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 Angeloudis[8], Borja Muniz-Pardos[9], Robert Gregory[10], Robert A. Huggins[11], Nancy S. Redeker[12], Stuart A. Weinzimer[13], Lauren A. Grieco[14], Kate Lyden[15], Esmeralda Megally[16], Ioannis Vogiatzis[17], LaurieAnn Scher[18,19], Xinxin Zhu[20], Julien S. Baker[21-23], Cynthia Brandt[1,2,24], Michael S. Businelle[25,26], Lisa M. Fucito[27-29], Stephanie Griggs[30], Robert Jarrin[31,32], Bobak J. Mortazavi[33], Temiloluwa Prioleau[34], Walter Roberts[27], Elias K. Spanakis[35,36], Andre Debruyne[37,38], Norbert Bachl[37-40], Fabio Pigozzi[37-38, 41-42], Farzin Halabchi[37,43-46], Dimakatso A. Ramagole[37,47], Dina C Janse van Rensburg[37,47], Bernd Wolfarth[37,48], Chiara Fossati[41], Sandra Rozenstoka[37,38,49], Kumpei Tanisawa[50], Mats Börjesson[37,51,52], José Antonio Casajus[9,37], Alex Gonzalez-Aguero[9,37], Irina Zelenkova[9,53], Jeroen Swart[37,54], Gamze Gursoy[55], William Meyerson[56,57], Jason Liu[55], Dov Greenbaum[55,58,59], Yannis P. Pitsiladis[8,37,38], Mark B. Gerstein[55,57,60,61]

1. Veterans Affairs Connecticut Healthcare System, West Haven, Connecticut, USA
2. Center for Medical Informatics, Yale University, New Haven, Connecticut, USA
3. Bariatric Surgery Program, 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. Consumer Technology Association, 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. Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran (MHPS)

45. Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran (MAM)

46. Department of Sports and Exercise Medicine, Tehran University of Medical Sciences, Tehran, Iran (FH)

47. Section Sports Medicine, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa.

48. Department of Sports Medicine, Humboldt University and Charité University School of Medicine, Berlin, Germany.

49. FIMS Collaboration Centre of Sports Medicine, Sports laboratory, Riga, Latvia.

50. Faculty of Sport Sciences, Waseda University, Tokorozawa, Japan.

51. Center for Health and Performance, Dept of Molecular and Clinical Medicine, Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden.

52. Sahlgrenska University Hospital, Dept og MGA, Region of Western Sweden, Gothenburg, Sweden.

53. I.M. Sechenov First Moscow State Medical University (Sechenov University), Ministry of Health of Russia, Moscow, Russia

54. Division of Exercise Science and Sports Medicine, University of Cape Town, Cape Town, South Africa

55. Program in Computational Biology & Bioinformatics, Yale University, New Haven, Connecticut, USA

56. Duke Psychiatry & Behavioral Sciences, Duke Medicine, Durham, North Carolina, USA

57. Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, Connecticut, USA

58. Zvi Meitar Institute for Legal Implications of Emerging Technologies, Interdisciplinary Center Herzliya, Herzliya, Israel

59. Harry Radyzner Law School, Interdisciplinary Center Herzliya, Herzliya, Israel

60. Department of Computer Science, Yale University, New Haven, Connecticut, USA

61. 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

The International Federation of Sports Medicine (FIMS) in association also with the European Federation of Sports Medicine Associations (EFSMA) seeks to establish a central resource at accredited laboratories to evaluate consumer sport and fitness wearables (CSFWs) for quality and data standardization. This will guide companies to achieve these aspects and educating consumers to critically consider them. A virtual panel was convened for formative discussions among industry and academic stakeholders regarding: (1) key facilitators and barriers to participation by CSFW manufacturers 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). This event served as part of the consultation process for FIMS to refine its roadmap towards full implementation of a global standard for wearables in sport and fitness. 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 adopting the global standard. There was general agreement to adopt Keadle et al.'s (2019) standard pathway for testing devices (i.e., benchtop, laboratory, field-based, implementation) without consensus on the prioritization of these steps. In conclusion, the panel identified facilitators to industry participation (e.g., added value to commercial return on investment and constructive critiques), and barriers, that were especially palpable for larger companies (e.g., inability to modify marketed devices at a benchtop level). An implementation roadmap was recommended that prioritized field-based testing with forthcoming small manufacturers, with the goal of subsequently attracting larger manufacturers and beginning to offer benchtop testing.

KEY POINTS

- The International Federation of Sports Medicine and the European Federation of Sports Medicine Associations seek to establish a central resource at accredited laboratories to evaluate consumer sport and fitness wearables for quality and data standardization.
- Stakeholders agree the resource could add value to commercial return on investment and provide constructive critiques to manufacturers, especially when quality and standardization procedures focus 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., batch effect corrections).

1. INTRODUCTION TO CONSUMER SPORT AND FITNESS WEARABLES

Scientific advances over the past 50 years have supported the evolution of wearable technology: the application of small, light-weight sensors to free-living conditions. These novel devices can be worn on the human body, inside vital organs (i.e., ingestible core body temperature sensor) or even mounted on sporting equipment such as skis, shoes, or clothing. Consumer-grade sport and fitness wearables (CSFWs) include devices that can measure position, motion, location, biomechanics (e.g., foot-worn inertial sensors), heart rate and blood and muscle tissue oxygen saturation, sweat composition and sweat lactate concentration, galvanic skin response, body temperature, autonomic function, and sleep. These portable sensors collect a wide range and volume of kinetic, kinematic, mechanical and bioenergetic data, and analyze them by interfacing with physical and server-based computers. An increasing number of physicians, sport scientists and other employees within international sports and medical federations, at rehabilitation centers, sports clubs, and sporting events use some form of CSFWs. Across professional and leisure contexts, CSFWs comprise a $19 billion industry worldwide [1].

2. MAJOR CONCERNS: 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 [2], resulting in two class action lawsuits [3, 4]. More recently, Peake and colleagues evaluated 61 wearables and found that only 5% matched their marketing claims according to accepted reference standards [5]. 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 [6]. A primary study of energy expenditure from 4 of these sensors and 8 others worn simultaneously by 19 adults drew a similar conclusion [7]. Moreover, variable gait patterns [8] (Figure 1) suggest the need for population-specific validations, which are currently lacking. Another concern is Lastly, these studies likely use new or well-maintained devices thus ignoring possible durability and long-term calibration concerns. Because the device market evolves rapidly, quickly outpacing this and other related research, this market necessitates fast and frequent au courant 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 [9]. 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 [10]. 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 to deny coverage for individuals exhibiting high-risk behavior patterns (e.g., poor sleep, physical inactivity) [11, 12]. Also, these decisions may be based on inaccurate data (e.g., periods of restful wakefulness may be interpreted as sleep) [13]. 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) [14]. 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 [15]. Many companies claim that data they share with outside parties are deidentified, but the United States Health Insurance Portability and Accountability Act 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 [16].

The best practices for interpreting and presenting CSFW data to consumers remain unclear and controversial. For instance, sleep watch data can harm consumers by eliciting "preoccupation or concern with improving or perfecting wearable sleep data" instead of accepting medical advice, standard sleep hygiene education, or validated laboratory sleep assessments [17]. Some research has addressed this problem by optimizing the timing of data presentation (i.e., just-in-time adaptive interventions) [18]. 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 [19]. 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 articles 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 [20].

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 [21]. This 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.

3. BODIES THAT COULD ADDRESS CONCERNS ASSOCIATED WITH CSFWs

The United States Food and Drug Administration (FDA, Washington, DC) 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 [22]. 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) [23]. 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.

Several international working groups have begun assembling knowledge that could address concerns in the CSFW field. The Consumer Technology Association (CTA, Washington, DC) 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 [24]. 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, 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 [25]. Turning from quality assurance to data standardization, the Personal Connected Health Alliance (PCHA) Continua Design Guidelines [26] and the Institute of Electrical and Electronics Engineers (IEEE) P1752 Open Mobile Health

Working Group [27] have specifications and open-source codes for standardization of mobile health data. The CTA, INTERLIVE, PCHA, and IEEE standards require a practical plan for practical refinement (e.g., synergistically testing multiple outcomes to improve workflow), protocols for keeping abreast of field developments apart from expert opinion (e.g., INTERLIVE), utilization, and implementation.

The International Federation of Sports Medicine (FIMS, Lausanne, Switzerland) in association also with the European Federation of Sports Medicine Associations (EFSMA) aims 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. FIMS advocates for both the consumers of CSFWs and the sports medicine researchers making use of their data. Both groups depend upon an essential third pillar: manufacturers of CSFWs. Therefore, FIMS aims to develop a novel solution that integrates the needs of all three entities. After several years of stakeholder discussions [28, 29], FIMS has determined that the best solution is to establish a central resource at accredited laboratories to evaluate CSFWs for quality and/or data standardization, thus, guiding companies to achieve these metrics, educating consumers to critically consider them, and creating a unique database with the evaluation results.

It is rare to find initiatives in any context that integrate the needs of two of these groups (consumers, manufacturers, researchers), let alone all three simultaneously. With CSFWs, it is essential to meet the mutual needs of all three constituents because they are growing rapidly in capacity: (1) Consumers of CSFWs increased from 350 to 441 million worldwide in 2020 [1]

and their usage of CSFWs is generating enormous databases. For example, the 100+ billion hours of heart rate data assembled by Fitbit to address questions such as the age dependence of resting heart rate [30]; (2) Manufacturers are continually expanding the scope of the variables their devices can measure, across biomechanical, biochemical, biophysical, and biobehavioral domains [31]; (3) Researchers recognize that these big datasets will contribute to the biological and behavioral phenotyping of individuals and populations, analyzing the relationships among these outputs and other big data sources such as genomics, and ultimately developing personalized interventions (e.g., recommended bedtime for optimal daytime performance) only if the data are accurate [32]. Therefore, it is essential to incorporate the needs of all three stakeholder groups into the new FIMS central resource. Yet, this novel undertaking has brought novel challenges.

4. PROGRESS TO DATE AND CHALLENGES

In January 2019, FIMS established a task force that began meeting to address the need for a quality assurance central resource among wearable devices. The first FIMS Collaborating Centre of Sports Medicine was established in the Growth, Exercise, Nutrition and Development (GENUD) Research Group at the University of Zaragoza (Zaragoza, Spain), which hosted the initiative in September 2019. The multidisciplinary GENUD Lab receives national and European Union funding to design and implement interventions that combine a nutritional-physical activity-psychological approach. Specifically, the GENUD group is formed by experts in body composition and functional capacity in a wide variety of populations, and has a long-standing record of performing clinical and public health investigations in collaboration with medical doctors, nurses, dieticians, and sports scientists. GENUD also has extensive experience with

method validation of wearable technologies (e.g., camera-based systems to measure movement velocity, accelerometers, brain stimulation wearables, and foot-wear inertial sensors) focusing on body composition, physical activity, and athletic performance assessment in both trained and sedentary children, adolescents, adults, and elderly individuals.

FIMS will also appoint additional testing centers. Current leading candidates include the University of Massachusetts Institute for Applied Life Sciences, the Hong Kong Baptist University Obesity Comorbidities Center, and the Yale University School of Medicine. The first one is specifically instrumented for testing wearable devices at various stages of device evolution including "core facilities for projects ranging from device prototyping, precision manufacturing and roll-to-roll fabrication, to human motion and gait studies, calorimetry, magnetic resonance imaging and spectroscopy, as well as, EEG and sleep studies" [33]. The second one regularly utilizes wearables for telemedicine studies and can also measure a wide range of obesity-related phenotypes to provide criterion (i.e., ground reaction forces) and construct validity (e.g., body composition, agility, and coordination). The third one has cloud computing management, information security, and behavioral science expertise to develop digital behavioral interventions (clinical trials for pilot, efficacy, effectiveness testing, implementation science) that incorporate CSFWs aiming to optimize their adoption, acceptance, and health impact. Authors from this institution (G.A., M.S., C.B., L.F., S.W., W.M., W.R., J.L., M.G.) and their collaborators (J.B, L.S., T.P., S.G., E.S., B.M.) are already developing two such interventions that use sensors to provide users with feedback about multiple health behaviors, their potential intersection, and responsive lifestyle guidance [34, 35]. This work responds to the concerns with best practices for interpreting and presenting CSFW data to consumers discussed earlier in the "major concerns"

section. FIMS will solicit additional centers through its partnership with WT - Wearable Technologies, an innovation and market development platform business supporting wearables manufacturers globally [31]. These plans for multicenter, intercontinental collaboration will establish inter-center reliability to prevent bias, productive writing groups for peer reviewed publications, and global meetings to disseminate findings and endorsements. GENUD is currently establishing protocols and standard operating procedures for validity/reliability testing and certification of CSFWs.

A full-time project coordinator has been deployed by FIMS to support the GENUD lab and will be assisted by staff and students as part of their normal scholarly activities. GENUD is providing specialist facilities and equipment. Other costs will be met from the testing fee passed on to the manufacturer of the wearable device(s) being evaluated. The fee will be determined on a case-by-case basis but it is not expected to exceed 10,000 € per assessment.

During Autumn 2020, FIMS conducted a consultation process to refine the next steps in its roadmap toward full implementation of the central resource. This included a virtual panel for formative discussions among industry and academic stakeholders regarding: (1) key facilitators and barriers to participation by CSFW manufacturers; and (2) stakeholder priorities. Venues were the Yale Center for Biomedical Data Science Digital Health Monthly Seminar Series and the New England Chapter of the American College of Sports Medicine Annual Meeting. By including both industry and academic stakeholders, the panel built upon its similar previous event in 2019 that only included academic stakeholders [29].

5. PANEL LOGISTICS AND RECRUITMENT

The panel 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). To fill the panelist spaces, we executed a recruitment strategy focused on attracting a mixture of large and small international and national companies. Invitations were sent electronically to the public relations departments and/or personal contacts within 6 large and 4 small companies 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 planning 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 the central resource 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 evaluate the role of CSFWs to develop large research initiatives. The third space was filled by the CTA (represented by author L.S.).

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 42 have made substantive contributions and were invited to coauthor this manuscript (24 kinesiologists, 9 data scientists, 2 endocrinologists, 3 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

253 authors commented on subsequent versions of the manuscript until all authors were able to
254 approve the final manuscript.
255
256 **6. DISCUSSION TOPICS**
257 **6.1. What could incentivize industry stakeholders to engage with the quality assurance and**
258 **data standardization central resource?**
259 Individuals from both manufacturers in attendance (Google Health, Xsensio) were supportive of
260 the FIMS central resource and expressed interest in joining. When these individuals were asked
261 what incentivized them to join the panel, two themes emerged. The first theme was value with
262 respect to consumer appeal and satisfaction. Third-party endorsement provided by the central
263 resource could help them dispel stereotypes about poor quality of CSFWs created by
264 controversies such as the Fitbit class action lawsuits [3, 4]. Also, user education provided by the
265 central resource would promote the more discerning selection of CSFWs and potentially foster a
266 greater appreciation of CSFWs that offer high validity, quality, and useful data. This education
267 would increase the commercial value yielded by their development efforts. For example, if users
268 know that they should expect a heart rate sensor to have <5% error based on a reference
269 standard, it will increase the return on investment for the development needed to reach that
270 standard; otherwise, the sensor with 5% error has no greater market value than one with 10%
271 error.
272
273 The second theme that emerged from the panel discussion was value with respect to scientific
274 endeavors. The two manufacturer panelists stated an interest to participate in data mining
275 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 [36]. 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 devices 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, it offers 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.

6.2. 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) [37], the panel debated which of these steps should be the focus of FIMS validation checks, quality assurance procedures, and standardization

of data outputs. Several members expressed support for focusing these efforts at the benchtop testing stage and for using the most basic physical units possible (e.g., gravitational force equivalents in benchtop testing of accelerometers). Assuring the validity and quality of these basic units could, in turn, contribute to the validity and quality of higher-level measures at later testing stages (e.g., estimated energy expenditure during free-living testing). Meanwhile, standardizing the data output from these basic units, could create algorithms that convert the units to higher-level measures portable between devices. 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 datasets to be combined and devices to interoperate. 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 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). An example is batch effect correction: machine-learned rules could scale

data between different devices worn by the same users (or populations of similar users) to a common consensus metric under which all devices report the same mean and variance for the same properties under similar conditions [38]. Existing devices could then incorporate these rules as a software update or a universal guide for researchers.

6.3. Will clinical applications raise the stakes?

CSFWs are originally conceptualized as end-user consumer devices, yet sometimes through unwitting marketing claims manufacturers may unexpectedly transform their products into regulated medical devices, as opposed to high-risk medical devices intended from early development to be regulated. An example is a high-risk medical device such as the artificial pancreas systems (e.g., closed-loop insulin delivery systems) for people with diabetes. Currently marketed systems infuse insulin in response to elevated blood glucose levels, as detected by a continuous glucose monitor, resulting in automated blood glucose stabilization. However, this stabilization improves when incorporating physical activity data (as detected by a CSFW accelerometer that is not necessarily a medical device) according to recent clinical trials [39]. Therefore, a marketed artificial pancreas system that is regulated may need to seek further FDA authorization for updated versions that incorporate data from a non-regulated CSFW accelerometer. In this instance, the FDA would intervene to assess safe and effective functioning of the artificial pancreas system that included a low-risk non-regulated accelerometer. Arguably, an accelerometer previously guided by FIMS central standards (e.g., holding quality assurance at the forefront of their development process) would be better positioned to embark upon the process of meeting FDA standards. A similar process previously occurred for continuous glucose monitors; the monitors were initially considered an end-user consumer technology product, but

the most successful versions passed FDA clearance as medical devices and were incorporated into the standard of care [40].

7. 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 1). 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 (see discussion topic #1, paragraph #2).

Table 1. Poll results. Participants were 24 of the 62 attendees from the Yale Center for Biomedical Data Science.

Median Priority Score	Objective	Number of Top Priority Votes
#1	Quality Assurance	18 (75%)
#2	Data standardization	5 (21%)
#3	Interoperability of devices with electronic health record	1 (4%)
#4	Interoperability of devices with each other	0 (0%)

8. CONCLUSIONS

Facilitators of industry participation in the FIMS central resource 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 requirements indicated should CSFWs become part of regulated medical devices. These facilitators (especially #4) can be best exploited if the central resource prioritizes 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.

9. NEXT STEPS

Altogether, FIMS recognizes a disconnect between the roadmap to optimizing the full potential of the central resource (benchtop testing, for large and small companies) and the more immediately achievable steps (field-based and implementation testing, for forthcoming small companies). Thus, an implementation roadmap was recommended (Figure 3), in which panel attendees prioritized field-based testing with forthcoming small manufacturers in the first instance, with the goal of subsequently attracting larger manufacturers with benchtop testing. We will also meta-analyze the literature for the CSFW endpoints to examine in future testing that are most clinically relevant (i.e., surrogate endpoints) [41] and grounded (e.g., pressure-sending treadmill to validate foot-worn inertial sensors), leading to a white paper with input from

academic and industry stakeholders. The roadmap steps have a relatively short timeframe compared to the longer timeframe of full regulatory processes (e.g., FDA); this efficiency is attributable to the outstanding in-kind resources provided by the GENUD research group.

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 who has 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).

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) 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. Wearables - worldwide. Statista. https://www.statista.com/outlook/319/100/wearables/worldwide2020. Accessed December 29, 2020.

2. 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.

3. Landers J, Fitbit I. Case no. 16-cv-00777-JD. N D Cal. 2016 November 14, 2016.

4. McLellan K, Fitbit I. Case no.16-cv-00036-JD. N D Cal. 2018 July 24, 2018.

5. 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.

6. 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.

7. 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.

8. Pirker W, Katzenschlager R. Gait disorders in adults and the elderly : A clinical guide. Wien Klin Wochenschr. 2017 Feb;129(3-4):81-95.

9. 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.

10. 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.

11. Brin DW. Wellness programs raise privacy concerns over health data. The Society for Human Resource Management. 2016. https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/wellness-programs-raise-privacy-concerns-over-health-data.aspx Accessed December 29, 2020.

12. Raber I, McCarthy CP, Yeh RW. Health insurance and mobile health devices: Opportunities and concerns. JAMA. 2019 May 14;321(18):1767-8.

13. 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;In Press.

14. Hsu J. The Strava Heat Map and the End of Secrets. Wired. 2018. https://www.wired.com/story/strava-heat-map-military-bases-fitness-trackers-privacy/. Accessed December 29, 2020.

15. Perakslis ED. Protecting patient privacy and security while exploiting the utility of next generation digital health wearables. *Br Med J Op* Online [eLetter] 18 January 2019. https://blogs.bmj.com/bmj/2019/01/18/protecting-patient-privacy-and-security-while-exploiting-the-utility-of-next-generation-digital-health-wearables/.

16. 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.

17. 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.

18. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. Ann Behav Med. 2018 May 18;52(6):446-62.

19. 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.

20. 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.

21. 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.

22. 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 December 29, 2020.

23. 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 December 29, 2020.

24. Consumer Technology Association Standards. https://shop.cta.tech/collections/standards. Accessed December 29, 2020.

25. 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.

26. Personal Connected Health Alliance. Continua Design Guidelines. 2019. https://www.pchalliance.org/continua-design-guidelines. Accessed December 29, 2020.

27. IEEE. IEEE P1752 Open Mobile Health Working Group. 2020. https://sagroups.ieee.org/1752/ Accessed December 29, 2020.

28. 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.

29. 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.

30. Fitbit. Fitbit's 100+ Billion Hours of Resting Heart Rate User Data Reveals Resting Heart Rate Decreases After Age 40. 2018. https://s2.q4cdn.com/857130097/files/doc_news/Fitbits-100-Billion-Hours-of-Resting-Heart-RateUser-DataReveals-Resting-Heart-Rate-Decreases-After-Age-40.pdf. Accessed December 29, 2020.

31. Wearable Technologies. https://www.wearable-technologies.com/. Accessed December 29, 2020.

32. Redeker NS. Sensor technology for nursing research. Nurs Outlook. 2020 Jun 21.

33. University of Massachusetts Amherst. Institute for Applied Life Sciences. https://www.umass.edu/ials/about. Accessed December 29, 2020.

34. Ash GI. Exercise Support for People with Type 1 Diabetes Plus Sedentary Lifestyle and/or Overweight. Invited Talk: American College of Sports Medicine Brown Bag Series in Science. September 2020 Edition. https://www.acsm.org/learn-develop-professionally/brown-bag-series. Accessed December 29, 2020.

35. Ash GI, Robledo DS, Ishii M, Pittman B, DeMartini KS, O'Malley SS, et al. Using web-based social media to recruit heavy-drinking young adults for sleep intervention: Prospective observation. J MED INTERNET RES. 2020;22(8):e17449.

36. 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, pages: 366 - 383).

37. 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.

38. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010 Oct;11(10):733-9.

39. 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.

40. American Diabetes Association. 7. diabetes technology: Standards of medical care in diabetes-2020. Diabetes Care. 2020 Jan;43(Suppl 1):S77-88.

41. 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 [8], 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 [37]

Fig 3 Implementation roadmap

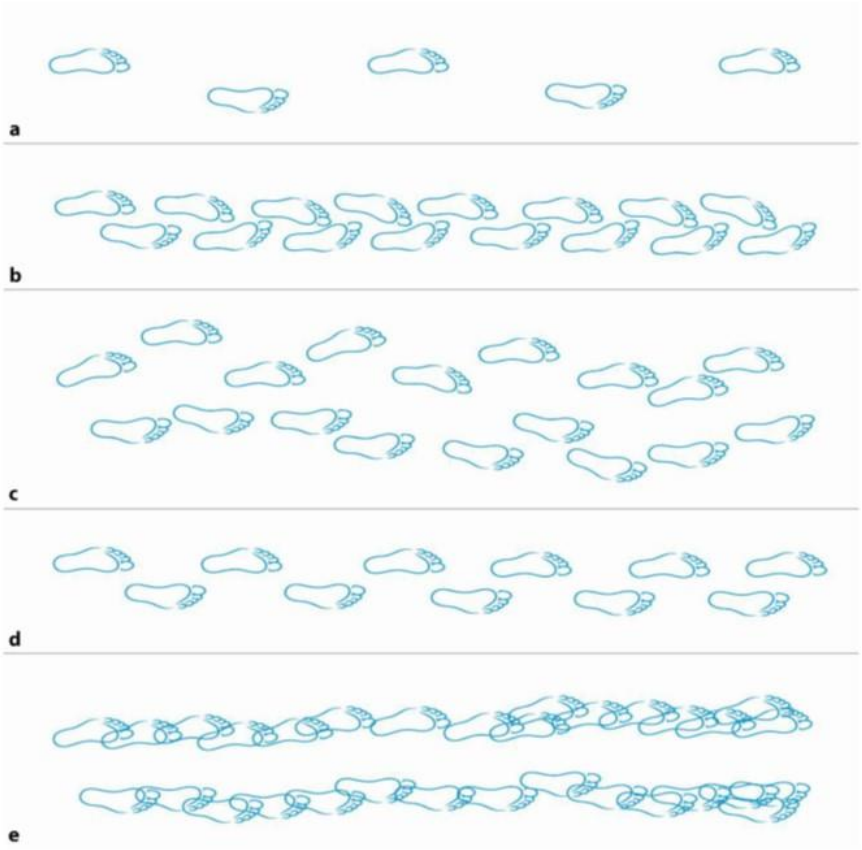
Normal gait

Spastic paraparetic gait

Cerebellar ataxic gait

Parkinsonian gait

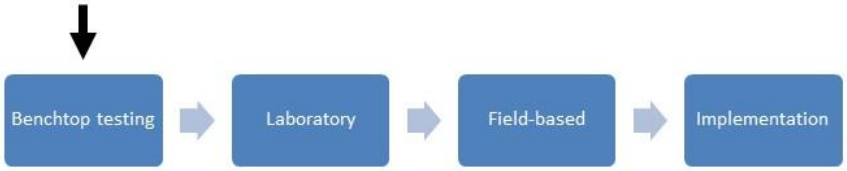
Frontal gait



Where in the pathway to test?

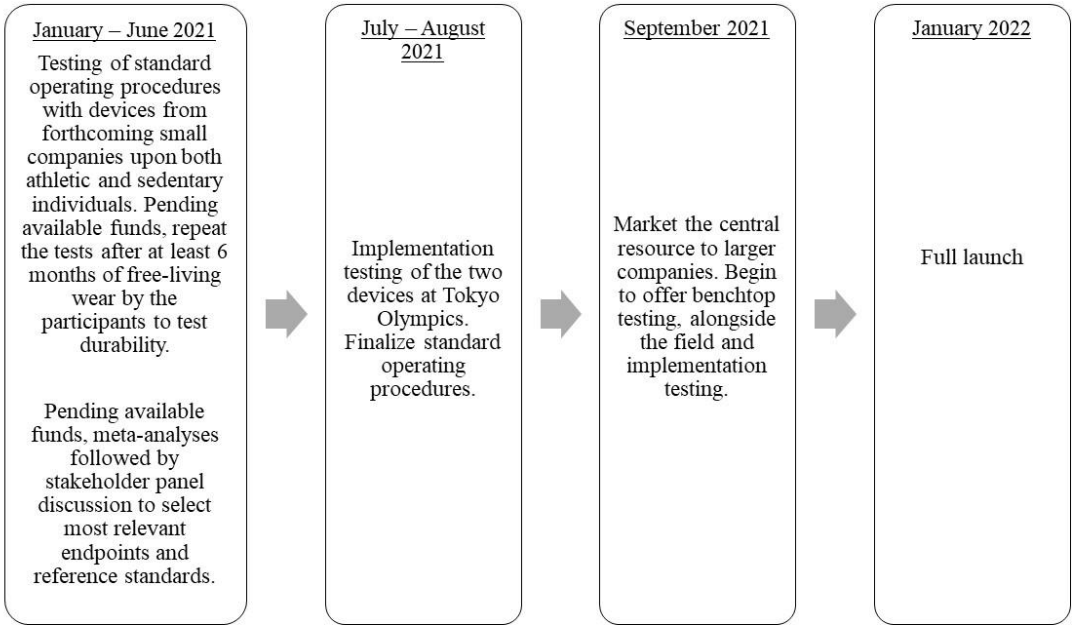
Most panel members: test here.

- Achieves nuanced, multilevel troubleshooting and constructive critiques



Large company representative: test here
(ie, data analytics to develop error-correction algorithms)

- Exploits large datasets available from marketed devices
- No redevelopment or risk to competitive advantage


January – June 2021
Testing of standard operating procedures with devices from forthcoming small companies upon both athletic and sedentary individuals. Pending available funds, repeat the tests after at least 6 months of free-living wear by the participants to test durability.
Pending available funds, meta-analyses followed by stakeholder panel discussion to select most relevant endpoints and reference standards.
July – August 2021
Implementation testing of the two devices at Tokyo Olympics. Finalize standard operating procedures.
September 2021
Market the central resource to larger companies. Begin to offer benchtop testing, alongside the field and implementation testing.
January 2022
Full launch