

Establishing a Global Standard for Wearable Devices in Sport and Exercise Medicine: Perspectives from Academic and Industry Stakeholders

Garrett I. Ash^{1,2}, Matthew Stults-Kolehmainen^{3,4}, Michael A. Busa^{5,6}, Allison E. Gaffey^{1,7},
Konstantinos Angeloudis⁸, Borja Muniz-Pardos⁹, Robert Gregory¹⁰, Robert A. Huggins¹¹, Nancy
S. Redeker¹², Stuart A. Weinzimmer¹³, Lauren A. Grieco¹⁴, Kate Lyden¹⁵, Esmeralda Megally¹⁶,
Ioannis Vogiatzis¹⁷, LaurieAnn Scher^{18,19}, Xinxin Zhu²⁰, Julien S. Baker²¹⁻²³, Cynthia
Brandt^{1,2,24}, Michael S. Businelle^{25,26}, Lisa M. Fucito²⁷⁻²⁹, Stephanie Griggs³⁰, Robert Jarrin^{31,32},
Bobak J. Mortazavi³³, Temiloluwa Prioleau³⁴, Walter Roberts²⁷, Elias K. Spanakis^{35,36},
Laura M. Nally¹³, Andre Debruyne^{37,38}, Norbert Bachl³⁷⁻⁴⁰, Fabio Pigozzi^{37-38,41-42}, Farzin
Halabchi^{37,43-44}, Dimakatso A. Ramagole^{37,45}, Dina C Janse van Rensburg^{37,45}, Bernd
Wolfarth^{37,46}, Chiara Fossati⁴¹, Sandra Rozenstoka^{37,38,47}, Kumpei Tanisawa⁴⁸, Mats
Börjesson^{37,49,50}, José Antonio Casajus^{9,37}, Alex Gonzalez-Aguero^{9,37}, Irina Zelenkova^{9,51}, Jeroen
Swart^{37,52}, Gamze GURSOY⁵³, William Meyerson^{54,55}, Jason Liu⁵³, Dov Greenbaum^{53,56,57}, Yannis
P. Pitsiladis^{8,37,38}, Mark B. Gerstein^{53,55,58,59}

1. Veterans Affairs Connecticut Healthcare System, West Haven, Connecticut, USA
2. Center for Medical Informatics, Yale University, New Haven, Connecticut, USA
3. Digestive Health Multispecialty Clinic, Yale-New Haven Hospital, New Haven, Connecticut, USA
4. Department of Biobehavioral Sciences, Teachers College, Columbia University, New York, New York, USA

5. Center for Human Health & Performance, Institute for Applied Life Sciences, University of Massachusetts Amherst, Amherst, Massachusetts, USA
6. Department of Kinesiology, University of Massachusetts, Amherst, Massachusetts, USA
7. Department of Internal Medicine (Cardiovascular Medicine), Yale School of Medicine, New Haven, Connecticut, USA
8. Collaborating Centre of Sports Medicine, University of Brighton, Eastbourne, UK
9. GENUD Research group, Faculty of Health and Sport Sciences, University of Zaragoza, Zaragoza, Spain
10. Department of Health and Movement Sciences, Southern Connecticut State University, New Haven, Connecticut, USA
11. Department of Kinesiology, Korey Stringer Institute, University of Connecticut, Storrs, Connecticut, USA
12. Yale School of Nursing, Orange, Connecticut, USA
13. Department of Pediatrics, Yale School of Medicine, New Haven, Connecticut, USA
14. Google Health, Palo Alto, California, USA
15. VivoSense, Denver, Colorado, USA
16. Xsensio, Lausanne, Switzerland
17. Department of Sport, Exercise and Rehabilitation, School Health & Life Sciences, Northumbria University, Newcastle, Upon Tyne, UK. Member of the European Respiratory Society Digital Health Working Group.
18. Member of Consumer Technology Association Working Groups for Health Technology Standards, Washington, District of Columbia, USA.
19. Fitscript^{LLC}, New Haven, Connecticut, USA

20. Center for Biomedical Data Science, Yale School of Medicine, New Haven, Connecticut, USA
21. Faculty of Sports Science, Ningbo University, China.
22. School of Health and Life Sciences, Institute for Clinical Exercise & Health Science, University of the West of Scotland, South Lanarkshire, Scotland, UK.
23. Centre for Health and Exercise Science Research, Department of Sport, Physical Education and Health, Hong Kong Baptist University, Kowloon Tong, Hong Kong.
24. Emergency Medicine, Yale School of Medicine, New Haven, Connecticut, USA
25. Department of Family and Preventive Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA
26. Tobacco Settlement Endowment Trust Health Promotion Research Center, Stephenson Cancer Center, Oklahoma City, Oklahoma, USA
27. Department of Psychiatry, Yale School of Medicine, New Haven, Connecticut, USA
28. Yale Cancer Center, New Haven, CT, USA.
29. Smilow Cancer Hospital, Yale-New Haven Hospital, New Haven, CT, USA.
30. Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, Ohio, USA
31. Department of Emergency Medicine, George Washington University, Washington, District of Columbia, USA
32. Department of Biochemistry and Molecular & Cellular Biology, Georgetown University Medical Center, Washington, District of Columbia, USA
33. Department of Computer Science and Engineering, Texas A&M University, College Station, Texas, USA

34. Department of Computer Science, Dartmouth College, Hanover, New Hampshire, USA
35. University of Maryland School of Medicine, Baltimore, Maryland, USA.
36. Division of Endocrinology, Baltimore Veterans Affairs Medical Center, Maryland, USA.
37. International Federation of Sports Medicine (FIMS), Lausanne, Switzerland
38. European Federation of Sports Medicine Associations (EFSMA), Lausanne, Switzerland
39. Institute of Sports Science, University of Vienna, Vienna, Austria
40. Austrian Institute of Sports Medicine, Vienna, Austria
41. Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Rome, Italy.
42. Villa Stuart Sport Clinic, FIFA Medical Center of Excellence, Rome, Italy
43. Sports Medicine Research Center, Neuroscience Institute, Tehran University of Medical Sciences, Tehran, Iran (MHPS, MSG)
44. Department of Sports and Exercise Medicine, Tehran University of Medical Sciences, Tehran, Iran (FH)
45. Section Sports Medicine, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa.
46. Department of Sports Medicine, Humboldt University and Charité University School of Medicine, Berlin, Germany.
47. FIMS Collaboration Centre of Sports Medicine, Sports laboratory, Riga, Latvia.
48. Faculty of Sport Sciences, Waseda University, Tokorozawa, Japan.
49. Center for Health and Performance, Dept of Molecular and Clinical Medicine, Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden.

50. Sahlgrenska University Hospital, Dept of MGA, Region of Western Sweden, Gothenburg, Sweden.
51. I.M. Sechenov First Moscow State Medical University (Sechenov University), Ministry of Health of Russia, Moscow, Russia
52. FIMS Collaboration Centre of Sports Medicine, Division of Physiological Sciences & HPALS Research Centre, University of Cape Town, Cape Town, South Africa
53. Program in Computational Biology & Bioinformatics, Yale University, New Haven, Connecticut, USA
54. Duke Psychiatry & Behavioral Sciences, Duke Medicine, Durham, North Carolina, USA
55. Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, Connecticut, USA
56. Zvi Meitar Institute for Legal Implications of Emerging Technologies, Interdisciplinary Center Herzliya, Herzliya, Israel
57. Harry Radzyner Law School, Interdisciplinary Center Herzliya, Herzliya, Israel
58. Department of Computer Science, Yale University, New Haven, Connecticut, USA
59. Department of Statistics and Data Science, Yale University, New Haven, Connecticut, USA

Corresponding Author:

Yannis P. Pitsiladis

Collaborating Centre of Sports Medicine

University of Brighton

Eastbourne

United Kingdom

Email: y.pitsiladis@brighton.ac.uk

ACKNOWLEDGMENTS

The authors thank Dr. David Korfhagen for transcribing the session recordings, Ms. Chanelle Simmons for providing edits and comments on the manuscript, and the organizers of the virtual events, especially Ms. Leslie Dawkins from the Yale Center for Biomedical Data Science

RUNNING HEADING

Global Standard for Wearable Devices: Perspectives from Academic and Industry Stakeholders

ABSTRACT

Millions of consumer sport and fitness wearables (CSFWs) are used worldwide, and millions of datapoints are generated by each device. Moreover, these numbers are rapidly growing, and they contain a heterogeneity of devices, data types, and contexts for data collection. Companies and consumers would benefit from guiding standards on device quality and data formats. To address this growing need, we convened a virtual panel of industry and academic stakeholders, and this manuscript summarizes the outcomes of the discussion. Our objectives were to identify: (1) key facilitators and barriers to participation by CSFW manufacturers in guiding standards and (2) stakeholder priorities. The venues were the Yale Center for Biomedical Data Science Digital Health Monthly Seminar Series (62 participants) and the New England Chapter of the American College of Sports Medicine Annual Meeting (59 participants). In the discussion, stakeholders outlined both facilitators (e.g., commercial return on investment in device quality, lucrative research partnerships, and transparent and multilevel evaluation of device quality) and barriers (e.g., competitive advantage conflict, lack of flexibility in previously developed devices) to participation in guiding standards. There was general agreement to adopt Keadle et al.'s standard pathway for testing devices (i.e., benchtop, laboratory, field-based, implementation) without consensus on the prioritization of these steps. Overall, there was enthusiasm not to add prescriptive or regulatory steps, but instead create a networking hub that connects companies to consumers and researchers for flexible guidance navigating the heterogeneity, multi-tiered development, dynamicity, and nebulosity of the CSFW field.

Abstract word count: 243/250

KEY POINTS

- We convened a virtual panel of industry and academic stakeholders to discuss the need for guiding standards of device quality and data formatting in the rapidly expanding market of consumer sport and fitness wearables.
- Stakeholders agree that such standards could add value to commercial return on investment and provide constructive critiques to manufacturers, especially when focused on the benchtop testing stage.
- The large company representative noted limited flexibility to unveil or modify devices at this basic level and suggested the alternative of analytics on big data generated by widely used devices (e.g., within-person consistency).
- Stakeholders recommended providing a networking hub that helps companies and researchers acquaint and synergistically navigate the concerns with the CSFW field.

1. SCOPE AND OBJECTIVE

The Internet of Things, a.k.a. the online connectivity of physical devices, has grown to encompass 35 billion devices worldwide [1]. Around 1 billion of these devices are wearables, defined by the European Commission as “body-borne computational and sensory devices which can sense the person who wears them and/or their environment” [2]. The wearables market is driven by the subset for consumer sport and fitness (CSFWs); fitness watches alone shipped over 100 million units in 2020 [3]. An increasing number of health care professionals, sport scientists and other employees within international sports and medical federations, at rehabilitation centers, sports clubs, and sporting events use some form of CSFWs.

Among these millions of CSFWs, each generates millions of biometric datapoints per year if used regularly. A typical smartwatch, for example provides the user with minute-by-minute readings of step counts, active energy expenditure, heart rate, and other variables. The raw data metrics that produce these minute-by-minute readings are even denser. For example, energy expenditure is derived from acceleration counts sampled several thousand times per minute.

This landscape is not only large but also heterogeneous. First, while smartwatches are common examples, there are many available CSFW devices and data types. Maintaining comprehensive lists is a massive undertaking beyond our scope that others have completed periodically [4], but we have shown some examples of types of devices in Table 1. Second, these freely marketed products can be used by anyone anywhere, creating a limitless number of contexts for data collection in terms of user demographics, health characteristics, ambient weather, air quality, magnetic fields, and other factors.

CSFWs bear some key differences from research-grade devices such as Actiwatches™ or Actigraph™ accelerometers. First, CSFWs present their data through software that can be operated and viewed by a lay user, whereas research devices recommend that a technician process the data before sharing with the consumer. An example is using Actiwatch™ software to manually adjust the automatically detected sleep times based on other indicators like light and user-marked bedtime [5]. Second, CSFWs typically issue data reports with a single (averaged) value for each minute, whereas research devices can include raw values which are sampled thousands of times per minute. Third, CSFWs each have a set method for converting raw data to health outcome metrics, whereas research devices offer the operator flexibility. An example is choosing the equation for converting actigraphy counts to activity intensity from various options that have been validated in various studies.

Table 1. Examples of CSFWs.

Product Category	Example Product	Technical characteristics
Position	Apple Watch TM	Accelerometer detects arm movement indicative of standing versus sitting.
Motion or activity	Apple Watch TM , Empatica TM E4, Garmin TM	Accelerometer detects arm movement indicative of body movement.
Location	Apple Watch TM , Garmin TM	Global positioning system.
Biomechanics	IMeasureU, Leomo TM	Inertial measurement unit measures 9 axes (i.e., x-y-z directions of accelerometer, gyroscope, magnetometer) attached to each limb of interest.
Heart rate	Apple Watch TM , Garmin TM , Empatica TM E4, Biostrap TM , Whoop TM , Oura TM ,	Photoplethysmography: shines light upon the skin surface and a photodetector measures arterial pulsatile flow.
Blood oxygen saturation	Biostrap TM	Photoplethysmography as above with a photodetector measuring blood color as an index of saturation.
Muscle tissue oxygen saturation	Moxy TM	Near-infrared spectroscopy: similar to photoplethysmography except light penetrates skin deeper, which allows comparison of arterial to capillaries and venous saturation to infer muscle tissue oxygen saturation.
Autonomic function	Apple Watch TM , Biostrap TM , Whoop TM , Oura TM ,	Photoplethysmography as above, followed by time-series analysis of heart rate variability.
Sweat composition and sweat lactate concentration	Xsensio Lab-on-Skin TM	A system of microfluidic channels transports sweat across electrodes which register pH and mineral contents, transmitting results by Bluetooth to a smartphone application. The system is held against the skin by a hydrophilic patch that powers the microfluidic motion.
Galvanic skin response (i.e., electrodermal activity)	Empatica TM E4	The watch measures skin conductivity indicated by the conductance of a current it runs between two electrodes that are pressed against the skin. Higher levels are associated with greater stress.
Body temperature	Empatica TM E4	Infrared thermopile sensor detects infrared energy emitted by the skin, which is directly associated with skin temperature.
Sleep	Whoop TM , Oura TM	Algorithm integrates accelerometer, gyroscope, and photoplethysmography heart rate.

2. QUALITY ASSURANCE, PRIVACY, AND DATA INTERPRETATION

As the CSFW market has rapidly expanded there has been increased focus on the quality assurance of CSFWs. For example, in an early study, researchers assessed the validity of two commercial wearables and determined that Fitbit™ heart monitoring was inaccurate, particularly with higher exercise intensity [7], resulting in two class action lawsuits [8, 9]. More recently, Peake and colleagues evaluated 61 wearables and found that only 5% matched their marketing claims according to accepted reference standards [4]. The validity and reliability of these devices also tend to vary depending on the variables that are measured. A review of 158 publications in which nine brands were examined, revealed that steps were generally measured accurately across brands in the laboratory but less so in field settings, and no device accurately measured energy expenditure [10]. A primary study of energy expenditure from 4 of these sensors and 8 others worn simultaneously by 19 adults drew a similar conclusion [11]. Moreover, variable gait patterns [12] (Figure 1) suggest the need for population-specific validations, which are currently lacking. Another concern is these studies likely use new or well-maintained devices thus leaving possible durability concerns such as the possibility of a dislodged accelerometer or faded photoplethysmography light understudied [13, 14]. Because the device market evolves rapidly, quickly outpacing the above studies of validity [10, 11], gait patterns [12] and other related research, this market necessitates fast and frequent, well informed comparisons to provide objective quality metrics. In this way, users can maximally benefit from CSFWs to monitor and understand their health behaviors.

Wearable devices typically lack the security that is afforded most personal data, thereby threatening an individual's privacy, which is often unbeknownst to them [15]. Privacy policies

are often ambiguous or extensive, so CSFW users may be largely unaware of the security policies of their data storage and sharing, including who may access, own, or sell their health data [16]. Data obtained from these devices generally do not fall under the regulatory purview of health privacy statutes. Consequently, workplace wellness programs could furnish wearable data to insurance companies, who may then choose to raise premiums or deny coverage for individuals exhibiting higher-risk behavior patterns (e.g., poor sleep, physical inactivity) [17, 18]. These decisions are particularly problematic when based on inaccurate data (e.g., periods of restful wakefulness may be interpreted as sleep) [19]. There is also the potential for data access and threats to confidentiality from outside parties, legally (sale of the company or its data) or illegally (hacking of databases or wireless transmissions) [20]. This disclosure is particularly concerning as Global Positioning System data can easily infer home address and 24 h biodata could theoretically carry a unique signature, akin to DNA and can be used for commercial purposes [21]. Many companies claim that data they share with outside parties are deidentified, but the United States Health Insurance Portability and Accountability Act (HIPAA) does not specify how to deidentify these data and there are several clear threats to privacy. Some protection against these threats may begin emerging in the European Union due to the recent General Data Protection Regulation (GDPR) designed to protect personal information. Unfortunately, a preliminary analysis suggests most consumer health applications fail to comply with the GDPR on numerous levels, especially regarding opaque privacy policies [22].

The best practices for interpreting and presenting CSFW data to consumers remain unclear and controversial. For instance, sleep watch data can harm consumers first by eliciting “preoccupation or concern with improving or perfecting wearable sleep data” and second by

accepting and believing wearable sleep data more readily than medical advice, standard sleep hygiene education, or validated laboratory sleep assessments [23]. Some research has addressed this problem by optimizing the timing of data presentation (i.e., just-in-time adaptive interventions) [24]. For example, if a night of sleep is inadequate, the Whoop™ smartwatch (Boston, MA) alerts the consumer to this problem when they should start getting ready for bed the next night [25]. A criticism of such an approach, however, is that it conveys paternalism, and furthermore, may impose overly generic sleep and physical activity requirements if their algorithms fail to capture individual physiological and psychological needs (e.g., benefit from positive versus negative reinforcement). In addition, brief message prompts may be an inadequate substitute for providing more comprehensive wellness education, in which consumer literacy and numeracy are considered. The latter, along with the relatively high cost of CSFWs, limits the diversity of consumers reached and subsequent research. A recent systematic review of 463 scientific papers found the most important research gap in the CSFW field was understanding the human-information interaction that determines the adoption, acceptance, and health impact of CSFWs [26].

In addition to issues surrounding data presentation to consumers, standardization of data for technical purposes is also a prominent concern. Various CSFWs collect data using different raw units, timescales, and coding languages. Data are also stored in different formats. Even the Coordinated Universal Time format for date and time stamping is often not followed. The United States' National Institutes of Health solved similar problems in the field of genomics with the Genomic Data Sharing Policy. Based on the Policy, federally funded researchers are required to format their data according to standards of the Genbank database, an annotated collection of all

publicly available DNA sequences that exchanges data with similar entities in Europe and Asia [27]. This requirement streamlines the process for other researchers and coders to download and integrate data. A similar process is needed for the large datasets derived from CSFWs to facilitate research, encourage market competition, and interoperability between devices and other systems such as the electronic health or medical record.

One body that could address these concerns is the United States Food and Drug Administration (FDA, Washington, DC) which is responsible for regulating medical devices. In the current digital age, this effort requires regulating not only the devices, but also their cybersecurity, software, artificial intelligence, and machine learning algorithms. This scope has led to an unprecedented focus on grey areas, such as defining the extent to which software can be updated before requiring reapproval. The FDA responded to these challenges by issuing dozens of formal guidance documents and recently launching the Digital Health Center of Excellence in September 2020. FDA has pledged extensive resources to develop the Center by raising awareness, engagement, and partnership with stakeholders [28]. However, the FDA does not oversee low risk products that are intended for general wellness use and unrelated to diagnosing or treating a chronic disease (i.e., most CSFWs) [29]. The FDA Digital Health Center of Excellence exemplifies the level of investment that is needed to keep regulatory processes abreast of the digital health revolution but does not offer tangible support to the CSFW field for issues like those described in the previous section. Tighter regulation could emerge in the European Union, which recently expanded the scope of its Medical Device Regulations (EU MDR), but legal opinions are mixed as to whether CSFWs will fall within the new scope [30-32] and the changes are too recent (May 26, 2021) to have judicial precedent.

3. PANEL LOGISTICS AND RECRUITMENT

The panel discussion topics were: (1) key facilitators and barriers to participation by CSFW manufacturers in global guiding standards and (2) stakeholder priorities. It was hosted on September 16, 2020, by the Yale Center for Biomedical Data Science Digital Health monthly seminar series using the Zoom™ video call platform (San Jose, CA). The seminar series previously has included panels, and we adopted their suggested maximum number of panelists (n = 5) and format: moderator introduction (7 min), 5 panelists giving self-introductions and explaining their company's or organization's profile (4 min each), audience questions (33 min).

The moderator introduced the meeting by briefly summarizing concerns with CSFWs (Sect. 2) and interest in guiding standards revealed by our previous panel discussion [6]. The moderator also introduced the International Federation of Sports Medicine (FIMS, Lausanne, Switzerland) and the European Federation of Sports Medicine Associations (EFSMA) noting both have pledged commitment to support guiding standards. FIMS advocates for both the consumers of CSFWs and the sports medicine researchers extracting data from CSFWs. This objective aligns with their overall mission to promote the well-being of all who are engaged in sports and exercise, to assist athletes in achieving optimal performance, and to promote the study and development of sports medicine throughout the world. A number of leaders from FIMS collaborating centers (A.D., N.B., F.P., F.H., D.A.R., D.C.J.V.M., B.W., S.R., M.B., J.A.C., A.G., J.S., Y.P.P.) some also having EFSMA memberships participated.

To fill the panelist spaces, we executed a recruitment strategy focused on attracting a mixture of large and small international and national companies. We invited the four largest worldwide

CSFW manufacturers [3], Google™ given their recent buy-in to CSFWs by acquiring Fitbit™, and five smaller companies within our professional network that we chose to cover critical categories of CSFWs (Table 1). Invitations were sent electronically to the public relations departments and/or personal contacts within each company and followed up with a postal letter if there was no initial reply. Google Health™ (Palo Alto, CA, represented by author L.G.) and Xsensio™ (Lausanne, Switzerland, represented by author E.M.) accepted the invitation. One large company declined the invitation stating the following reasons: (1) the company is already involved in numerous research efforts so do not see the added value of data standardization; (2) they are concerned about protecting the privacy of their customers' data; and (3) they have limited resources and would prefer to invest those resources once the strategy has come to fruition, versus in these early discussion stages. One small company also declined the invitation for this year but welcomed us to contact them in future years. The other 6 companies did not reply. Thus, 40% of companies expressed some interest, although only 20% agreed to participate.

We interpreted this recruitment result to mean that the idea of guiding standards has the potential to gain industry stakeholder attention, but it was not possible to convene a large discussion at this time. Therefore, as a short-term strategy to increase scope, the last 3 panelist spaces were used to include individuals who have experience collaborating with a variety of CSFW companies. The first space was filled by VivoSense™ (Denver, CO, represented by author K.L.) who consults with pharmaceutical companies by interpreting wearable sensor outcomes and has worked with hundreds of devices in this manner. The second panelist space was filled by a European Respiratory Society Digital Health Working Group (Lausanne, Switzerland) member (author

I.V.), who evaluates the role of CSFWs to develop large research initiatives. The third space was filled by GlucoseZone™ (author L.S.), a consumer mobile exercise app interfacing with CSFW wearables. Author L.S. also belonged to the Consumer Technology Association (CTA™) working groups for health technology standards on industry standards for product quality.

The panel audience was recruited by mass advertising on the Yale Center for Biomedical Data Science listserv (n = 355 faculty and graduate students) as well as via personal invitations that were extended to researchers and clinicians working with wearable devices from Yale University, Yale-New Haven Hospital, the United States Veterans Affairs Healthcare System, the United States National Institutes of Health Mobile Health Shared Resource, the New England Chapter of the American College of Sports Medicine (NEACSM), FIMS, and EFSMA. In total, 62 individuals attended the panel, among whom 43 have made substantive contributions and were invited to coauthor this manuscript (24 kinesiologists, 9 data scientists, 3 endocrinologists, 1 nurse, 2 sleep researchers, 3 behavioral psychologists, 1 strategic advisor). A condensed summary of the proceedings was broadcast on-demand at the NEACSM Annual Meeting (October 1-15, 2020) followed by a live discussion when attendees were invited to ask questions and provide comments (October 16, 2020). The session recordings were professionally transcribed and circulated to all authors so they could review and edit their contributions as desired. Authors G.A. and Y.P. then reviewed the edited transcript and wrote the first draft of this manuscript. All authors commented on subsequent versions of the manuscript until all authors were able to approve the final manuscript.

4. DISCUSSION TOPICS

4.1. What could incentivize industry stakeholders to engage with guiding standards of device quality and data formatting?

Individuals from both manufacturers in attendance (Google Health™, Xsensio™) were supportive of global guiding standards and expressed interest in joining. When these individuals were asked what incentivized them to join the panel, two themes emerged. The first theme was value with respect to consumer appeal and satisfaction. Third-party endorsement provided by global guiding standards could help them dispel stereotypes about poor quality of CSFWs created by controversies such as the Fitbit class action lawsuits [8, 9]. Also, user education provided by the central resource would promote more discerning selection of CSFWs and potentially increase appreciation of CSFWs that offer high validity, quality, and useful data. This education would increase the commercial value yielded by their development efforts. For example, user expectations of a tighter error margin will increase return on investment to achieve such expectations.

The second theme that emerged from the panel discussion was value with respect to scientific endeavors. The two manufacturer panelists stated an interest to participate in data mining research that would be facilitated by data standardization. For example, it is very challenging to compile and interpret physical activity accelerometer information from different populations and datasets because of the myriad of inter-study and inter-device variations in protocols for converting raw to clinical units [33]. Some examples are epoch lengths, count thresholds demarcating activity intensity, and detection and handling of non-wear time. Data

standardization would allow multicenter projects with data from thousands of individuals, thus, increasing the impact of associated research and potential health outcomes.

Two of the consultant panelists also noted observing scenarios where companies benefit from having high quality and accessible data, as defined by an unambiguous list of endpoints and reference standards. The author from VivoSense™ reported that device manufacturers often miss opportunities to collaborate on drug trials if they are incompatible with the analytic software the trial is using. Evidence presented at the 2020 Annual Congress of the European Respiratory Society suggested that within the new ecosystem of clinical trials when companies have validated and accessible data, this presents a number of business opportunities: they can supply data directly to researchers and pharmaceutical companies, collect data directly from hospitals and universities, and collaborate with leading bioinformaticians to improve their algorithms for data processing and interpretation.

4.2. What stage of device development should the central resource target in order to achieve the quality assurance and data standardization objectives?

Since Keadle et al.'s standard testing pathway for wearable technology has multiple steps (benchtop, laboratory, free-living, implementation) [34] (Table 2), the panel debated which of these steps should be the focus of guiding standards' validation checks, quality assurance procedures, and standardization of data outputs.

Table 2. Device evaluation stages using a wrist accelerometer as a case example.

Step of Evaluation	Definition	Example	Goal achieved for the guiding reference	Desired Standard
Benchtop	Evaluate response to standardized synthetic signals.	Attach accelerometer to calibrated shaker plate and compare its outputs to the expected accelerations.	*Quality assurance by troubleshooting at the most basic level *Data formatting by developing common basic physical units *Interoperability of devices by standardizing units to facilitate algorithms that are transferable	Agreement with shaker plate of 3% (to match laboratory standard proposed below).
Laboratory	Compare device outputs against criterion measures, upon human participants wearing the device and a criterion instrument under conditions of controlled physiological inputs (e.g., controlled graded exercise).	Energy expenditure outputs from the device are compared to oxygen consumption.	*Quality assurance in laboratory context *Data formatting by reaching consensus on the best-validated equations for converting physical units to clinical metrics	Agreement of 3% [10]
Field-based	Compare device outputs against a reference standard device, upon human participants wearing the device and a reference standard during naturalistic and variable conditions of daily living, to assess metrics like reliability and time delay from the reference standard.	Participants wear the device concurrently with a reference standard hip accelerometer.	*Quality assurance in field context *Data formatting by reaching consensus on the best-validated equations for converting physical units to clinical metrics	Agreement of 10% [10]
Implementation	Follow consumers wearing the device post-marketing for metrics like user satisfaction and device durability, the latter assessed as consistency of readings within-persons over time.	Participants are tracked for consistency of smartwatch-measured weekly energy expenditure, a known stable variable [35]	*Quality assurance in implementation context	Novel testing approaches proposed by this panel for which standard cutoffs should be developed

Several members expressed support for focusing these efforts at the benchtop testing stage with the most basic physical units possible (e.g., gravitational force equivalents for accelerometry; precision of wavelength measurement through various media for photoplethysmography). Assuring the validity and quality of these basic units could, in turn, contribute to the evaluation of higher-level measures at later testing stages (e.g., estimated energy expenditure during free-living testing) while still allowing for researchers and companies to pursue innovation. The standardization of these basic units would allow for algorithms that are transferable between devices needing only minor refinements such as a transfer function. For example, Fitbit's formula to convert gravitational force equivalents to estimated energy expenditure could be tested with Apple Watch hardware. It would similarly allow the combination of datasets and interoperability of devices. Overall, these achievements would facilitate detailed, collaborative evaluation of each device at multiple levels, rather than a simplistic confirmation/refutation of the entire device. This process would yield transparency to troubleshoot poor performance and potential cost-savings during development, which would incentivize companies to participate.

The lone author from a large company (Google Health™), however, pointed out that such collaboration may present a competitive advantage conflict for some companies. Thus, they may prefer to have non-standardized basic physical units and hardware-level data smoothing that are proprietary and novel. Furthermore, even those companies interested in having standardized basic units may be unable to comply because they have already completed downstream development around their existing units. Therefore, an alternative strategy was proposed: rather than focusing on the basic physical units (i.e., the earliest possible stage), to look at the other end of the testing pathway spectrum; i.e., analytics on big data generated by CSFWs that are widely

used already (Figure 2). Panelists considered this request and formulated two possible evaluation strategies using these big data despite lacking a ground truth. First, evaluate within-person consistency of known stable variables such as weekly energy expenditure [35] to indicate device reliability and durability. Second, use implemented devices that *do* have benchtop testing validity as convergent validity standards for other devices worn by the same or clinically similar users (i.e., “virtual cohort”). In the case of multiple devices with varying degrees of benchtop testing validity, impose mathematical weighting according to the degree of validity. To utilize the information gained from these evaluations, circumvent the inability to modify the hardware of existing devices by calculating correction factors [38] that are released through a software update or a universal guide for researchers.

4.3. Will clinical applications raise the stakes?

CSFWs can unexpectedly evolve from end user consumer devices into high-risk regulated medical devices. For example, recent evidence suggests the high-risk artificial pancreas systems that automatically titrate insulin delivery (e.g., closed-loop insulin delivery systems) according to continuous glucose monitors benefit from additional input by CSFW smartwatches [36].

Incorporating the CSFW into the marketed high-risk system would require FDA or EU MDR authorization, so we propose CSFWs anticipate this step from their inception by utilizing global guiding standards. In fact even continuous glucose monitors were initially considered an end-user consumer technology product, before the most successful versions passed FDA clearance as medical devices and were incorporated into the standard of care [37].

4.4 Guidance on privacy

Privacy concerns were not considered during the live discussion after being raised during introduction presentations, so participants with privacy expertise from the academic (G.G., D.G., M.B.G.) and industry cohorts (L.A.G.) led a written section presented here. They first recommended helping companies evaluate their CSFWs for the data's level of identifiability and traceability to protected health information. For example, data with higher sampling frequency more readily identify users and predict health outcomes [38, 39]. External assessment tools or consultants could help companies evaluate their data relative to such concerns to inform decisions regarding the tradeoffs of risk versus costs of safeguards. Second, a similar perspective as the clinical applications (Sect. 4.3) emerged: evolving external regulations from bodies like the FDA and GDPR are likely to impact CSFW manufacturers soon, thus heightening the potential benefit of a low-cost pre-evaluation from CSFW guiding standards.

4.5 Guidance on interpreting and presenting CSFW data to consumers

Concerns about data interpretation and presentation to consumers were also not considered during live discussion after being raised during introduction presentations, so participants with behavioral science expertise from the academic (G.I.A., A.E.G., M.S.B., L.M.F., W.R., S.G.) and industry (L.A.G.) cohorts led a written section presented here. They first recommended developing a taxonomy tree of behavior change techniques specific to delivering CSFW data, to standardize and compare techniques across studies. Such taxonomy trees have long existed for older behavior change techniques [40]. Second, guiding standards could include best practice recommendations for devices to deliver CSFW feedback, if the present evidence gaps [26] could be filled and tailored to the user's demographic and health profile. Third, quality assurance

considerations (Sect. 4.2) should inform stated limitations of the data that qualify presentation.

For example, consumers could be advised to consider sleep/wake data more seriously than sleep stages [41]. As with earlier points, participants did not favor enacting legal regulations but rather providing evidence-based recommendations to help companies optimize their CSFW products.

5. POLL OF CSFW IMPROVEMENT PRIORITIES

At the end of the session, we asked attendees to complete a poll, assigning a 1 to 4 priority score to each possible objective of the FIMS central standards. Results revealed that the majority of attendees were most concerned about quality assurance (Table 3). One participant justified this response by noting that “without high quality data none of the other priorities are meaningful”. These sentiments are consistent with the preference to deprioritize big data analytics on devices that have not completed earlier stages of quality testing (Sect. 4.2, paragraph #2). Furthermore, panelists noted that the greatest number of the goals are achieved at the earlier rather than later testing stages, as evident from the concentration of bullet items toward the top versus bottom of Table 2, column 4.

Poll results also indicated marginally higher priority for devices to interoperate with the electronic health record rather than each other (Table 3). However, the former was incidentally achieved by the overall strategy, as evident from its listing in the first row of Table 2, column 4).

Table 3. Results of a poll of consumer sport and fitness wearables improvement priorities.

Median Priority Score	Objective	Definition	Number of Top Priority Votes
#1	Quality assurance	Data accuracy	18 (75%)
#2	Data standardization	Formatting raw units, timescales, coding languages, and storage so that datasets can be more readily shared with other human researchers	5 (21%)
#3	Interoperability of devices with electronic health record	Similar as above but with goal of sharing with medical record	1 (4%)
#4	Interoperability of devices with each other	Similar as above but with goal of combining data and processing algorithms between devices within common software	0 (0%)

Participants were 24 of the 62 attendees from the Yale Center for Biomedical Data Science.

The authors attribute the low response rate to Zoom™ lacking a direct audience poll feature at the time of the panel, forcing additional clicks to reach the poll.

Poll options were set prior to the panel so do not necessarily align with all concerns that motivated the discussion, since some of the latter (privacy, interpreting and presenting CSFW data to consumers) were added by panelists during opening presentations.

6. SUMMARY OF DISCUSSION

Facilitators of industry participation in the global guiding standards were identified and agreed upon by all stakeholders: (1) consumer appeal and satisfaction by increasing the return on investment in device quality; (2) unambiguous targets regarding endpoints and reference standards; (3) lucrative research partnerships; (4) transparent, multilevel evaluation of device quality with specific, constructive criticisms to inform further development; and (5) priming for the more rigorous FDA and EU MDR requirements indicated should CSFWs become part of regulated medical devices. These facilitators (especially #4) can be best exploited if the guiding standards prioritize the benchtop stage of testing.

Benchtop testing was the stage most affected by the barriers to industry participation that were identified: competitive advantage conflict and lack of flexibility in previously developed devices.

These barriers are heavily pertinent to the benchtop stage of testing because it focuses upon basic physical units that are often proprietary. These barriers were all noted by the representative from a large manufacturer (Google Health™) rather than the small one (Xsensio™), suggesting they may be most pertinent to larger companies market-wide.

7. RECOMMENDED ADDITIONS TO EXISTING EFFORTS

Several other international working groups have begun assembling knowledge that could address concerns in the CSFW field. The CTA™ has standard guidelines for testing protocols and performance criteria of CSFWs, including those that measure energy expenditure, heart rate, step counting, sleep, and stress indicators such as autonomic function [42]. These guidelines were developed by panels of experts (vendors, regulators, other industry leaders), to establish a common understanding that sets a foundation for the industry to develop. In the case of step counting and heart rate, the Towards Intelligent Health and Well-Being Network of Physical Activity Assessment (INTERLIVE) consortium has refined guidelines via expert panel discussion supported by a systematic literature review of existing validation protocols and possible sources of bias [43, 44]. Turning from quality assurance to data standardization, the Personal Connected Health Alliance (PCHA) Continua Design Guidelines [45] and the Institute of Electrical and Electronics Engineers (IEEE) P1752 Open Mobile Health Working Group [46] have specifications and open-source codes for standardization of mobile health data.

These protocols are targeted to make a large dent in concerns about CSFWs. By focusing on metrics that account for the largest share of CSFW sales (step counting, heart rate) and validation among the general population without chronic disease under controlled laboratory conditions,

they have maximized the coverage that can be attained from concise published documents.

Nonetheless, such documents cannot fully respond to needs created by unique aspects of the CSFW field. Specifically, there is a need to accommodate:

- 1) Heterogeneity of devices, data types, and contexts of data collection (Sect. 1)
- 2) Engagement of companies that is pragmatic and appealing for them (Sect. 4.1)
- 3) Development in multiple tiers that must be unified (Sect. 4.2)
- 4) Constructive and low-cost guidance while products are in early development stages and/or have not yet reached a space of regulation by FDA, EU MDR, or GDPR (Sect. 4.2-4.4). When it comes time to engage those regulators, fees are much higher, and the only feedback is a binary approval or disapproval.
- 5) Nebulous areas that lack even basic taxonomic standardization, such as best practices for CSFW data presentation to consumers (Sect. 4.5)

These nuances would be overwhelming to capture in stagnant published documents. We propose instead a dynamic networking “hub” that connects companies with consumers and researchers. The hub would amalgamate input from consumers and researchers about desired standards for CSFW quality and data formats, then provide companies with networking introductions to prospective colleagues that have the needed resources. These resources may include laboratories that have essential experts (e.g., a pharmacologist to advise on developing photoplethysmography that works for users taking beta blockers), the needed overhead resources so companies would only support incremental costs, or clinical trials that could add additional devices and surveys for validation sub-studies. For example, a Yale clinical trial recruited a cohort from a hard-to-reach population (heavy-drinking young adults) to wear reference standard

alcohol sensor devices for two weeks at a cost of US \$722,000 but added an experimental alcohol sensor for a supplement cost of just \$31,000 [47]. Other similar trials are available from other panelists [48-50]. Connections made in a networking hub would pass these savings to CSFW companies.

8. CONCLUSIONS AND RECOMMENDED NEXT STEPS

Potential strategies to develop the hub include networking events, peer-reviewed journals for the resulting studies, webinars and consultations providing companies with a needs assessment and initial networking introductions.

We also considered how to start eliciting company buy-in. Altogether, our panel discussion revealed a disconnect between optimizing the full potential of a global guiding reference (benchtop testing, for large and small companies) and more immediately achievable steps (field-based and implementation testing, for forthcoming small companies). Therefore, the panel recommended that immediate future endeavors should prioritize field-based testing with forthcoming small manufacturers, to subsequently attract larger manufacturers and begin to offer benchtop testing.

Finally, we acknowledged there is uncertainty over which endpoints should underlie the testing we have outlined (Table 2). Therefore, there is a need to meta-analyze the literature for the CSFW endpoints to examine in future testing that are most clinically relevant (i.e., surrogate endpoints) [51] and grounded (e.g., pressure-sensing treadmill to validate foot-worn inertial sensors), leading to a white paper with input from academic and industry stakeholders. These

would parallel the efforts of INTERLIVE [43, 44] but extend beyond step counting and heart rate to other outputs such as those enumerated in Table 1. Overall, we recommend that guiding standards for CSFWs provide companies feedback that is 1) constructive, 2) minimal cost, and 3) facilitates flexibility in future directions. Companies should use the standards to the extent that they find this feedback beneficial; the aim is not to introduce mandatory regulatory costs. We envision a non-profit venture that benefits companies, researchers, and consumers.

DECLARATIONS

Conflicts of interest/Competing interests

Dr. Robert Huggins is currently employed by the Korey Stringer Institute who is a 501(c)3 not for profit organization with corporate partners that support the mission of the institute. These partners include the National Football League, Gatorade, the National Athletic Trainers' Association, Mission Athletecare, Kestrel by Neilsen Kellerman, Eagle Pharmaceuticals, and DeFibtech. These entities provided no financial support, other support, or other influence toward the manuscript.

Dr. Stuart Weinzimer has received honoraria for serving as Speaker and/or Consultant for Medtronic, Insulet, and Tandem, manufacturers of diabetes technologies that are relevant to the subject of the manuscript; these commercial entities were not in any manner involved with the research, preparation, or review of the manuscript.

Mr. Robert Jarrin has been compensated as a strategic advisor by the CTA, MiCare Path (consulting fees or honorarium), and Strive Orthopedics, Inc. (stock/stock options). In addition, he serves as Member/Advisor to the American Medical Association (AMA) Digital Medicine Payment Advisory Group (DMPAG).

Drs. Garrett Ash, Matthew Stults-Kolehmainen, Michael Busa, Allison Gaffey, Mr. Konstantinos Angeloudis, Drs. Borja Muniz-Pardos, Robert Gregory, Nancy Redeker, Lauren Grieco, Kate Lyden, Ms. Esmeralda Megally, Dr. Ioannis Vogiatzis, Ms. LaurieAnn Scher, Drs. Xinxin Zhu, Julien Baker, Cynthia Brandt, Michael Businelle, Lisa Fucito, Stephanie Griggs, Bobak Mortazavi, Temiloluwa Prioleau, Walter Roberts, Elias Spanakis, Laura Nally, Andre Debruyne, Norbert Bachl, Fabio Pigozzi, Farzin Halabchi, Dimakatso Ramagole, Dina Janse van Rensburg, Bernd Wolfarth, Chiara Fossati, Sandra Rozenstoka, Kumpei Tanisawa, Mats Börjesson, José Casajus, Alex Gonzalez-Aguero, Irina Zelenkova, Jeroen Swart, Gamze Gursoy, William Meyerson, Mr. Jason Liu, Drs. Dov Greenbaum, Yannis Pitsiladis and Mark Gerstein declare that they have no conflicts of interest relevant to the content of this article.

Funding

Dr. Garrett Ash was supported by a fellowship from the Office of Academic Affiliations at the United States Veterans Health Administration. Dr. Elias Spanakis was partially supported by the VA MERIT award (#1I01CX001825) from the United States Department of Veterans Affairs Clinical Sciences Research and Development Service. Dr. Allison Gaffey was supported by a research grant from the National Institutes of Health (R01HL126770). Dr. Stephanie Griggs was supported by mentored research scientist awards from the National Institutes of Health (K99NR018886) and the American Academy of Sleep Medicine (220-BS-19). Dr. Walter Roberts (K23AA026890), Dr. Laura Nally (K12DK094714-10), and Dr. Gamse Gursoy (K99HG010909) were supported by mentored research scientist awards from the National Institutes of Health. No other sources of funding were used to assist in the preparation of this manuscript.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Author contributions

The first draft of the manuscript was written by Garrett Ash and Yannis Pitsiladis. All authors commented on subsequent versions of the manuscript until all authors were able to approve the final manuscript.

Data Availability

The data are the transcription of the session recordings, available from author Garrett Ash (<https://orcid.org/0000-0002-8655-7525>, garrett.ash@yale.edu) and permitted for reuse with his permission.

REFERENCES

1. How many IoT devices are there in 2021? Techjury. <https://techjury.net/blog/how-many-iot-devices-are-there/#gref>. 2021. Accessed June 18, 2021.
2. European Commission. Smart wearables: Reflection and orientation paper including feedback from stakeholders. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50020. 2017. Accessed June 18, 2021.
3. Smartwatch shipments forecast worldwide from 2016 to 2025(in millions). Statista. 2021. <https://www.statista.com/statistics/878144/worldwide-smart-wristwear-shipments-forecast/>. Accessed June 18, 2021.
4. Peake JM, Kerr G, Sullivan JP. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Front Physiol*. 2018 Jun 28;9:743.
5. Gaffey AE, Jeon S, Conley S, Jacoby D, Ash GI, Yaggi HK, et al. Perceived stress, subjective, and objective symptoms of disturbed sleep in men and women with stable heart failure. *Behav Sleep Med*. 2021 May-Jun;19(3):363-77.
6. Ash GI, Stults-Kolehmainen M, Busa MA, Gregory R, Garber CE, Liu J, et al. Establishing a global standard for wearable devices in sport and fitness: Perspectives from the new england chapter of the american college of sports medicine members. *Curr Sports Med Rep*. 2020 Feb;19(2):45-9.
7. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. *J Sports Sci Med*. 2016 Aug 5;15(3):540-7.
8. Landers J, Fitbit I. Case no. 16-cv-00777-JD. N D Cal. 2016 November 14, 2016.
9. McLellan K, Fitbit I. Case no.16-cv-00036-JD. N D Cal. 2018 July 24, 2018.
10. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: Systematic review. *JMIR Mhealth Uhealth*. 2020 Sep 8;8(9):e18694.
11. Murakami H, Kawakami R, Nakae S, Yamada Y, Nakata Y, Ohkawara K, et al. Accuracy of 12 wearable devices for estimating physical activity energy expenditure using a metabolic chamber and the doubly labeled water method: Validation study. *JMIR Mhealth Uhealth*. 2019 Aug 2;7(8):e13938.
12. Pirker W, Katzenschlager R. Gait disorders in adults and the elderly : A clinical guide. *Wien Klin Wochenschr*. 2017 Feb;129(3-4):81-95.

13. Duking P, Stammel C, Sperlich B, Sutehall S, Muniz-Pardos B, Lima G, et al. Necessary steps to accelerate the integration of wearable sensors into recreation and competitive sports. *Curr Sports Med Rep*. 2018 Jun;17(6):178-82.
14. Duking P, Fuss FK, Holmberg HC, Sperlich B. Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity. *JMIR Mhealth Uhealth*. 2018 Apr 30;6(4):e102.
15. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and security in mobile health: A research agenda. *Computer (Long Beach Calif)*. 2016 Jun;49(6):22-30.
16. Galvin HK, DeMuro PR. Developments in privacy and data ownership in mobile health technologies, 2016-2019. *Yearb Med Inform*. 2020 Aug;29(1):32-43.
17. Wellness programs raise privacy concerns over health data. The Society for Human Resources Management. 2016. <https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/wellness-programs-raise-privacy-concerns-over-health-data.aspx>. Accessed June 18, 2021.
18. Raber I, McCarthy CP, Yeh RW. Health insurance and mobile health devices: Opportunities and concerns. *JAMA*. 2019 May 14;321(18):1767-8.
19. Conley S, Knies A, Batten J, Ash GI, Miner B, Hwang Y, et al. Agreement between actigraphic and polysomnographic measures of sleep in adults with and without chronic conditions: A systematic review and meta-analysis. *Sleep Med Rev*. 2019 Aug;46:151-160.
20. The strava heat map and the end of secrets. 2018. <https://www.wired.com/story/strava-heat-map-military-bases-fitness-trackers-privacy/>. Accessed June 18, 2021.
21. Protecting patient privacy and security while exploiting the utility of next generation digital health wearables. 2019. <https://blogs.bmj.com/bmj/2019/01/18/protecting-patient-privacy-and-security-while-exploiting-the-utility-of-next-generation-digital-health-wearables/>. Accessed June 18, 2021.
22. Mulder T. Health Apps, their Privacy Policies and the GDPR. *Eur J Law Tech* 2019;10(1). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3506805. Accessed December 29, 2020.
23. Baron KG, Abbott S, Jao N, Manalo N, Mullen R. Orthosomnia: Are some patients taking the quantified self too far? *J Clin Sleep Med*. 2017 Feb 15;13(2):351-4.
24. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAs) in mobile health: Key components and design principles for ongoing health behavior support. *Ann Behav Med*. 2018 May 18;52(6):446-62.

25. Berryhill S, Morton CJ, Dean A, Berryhill A, Provencio-Dean N, Patel SI, et al. Effect of wearables on sleep in healthy individuals: A randomized crossover trial and validation study. *J Clin Sleep Med*. 2020 May 15;16(5):775-83.
26. Shin G, Jarrahi MH, Fei Y, Karami A, Gafinowitz N, Byun A, et al. Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review. *J Biomed Inform*. 2019 May;93:103153.
27. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D36-42.
28. United States Food and Drug Administration. FDA Launches the Digital Health Center of Excellence. 2020. <https://www.fda.gov/news-events/press-announcements/fda-launches-digital-health-center-excellence>. Accessed June 18, 2021.
29. United States Food and Drug Administration. General Wellness: Policy for Low-Risk Devices. Guidance for Industry and Food and Drug Administration Staff. 2019. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices>. Accessed June 18, 2021.
30. Ravizza A, De Maria C, Di Pietro L, Sternini F, Audenino AL, Bignardi C. Comprehensive review on current and future regulatory requirements on wearable sensors in preclinical and clinical testing. *Front Bioeng Biotechnol*. 2019 Nov 8;7:313.
31. Spinner, J. Trelleborg helps wearables firms navigate new EU regulation. 2021. <https://www.outsourcing-pharma.com/Article/2021/05/25/Trelleborg-helps-wearables-firms-navigate-new-EU-regulation>. Accessed June 18, 2021.
32. Are wearables medical devices requiring a CE-mark in the EU? Covington Digital Health. 2019. <https://www.covingtondigitalhealth.com/2019/01/are-wearables-medical-devices-requiring-a-ce-mark-in-the-eu/>. Accessed June 18, 2021.
33. Esliger DW, Copeland JL, Barnes JD, Tremblay MS. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *Journal of Physical Activity and Health*. 2005;2(3):366-383.
34. Keadle SK, Lyden KA, Strath SJ, Staudenmayer JW, Freedson PS. A framework to evaluate devices that assess physical behavior. *Exerc Sport Sci Rev*. 2019 Oct;47(4):206-14.
35. Dillon CB, Fitzgerald AP, Kearney PM, Perry IJ, Rennie KL, Kozarski R, et al. Number of days required to estimate habitual activity using wrist-worn GENEActiv accelerometer: A cross-sectional study. *PLoS One*. 2016 May 5;11(5):e0109913.
36. Jacobs PG, Resalat N, El Youssef J, Reddy R, Branigan D, Preiser N, et al. Incorporating an exercise detection, grading, and hormone dosing algorithm into the artificial pancreas using accelerometry and heart rate. *J Diabetes Sci Technol*. 2015 Oct 5;9(6):1175-84.

37. American Diabetes Association. 7. diabetes technology: Standards of medical care in diabetes-2021. *Diabetes Care*. 2021 Jan;44(Suppl 1):S85-99.
38. Jacobson NC, Lekkas D, Huang R, Thomas N. Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years. *J Affect Disord*. 2021 Mar 1;282:104-11.
39. Perakslis E, Coravos A. Is health-care data the new blood? *Lancet Dig Health*. 2019 May;1:e8-9.
40. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: The development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). *Health Technol Assess*. 2015 Nov;19(99):1-188.
41. Stone JD, Rentz LE, Forsey J, Ramadan J, Markwald RR, Finomore VS, et al. Evaluations of commercial sleep technologies for objective monitoring during routine sleeping conditions. *Nat Sci Sleep*. 2020 Oct 27;12:821-42.
42. Consumer Technology Association Standards. <https://shop.cta.tech/collections/standards>. Accessed June 18, 2021.
43. Johnston W, Judice PB, Molina García P, Mühlen JM, Lykke Skovgaard E, Stang J, et al. Recommendations for determining the validity of consumer wearable and smartphone step count: Expert statement and checklist of the INTERLIVE network. *Br J Sports Med*. 2020 Dec 24;bjsports-2020-103147. Online ahead of print.
44. Mühlen JM, Stang J, Lykke Skovgaard E, Judice PB, Molina-Garcia P, Johnston W, et al. Recommendations for determining the validity of consumer wearable heart rate devices: Expert statement and checklist of the INTERLIVE network. *Br J Sports Med*. 2021 Jan 4. bjsports-2020-103148. Online ahead of print.
45. Continua Design Guidelines. Personal Connected Health Alliance. 2019. <https://www.pchalliance.org/continua-design-guidelines>. Accessed June 18, 2021.
46. IEEE P1752 Open Mobile Health Working Group. 2020. <https://sagroups.ieee.org/1752/> Accessed June 18, 2021.
47. Fucito LM, Ash GI, DeMartini KS, Pittman B, Barnett NP, Li CR, et al. A multimodal mobile sleep intervention for young adults engaged in risky drinking: Protocol for a randomized controlled trial. *JMIR Res Protoc*. 2021 Feb 26;10(2):e26557.

48. Griggs S, Redeker NS, Crawford SL, Grey M. Sleep, self-management, neurocognitive function, and glycemia in emerging adults with type 1 diabetes mellitus: A research protocol. *Res Nurs Health*. 2020 Aug;43(4):317-28.

49. Ash GI, Nally LM, Stults-Kolehmainen M, De Los Santos M, Jeon S, Brandt C, et al. Personalized big data for type 1 diabetes exercise support. *SportRxiv*. 2021. 34vdc [Preprint]. <https://doi.org/10.31236/osf.io/34vdc>. Accessed June 18, 2021.

50. Singh LG, Satyarengga M, Marcano I, Scott WH, Pinault LF, Feng Z, et al. Reducing inpatient hypoglycemia in the general wards using real-time continuous glucose monitoring: The glucose telemetry system, a randomized clinical trial. *Diabetes Care*. 2020 Nov;43(11):2736-43.

51. Puente-Maestu L, Palange P, Casaburi R, Laveneziana P, Maltais F, Neder JA, et al. Use of exercise testing in the evaluation of interventional efficacy: An official ERS statement. *Eur Respir J*. 2016 Feb;47(2):429-60.

FIGURE CAPTIONS

Fig 1 Graphical representation of the step sequence in people with and without classical gait disorders [reprinted from [12], Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>)]

Fig 2 Discussion of where to focus testing efforts, based upon Keadle et al.'s standard testing pathway for wearable technology [34]