), whereas 8.1–23° was observed for the upper extremity in the present work. These higher values could be a consequence of weaker estimation of segment poses during a dynamic activity as compared to gait, respecting the previous evidence of pose estimation performance being task-specific (Nakano et al., 2020; Needham et al., 2021). With regard to all statistical parameters, the shoulder joint demonstrated better agreement between the methods across all DoFs, except Flex/Ext axis along which elbow joint was seen superior. Furthermore, relatively lesser agreement was observed for the front-side joints in general. It is perhaps because of relatively higher and faster movement of the front-side segments resulting higher differences. The front arm is also more often fully extended, a configuration in which determining Int/Ext rotation becomes problematic.

With regard to the segment velocities, the markerless system performed a good-to-strong level of agreement, with maximum *RMSD* ≤ 0.17 m/s and with a strong R^2 (0.96–0.98). Both systems captured highest velocity at the hand (7–7.5 m/s) followed by the other body segments in the kinematic chain. These tendencies corroborate the findings who reported average punch contact velocities in the range 5.9–8.2 m/s for combination of punches (Whiting et al., 1988; Piorkowski et al., 2011).

Overall, we have noticed a higher degree of agreement at the proximal joints/segments between the data collecting methods. This could be a result of the pose algorithm, which may perform less for distal segments, especially for the hand in the considered boxing task as it moves relatively quicker. It could also be a consequence of the multibody kinematic optimization, in which the proximal segments are more constrained than distal segments (they "inherit" the constraints from distal segments) and thus less sensitive to measurement errors. Nevertheless, further investigations are required to confirm these hypotheses.

While interpreting the degree of agreement or differences between the two methods, we would like to highlight few potential sources of errors and assumptions which may influence the results. Marker-based kinematics are prone to misplacement or inconsistent placement of markers. Although the markers were placed by the same operator with adequate palpation skills, some degrees of variability/inconsistency cannot be denied. On the other hand, markerless kinematics are normally susceptible to the quality of 2D video data determined by particularly spatial resolution, exposure time, and angle of view specified for the motion under study. As such we have not studied the sensitivity of these parameters on the kinematic accuracy, we can expect some changes (improvement/deterioration) in the kinematics as reported in other studies (Nakano et al., 2020). Nevertheless, we believe that these impacts would likely to be minimal as data collection was carried out under proficient supervision, and the video data were randomly and qualitatively checked after each acquisition. Furthermore, a repeatability study for the upper extremity seems relevant in the future, although the same has been assessed during gait showing reliable estimation of the lower-limb kinematics using Theia3D (Kanko et al., 2021b). Another source of error commonly known as soft tissue artifact (Camomilla et al., 2017) may impact marker-based kinematics to certain extent, although multibody kinematic optimization was implemented to compensate for it (Begon et al., 2018). Apart from the probable sources of errors, there are some likely differences in defining segment reference frames between the Theia3D kinematic model and marker-based model. For instance, in the marker-based model, the long axis of the thorax is defined between the thoracic joint center and the cervical joint center estimated with regression equations (Dumas and Wojtusch, 2018). Although the Theia3D model uses identical landmarks derived from pose matrices to define the axis, any differences in estimating joint centers would impact the segment frame and thereby resulting in offsets and distortions in the joint angles. We acknowledge that such discrepancies could not be avoided; nevertheless, definition of marker-based joint centers for the shoulder, elbow, and wrist was in accordance with the study carried out by the team involving in Theia3D development (Kanko et al., 2021a,b).

The present work may provide practical avenues to analyze the performance and skill levels of athletes by assessing upper extremity kinematics. For instance, the joint center trajectories and angles have been shown to vary based on the characteristics of boxing type and level of expertise (Whiting et al., 1988; Dinu and Louis, 2020). Measuring these kinematic variables would be necessary to analyze and enhance punches that require a distinct segment orientation in different planes. Information on trajectories of different punch types will also help the combatant to deflect or escape blows (Piorkowski et al., 2011). Furthermore, ranges of motion may provide insights on predisposing factors of injury, as larger joint motion has been reported to implicate joint injury, particularly at the shoulder (Lenetsky et al., 2015). The literature on assessing segment velocities suggests that punch velocity is crucial for optimal performance in boxing (Whiting et al., 1988), with higher values reported for elite boxers (Dinu and Louis, 2020). Attaining high velocity at the fist is typically a result of contribution of other body segments in the kinematic chain. The latter study indicated that the shoulder contributed most to hook and uppercut punch both in junior and in elite boxers. They also noticed moderate differences in segment contribution between high- and low-performing boxers, underlining the need for accurate estimation of segment velocities. Despite such knowledge on the biomechanical distinctions across boxing styles and athletes, receiving timely feedback has been a major burden to coaches and athletes with marker-based methods. In this context, the markerless system is easily deployable both in training sessions and in live combats, while maintaining comparable kinematics to marker-based approach. It is worthwhile to mention that the usability of the kinematic variables as performance measures will depend upon the context of application within an allowable error limit, and this remains an explorable avenue. In marker-based motion analysis, with reference to gait analysis, 5° of error is generally considered the maximum accepted (McGinley et al., 2009). This 5° error correspond to the lower limb, and no such value seems to be present in the literature for the upper limb. That said, we expect segment velocities across all the segments, and joint angles along Abd/Add and Flex/Ext axes can be used with reasonable confidence. The wrist joint angles should be dealt with caution due to relatively poor agreement.

There are few limitations of the present work to acknowledge. We assessed only the upper extremity kinematics, although may not be sufficient to underscore a wide range of performance descriptors such as athletes' stability and kinetic characteristics such as punching force. For example, distribution of the forces between the legs has a considerable effect on punching performance in terms of both stability and fist velocity (Stanley et al., 2018; El-Oujaji et al., 2019). For direct measurement, this would, however, require additional arrangements such as force plates (Piorkowski et al., 2011; Stanley et al., 2018) and instrumented punch bags, making it cumbersome and unsuitable for monitoring "in the field" matches. Studies have also also showed the possibility of estimating punching force using wearable devices and external contact loads (Robert et al., 2013; Muller et al., 2020) using marker-based approach without the need of force sensors. Estimation of these variables using markerless video-based approach seems relevant in assessing sports-performance. One important limitation to highlight is the small number of elite athletes participated in the study. It is also noteworthy to remark that comparative assessment was performed for shadow boxing trials, i.e., one single athlete throwing punches without interaction with the opponent. Although it would be pertinent to analyze the performance of both the athletes in close-combats, evaluating with marker-based motion capture system would be questionable due to its inherent limitations. Moreover, the boxers performed trials without the usage of gloves and boxing outfit as it was convenient to place markers on body landmarks. It would be interesting to analyze the sensitivity of the markerless kinematics in response to traditional boxing attire.

Conclusion

As a first "in the field" study of a highly dynamic sport, we evaluated 3D joint center positions, joint angles, and segment velocities of the upper extremity of three elite athletes estimated with a markerless approach in comparison with those obtained with marker-based method. We observed a median difference of <2.5 cm for the shoulder and wrist, and slightly higher than 3 cm for the elbow joint between the two approaches. While assessing the joint angles, the shoulder joint largely exhibited a higher level of agreement with RMSD in the range of 6–12°, whereas the wrist and elbow joint displayed more than or equal to 20° in some DoFs. The agreement along the Int/Ext axis was consistently poor across all the DoFs. Segment velocities demonstrated a strong level agreement between the two methods showing a maximum RMSD of 0.17 m/s. Overall results indicated higher levels of agreement between the methods for segment velocities compared to joint angles. Given the practicality of the markerless motion capture system out of the laboratory environment, the results will help both athletes and coaches to analyze sports-performance. Future studies will focus on analyzing both the athletes in close-combat situations with markerless method.

Data availability statement

The datasets presented in this article are not readily available because of ethical concerns regarding confidentiality. Requests to access the datasets should be directed to thomas.robert@univ-eiffel.fr.

Ethics statement

The studies involving human participants were reviewed and approved by Local Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

BL, AM, RD, and TR: design and conceptualization. BL, LR, and TR: data collection. BL: writing—original draft preparation. AM, RD, LR, and TR: writing, reviewing, and editing. TR: supervision. LR: funding. All authors have read and agreed to the submitted version of the manuscript.

Funding

This study was partly funded by the ANR PPR STHP 2020 (project PerfAnalytics, ANR 20-STHP-0003).

Acknowledgments

The authors are grateful to the National Institute of Sports, Expertise, and Performance (INSEP) for providing the boxing platform and inviting their elite boxers to participate in the study.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fspor.2022.939980/full#supplementary-material

References

Al Borno, M., O'Day, J., Ibarra, V., Dunne, J., Seth, A., Habib, A., et al. (2022). OpenSense: an open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations. *J. Neuroeng. Rehabil.* 19, 1–11. doi: 10.1186/s12984-022-01001-x

Armitano-Lago, C., Willoughby, D., and Kiefer, A. W. (2022). A SWOT analysis of portable and low-cost markerless motion capture systems to assess lower-limb musculoskeletal kinematics in sport. *Front. Sport. Act. Living* 3, 809898. doi: 10.3389/fspor.2021.809898

Begon, M., Andersen, M. S., and Dumas, R. (2018). Multibody kinematics optimization for the estimation of upper and lower limb human joint kinematics: a systematized methodological review. *J. Biomech. Eng.* 140, 1–11. doi: 10.1115/1.4038741

Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1, 307–310. doi: 10.1016/S0140-6736(86)90837-8

Camomilla, V., Dumas, R., and Cappozzo, A. (2017). Human movement analysis: the soft tissue artefact issue. *J. Biomech.* 62, 1–4. doi: 10.1016/j.jbiomech.2017.09.001

Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., and Sheikh, Y. (2021). OpenPose: realtime multi-person 2D pose estimation using part affinity fields. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 172–186. doi: 10.1109/TPAMI.2019.2929257

Colyer, S. L., Evans, M., Cosker, D. P., and Salo, A. I. T. (2018). A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports Med. Open* 4, 24. doi: 10.1186/s40798-018-0139-y

Corazza, S., Mündermann, L., and Andriacchi, T. (2007). A framework for the functional identification of joint centers using markerless motion capture, validation for the hip joint. *J. Biomech.* 40, 3510–3515. doi: 10.1016/j.jbiomech.2007.05.029

Corazza, S., Mündermann, L., Chaudhari, A. M., Demattio, T., Cobelli, C., and Andriacchi, T. P. (2006). A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach. *Ann. Biomed. Eng.* 34, 1019–1029. doi: 10.1007/s10439-006-9122-8

Cronin, N. J., Rantalainen, T., Ahtiainen, J. P., Hynynen, E., and Waller, B. (2019). Markerless 2D kinematic analysis of underwater running: a deep learning approach. *J. Biomech.* 87, 75–82. doi: 10.1016/j.jbiomech.2019.02.021

Dinu, D., and Louis, J. (2020). Biomechanical analysis of the cross, hook, and uppercut in junior vs. elite boxers: implications for training and talent identification. *Front. Sport. Act. Living* 2, 598861. doi: 10.3389/fspor.2020.598861

Drazan, J. F., Phillips, W. T., Seethapathi, N., Hullfish, T. J., and Baxter, J. R. (2021). Moving outside the lab: markerless motion capture accurately quantifies sagittal plane kinematics during the vertical jump. *J. Biomech.* 125, 110547. doi: 10.1016/j.jbiomech.2021.110547

Dumas, R., and Wojtusch, J. (2018). "Estimation of the body segment inertial parameters for the rigid body biomechanical models used in motion analysis." in Handbook of Human Motio, eds B. Müller and S. I. Wolf (Cham: Springer), 47–77. doi: 10.1007/978-3-319-14418-4_147

El-Oujaji, S., Provot, T., Bourgain, M., and Dinu, D. (2019). Difference of stability between two elite boxing groups: a preliminary study. *Comput. Methods Biomech. Biomed. Eng.* 22, S481–S483. doi: 10.1080/10255842.2020.1714989

El-Sallam, A., Bennamoun, M., Sohel, F., Alderson, J., Lyttle, A., and Rossi, M. (2013). A low cost 3D markerless system for the reconstruction of athletic techniques. *Proc. IEEE Work. Appl. Comput. Vis.* 222–229. doi: 10.1109/WACV.2013.6475022

Johnson, W. R., Mian, A., Donnelly, C. J., Lloyd, D., and Alderson, J. (2018). Predicting athlete ground reaction forces and moments from motion capture. *Med. Biol. Eng. Comput.* 56, 1781–1792. doi: 10.1007/S11517-018-1802-7

Kanko, R. M., Laende, E., Selbie, W. S., and Deluzio, K. J. (2021b). Inter-session repeatability of markerless motion capture gait kinematics. *J. Biomech.* 121, 110422. doi: 10.1016/j.jbiomech.2021.110422

Kanko, R. M., Laende, E. K., Davis, E. M., Selbie, W. S., and Deluzio, K. J. (2021a). Concurrent assessment of gait kinematics using marker-based and markerless motion capture. *J. Biomech.* 127, 110665. doi: 10.1016/j.jbiomech.2021.110665

Kimm, D., and Thiel, D. V. (2015). Hand speed measurements in boxing. *Proc. Eng.* 112, 502–506. doi: 10.1016/j.proeng.2015.07.232

Lahkar, B. K., Rohan, P.-Y., Assi, A., Pillet, H., Bonnet, X., Thoreux, P., et al. (2021). Development and evaluation of a new methodology for Soft Tissue Artifact compensation in the lower limb. *J. Biomech.* 122, 110464. doi: 10.1016/j.jbiomech.2021.110464

Lenetsky, S., Brughelli, M., and Harris, N. K. (2015). Shoulder function and scapular position in boxers. *Phys. Ther. Sport* 16, 355–360. doi: 10.1016/j.ptsp.2015.02.003

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., et al. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* 21, 1281–1289. doi: 10.1038/s41593-018-0209-y

McGinley, J. L., Baker, R., Wolfe, R., and Morris, M. E. (2009). The reliability of three-dimensional kinematic gait measurements: a systematic review. *Gait Posture* 29, 360–369. doi: 10.1016/j.gaitpost.2008.09.003

Mcguirk, T. E., Perry, E. S., Sihanath, W. B., Riazati, S., Patten, C., and Wade, E. (2022). Feasibility of markerless motion capture for three-dimensional gait assessment in community settings. *Front. Hum. Neurosci.* 16, 86748516. doi: 10.3389/fnhum.2022.867485

Muller, A., Pontonnier, C., and Dumont, G. (2020). Motion-based prediction of hands and feet contact efforts during asymmetric handling tasks. *IEEE Trans. Biomed. Eng.* 67, 344–352. doi: 10.1109/TBME.2019.2913308

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., et al. (2020). Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras. *Front. Sport. Act. Living* 2, 50. doi: 10.3389/fspor.2020.00050

Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., et al. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. *Sci. Rep.* 11, 1–11. doi: 10.1038/s41598-021-00212-x

Pagnon, D., Domalain, M., and Reveret, L. (2022). Pose2Sim: an end-to-end workflow for 3D markerless sports kinematics—part 2: accuracy. *Sensors* 22, 2712. doi: 10.3390/s22072712

Piorkowski, B. A., Lees, A., and Barton, G. J. (2011). Single maximal versus combination punch kinematics. *Sport. Biomech.* 10, 1–11. doi: 10.1080/14763141.2010.547590

Punchihewa, N. G., Miyazaki, S., Chosa, E., and Yamako, G. (2020). Efficacy of inertial measurement units in the evaluation of trunk and hand kinematics in baseball hitting. *Sensors* 20, 1–11. doi: 10.3390/s20247331

Putnam, C. A. (1993). Sequential motions of body segments in striking and throwing skills: descriptions and explanations. *J. Biomech.* 26, 125–135. doi: 10.1016/0021-9290(93)90084-R

Riazati, S., Mcguirk, T. E., Perry, E. S., Sihanath, W. B., and Patten, C. (2022). Absolute reliability of gait parameters acquired with markerless motion capture in living domains. *Front. Hum. Neurosci.* 16, 867474. doi: 10.3389/fnhum.2022.867474

Robert, T., Causse, J., and Monnier, G. (2013). Estimation of external contact loads using an inverse dynamics and optimization approach: general method and application to sit-to-stand maneuvers. *J. Biomech.* 46, 2220–2227. doi: 10.1016/j.jbiomech.2013.06.037

Stanley, E., Thomson, E., Smith, G., and Lamb, K. L. (2018). An analysis of the three-dimensional kinetics and kinematics of maximal effort punches among amateur boxers. *Int. J. Perform. Anal. Sport* 18, 835–854. doi: 10.1080/24748668.2018.1525651

Stenum, J., Rossi, C., and Roemmich, R. T. (2021). Two-dimensional video-based analysis of human gait using pose estimation. *PLoS Comput. Biol.* 17, e1008935. doi: 10.1371/journal.pcbi.1008935

Whiting, W. C., Gregor, R. J., and Finerman, G. A. (1988). Kinematic analysis of human upper extremity movements in boxing. *Am. J. Sports Med.* 16, 130–136. doi: 10.1177/036354658801600207

Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., et al. (2002). ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion – part I: ankle, hip, and spine. *J. Biomech.* 35, 543–548. doi: 10.1016/S0021-9290(01)00222-6

Wu, G., Van Der Helm, F. C. T., Veeger, H. E. J., Makhsous, M., Van Roy, P., Anglin, C., et al. (2005). ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion – part II: shoulder, elbow, wrist and hand. *J. Biomech.* 38, 981–992. doi: 10.1016/j.jbiomech.2004.05.042

Zajac, F. E., and Winters, J. M. (1990). Modeling musculoskeletal movement systems: joint and body segmental dynamics, musculoskeletal actuation, and neuromuscular control. *Mult. Muscle Syst.* 121–148. doi: 10.1007/978-1-4613-9030-5_8