Citation: Clin Transl Sci (2021) 14, 94–101; doi:10.1111/cts.12851

MINI–REVIEW

Remote Digital Monitoring for Medical Product Development

Elena S. Izmailova[1], John A. Wagner[2], Nadir Ammour[3], Ninad Amondikar[4], Andrea Bell-Vlasov[5], Steven Berman[6], Dan Bloomfield[7], Linda S. Brady[8], Xuemei Cai[9], Roberto A. Calle[9], Michelle Campbell[6], Francesca Cerreta[10], Ieuan Clay[11], Luca Foschini[11], Pat Furlong[12], Rob Goldel[5], Jennifer S. Goldsack[13], Peter M.A. Groenen[14], Amos Folarin[15], Jill Heemskerk[16], Peter Honig[17], Matthew Hotopf[15], Tania Kamphaus[18], Daniel R. Karlin[19], Christopher Leptak[6], Qi Liu[6], Husseini Manji[20], Robert J. Mather[9], Joseph P. Menetski[18], Vaibhav A. Narayan[20], Elektra Papadopoulos[6], Bakul Patel[5], Bray Patrick-Lake[11], Jagdeep T. Podichetty[21], Abhishek Pratap[22], Laurent Servais[23,24], Diane Stephenson[21], Pam Tenaerts[25], Bruce J. Tromberg[16], Steve Usdin[26], Srikanth Vasudevan[5], Vadim Zipunnikov[27] and Steven C. Hoffmann[18,*]

The use of digital health products has gained considerable interest as a new way to improve therapeutic research and development. Although these products are being adopted by various industries and stakeholders, their incorporation in clinical trials has been slow due to a disconnect between the promises of digital products and potential risks in using these new technologies in the absence of regulatory support. The Foundation for the National Institutes of Health (FNIH) Biomarkers Consortium hosted a public workshop to address challenges and opportunities in this field. Important characteristics of tool development were addressed in a series of presentations, case studies, and open panel sessions. The workshop participants endorsed the usefulness of an evidentiary criteria framework, highlighted the importance of early patient engagement, and emphasized the potential impact of digital monitoring tools and precompetitive collaborations. Concerns were expressed about the lack of real-life validation examples and the limitations of legacy standards used as a benchmark for novel tool development and validation. Participants recognized the need for novel analytical and statistical approaches to accommodate analyses of these novel data types. Future directions are to harmonize definitions to build common methodologies and foster multidisciplinary collaborations; to develop approaches toward integrating digital monitoring data with the totality of the data in clinical trials, and to continue an open dialog in the community. There was a consensus that all these efforts combined may create a paradigm shift of how clinical trials are planned, conducted, and results brought to regulatory reviews.

The use of digital health technologies has gained considerable interest from consumers, providers, and researchers as a new way to improve therapeutic research and development. Although digital health technologies are being adopted by various industries and stakeholders, clinical trials have been slow to incorporate these technologies due, in part, to the complex global regulatory environment. As these products are developed by academic or industry experts in disparate fields, such as engineering, medicine, and computer modeling, adopting a shared framework and language is crucial to the advancement of digital tools acceptance and adoption in human research.

The Foundation for the National Institutes of Health (FNIH) Biomarkers Consortium hosted a public workshop entitled Remote Digital Monitoring for Medical Product Development on February 17–18, 2020. This workshop provided a venue to address challenges and opportunities in the use of remote sensing technologies for improving the probability of success of therapeutic clinical trials. The goals of this 2-day workshop included gathering diverse stakeholders in the field to apply a biomarker qualification evidentiary framework, to assess the use of remote digital monitoring for medical product development by discussing how well the existing definitions of biomarkers describe digital monitoring tools, examining the utility of the framework's key elements

[1]Koneksa Health, New York, New York, USA; [2]Foresite Capital, Boston, Massachusetts, USA; [3]Sanofi Research & Development, Chilly-Mazarin, France; [4]The Michael J. Fox Foundation for Parkinson's Research, New York, New York, USA; [5]Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, Maryland, USA; [6]Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [7]Anthos Therapeutics, Cambridge, Massachusetts, USA; [8]National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland, USA; [9]Pfizer, Inc, Cambridge, Massachusetts, USA; [10]European Medicines Agency, Amsterdam, The Netherlands; [11]Evidation Health, San Mateo, California, USA; [12]Parent Project Muscular Dystrophy, Hackensack, New Jersey, USA; [13]The Digital Medicine Society, Boston, Massachusetts, USA; [14]Idorsia Pharmaceuticals, Allschwil, Switzerland; [15]King's College London, London, UK; [16]National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Bethesda, Maryland, USA; [17]Pfizer, Inc., Collegeville, Pennsylvania, USA; [18]Foundation for the National Institutes of Health, North Bethesda, Maryland, USA; [19]HealthMode, Inc, New York, New York, USA; [20]Janssen Research and Development LLC, Titusville, New Jersey, USA; [21]Critical Path Institute, Tucson, Arizona, USA; [22]Sage Bionetworks, Seattle, Washington, USA; [23]University of Liège, Neuromuscular Reference Center Disease, Liege, Belgium; [24]MDUK Neuromuscular Center, University of Oxford, Oxford, UK; [25]Clinical Trials Transformation Initiative, Durham, North Carolina, USA; [26]BioCentury Publications, Washington, District of Columbia, USA; [27]Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.
*Correspondence: Steven C. Hoffmann (shoffmann@fnih.org)
Received: May 28, 2020; accepted: July 1, 2020. doi:10.1111/cts.12851

in evaluating the maturity of digitally measured biomarkers, and defining future directions to enable a broader adoption of digital tools in human research. Moreover, case studies on remote monitoring tools were examined to highlight areas in which new vocabulary may be needed, identify areas of high medical need that could be addressed using digital technologies, and ensure stakeholder alignment on the application of a biomarker qualification evidentiary framework (**Figure 1**) for the use of digital health technologies for therapeutic research and development. Ultimately, the insights and proceedings of the workshop will be incorporated into a white paper to inform operational and regulatory guidance. This minireview provides a timely summary of the workshop and accounting of both scientific and real-world challenges that impact remote monitoring measurement development. For the purposes of this workshop, "digital" is objective, quantifiable, physiological, functional, and/or behavioral data collected by means of wearables, ingestibles, implantables, and mobile technologies for the remote capture of data. The term "device" should be consistent with the US Food and Drug Administration (FDA) definition of a medical device; if a tool is not an FDA-cleared device, it should be called "technology." Biomarker is defined as "a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or biological responses to an exposure or intervention, including therapeutic interventions" according to the Biomarkers, Endpoints, and other Tools (BEST) Resource.

Important characteristics of tool development in this growing digital landscape were addressed in a series of presentations, case studies, and open panel sessions. The presentations covered a variety of topics, including challenges for the implementation of digital monitoring devices and technologies in drug development, regulatory guidance and evidentiary criteria for drug development tools (DDTs), and the digital National Institutes of Health (NIH) pipeline. As illustrated in **Figure 2**, a data-driven approach was used to choose the case studies based on identifiable criteria (e.g., type of sensor, its regulatory status, and algorithm transparency). A broad range of potential cases incorporating the use of technologies or devices were identified, and ultimately selected, based on several key parameters (e.g., novel or existing measures; commercial technologies or FDA-cleared devices; if the algorithm is open source or black box; and number of variables collected). These criteria were chosen to present different challenges in evidence collection and enrich the discussion during the workshop. Using this thought process, the team identified five case studies, across multiple therapeutic areas, that they felt would provide a good basis for identifying the types of evidence needed to develop reliable decision-making tools for drug research and development. The summary of case studies is presented in **Table 1**.

ISSUES, CONSIDERATIONS, AND CHALLENGES FOR REMOTE DIGITAL MONITORING

Advances in the field of digital monitoring technologies and devices with corresponding available data have created opportunities for expanded clinical trial participation and novel data streams. For example, electronic medical records adoption has more than doubled since 2008.[1] As such, years of longitudinal data are available along with the computing power to analyze large volumes of data. There are government initiatives driving forward the use of digital technology.[2–4] However, despite some initial progress of digital tool adoptions in human research, the widespread adoption is yet to happen because of a disconnect between promises of digital products and potential risks and a lack of trust for these new technologies.[5]

The evidentiary framework, originally designed for biomarker development and proposed as applicable to remote digital monitoring, (**Figure 1**) provides: (i) a biomarker description and related platform type along with a clear

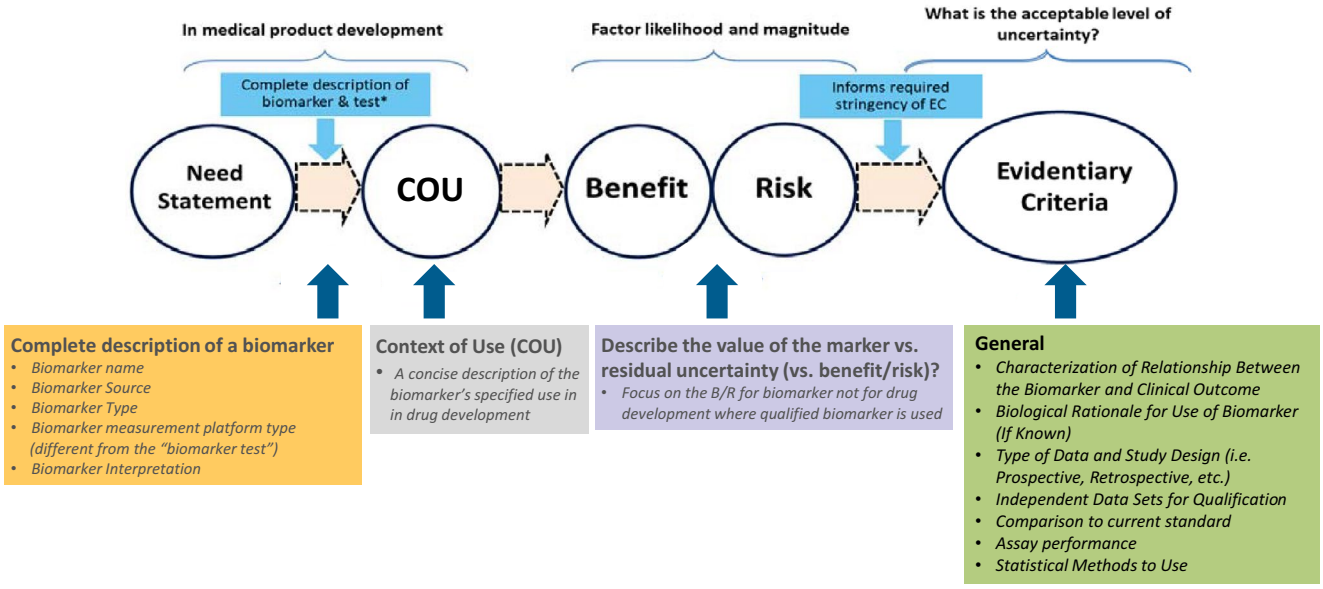

Figure 1 Evidentiary criteria (EC) framework.

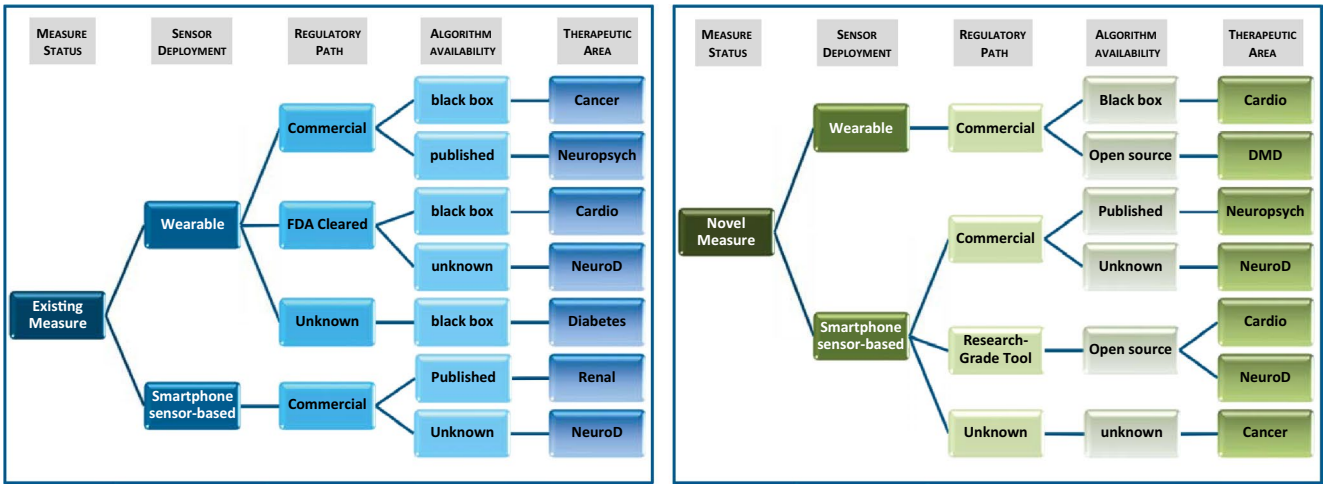

Figure 2 Strategy for selection of case studies. DMD, Duchenne's Muscular Dystrophy; FDA, US Food and Drug Administration.

explanation of the way the tool is to be used; (ii) key areas for defining unmet need; (iii) assessment of benefits and risks of deploying such a biomarker; and (iv) evidentiary criteria to be fulfilled to have a digital monitoring tool qualified.[6,7] Assumptions include a clearly defined goal, which provides a path to drug development decision making and regulatory approval and that the framework provides a context for discussion between sponsor and a regulatory agency.

The flow of data emerging from digital monitoring technologies, from unprocessed data to data transmittal, requires evidence at each step, for a biomarker or clinical outcome assessment (**Figure 3**). There are varying levels of evidentiary criteria needed for each digital DDT and these can change depending on benefit and risk of the tool and its use: universality, plausibility, causality, proportionality, specificity, and potential for off-target effects. A clinical outcome is defined as an outcome that describes or reflects how an individual feels, functions, or survives. Digital health technology tools can also be used to assess clinical outcomes. For example, activity monitor-based end points can be designed to reflect aspects of a patient's physical functioning (e.g., ambulation). To be truly patient-centric, such an end point often necessitates input from multiple stakeholders, including patients, disease experts, clinical experts, caregivers, engineers, and others to ensure its clinical relevance.

In terms of the development process, specific components of a digital measure will help define which stakeholders should be included to optimize development. The performance characteristics of a technology should fit its purpose and intended use.[8] Because technologies evolve rapidly, a specific product could become obsolete before all the data are collected to satisfy a regulatory agency. Finally, having established performance characteristics does not necessarily translate to readiness for use in clinical trials or patient care; the security, data rights/governance, utility/usability, and economic feasibility also need to be evaluated.[9]

REGULATORY GUIDANCE AND EVIDENTIARY CRITERIA FOR DRUG DEVELOPMENT TOOLS

Existing regulatory evidence frameworks and processes for DDTs were reviewed to generate discussion regarding the needs for digital tool evaluation.[6] The term DDT is defined broadly and includes biomarkers, clinical outcome assessments, or any other method, material, approach, or measure that can aid drug development and regulatory review. The drawback of such a broad definition is that regulators find it difficult to provide generalizable advice that spans all potential DDTs.

The conceptual framework for biomarker development for regulatory acceptance includes: a drug development needs statement, a context of use (COU), the benefits and risks to the patient, and evidentiary criteria, including data on analytical and clinical validity. There are three development pathways that can be considered in parallel: the

Table 1 Summary of case studies

Case study name	Therapeutic area/ disease indication	Measurement
mPower	CNS/PD	Mobile tools for PD natural history
Cardiac monitoring in phase I clinical trials	Normal healthy volunteers	Heart rate, respiratory rate, skin temperature
VERKKO study	Metabolic disorders/diabetes type II	CGM
Stride Velocity 95th Centile 2° Endpoint in DMD	Muscle skeletal disorders/ DMD	Stride Velocity 95th Centile measured by a wearable sensor
RADAR – MDD	Mental health/MDD	Diverse smartphone and wearable sensors to monitor physical activity, sociability, sleep, cognition, speech, mood and stressors

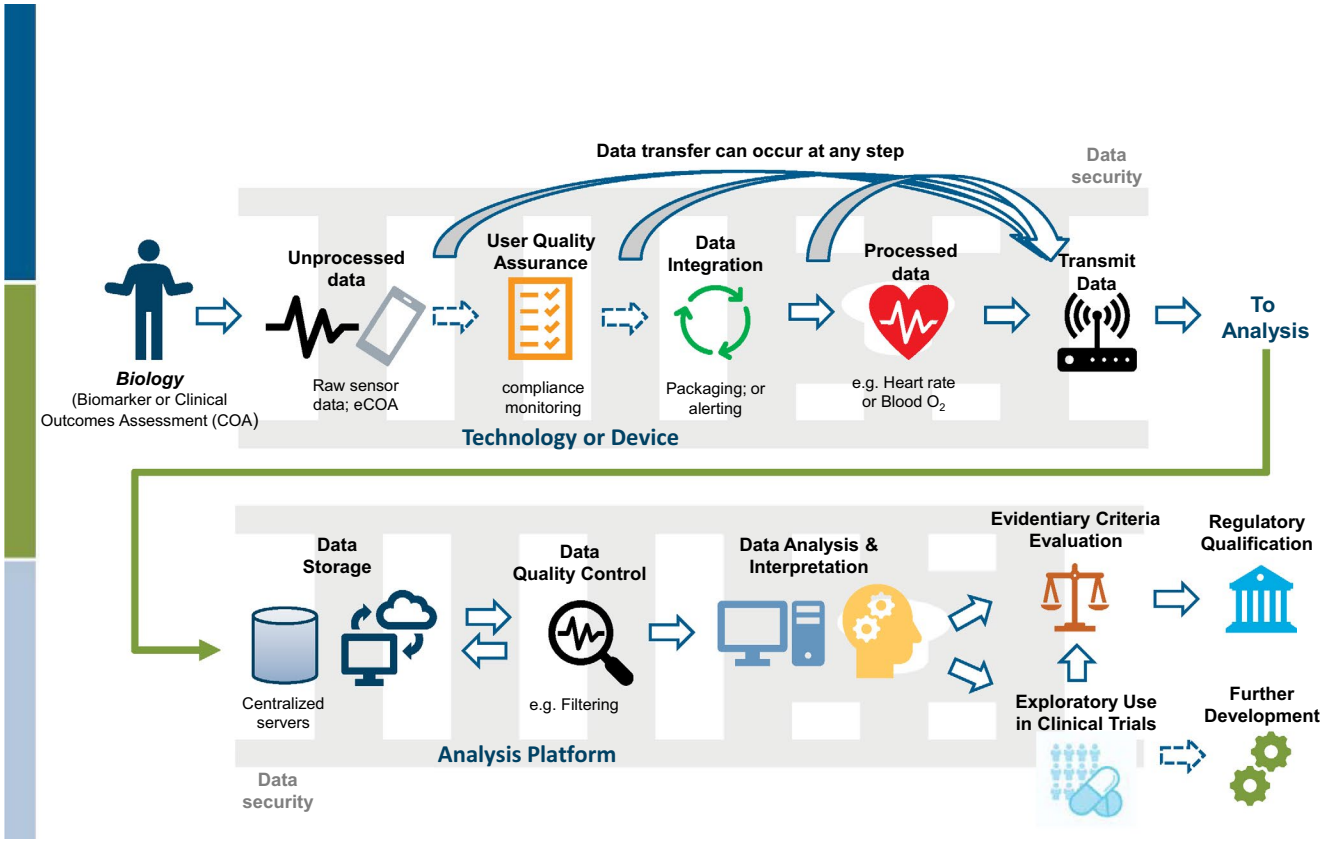

Figure 3 Flow of medical product data: from biology to decision.

Investigative New Drug pathway in the context of a drug development program; scientific community consensus through thought leaders and published papers; and the DDT qualification program.

The BEST resource is a living document created by the NIH-FDA Biomarker Working Group that defines terminology and uses of biomarkers and end points in basic biomedical research, medical product development, and clinical care.[10] Of eight different classes of biomarkers described, two are particularly relevant for digital technologies: (i) monitoring biomarker: used to serially assess subject status and (ii) pharmacodynamic response biomarker: used to detect if a biological response has occurred in a patient who has received a therapeutic intervention and it may become a clinical trial end point.

Elements that enable DDT development include data quality, reproducibility, standards; standard operating procedures or protocols; preanalytical standardization; and evaluation of impact on clinical trial elements. In addition to COU and benefit/risk assessment, the drug development need must be established for the proposed DDT. Some of the following topics may be considered when establishing the drug development need: What question is the tool intended to address? If there are existing tools, what improvement does the proposed new tool provide? How does the new technology address the unmet need? If there is success with the new tool, how would decision making change in a clinical trial? The DDT qualification process offers a means for direct engagement with the FDA and allows DDT development to occur outside the path of a specific drug development program.

CASE STUDIES

Case study #1: Using consumer sensor technologies to measure Parkinson's disease - Lessons learned from the mPower study

The mPOWER study was a fully remote, proof-of-principle, large scale study (N = 16,000) with subjective and objective measures taken over 50 months. Four objective sensor-based measurements of symptom domains (voice, gait/balance, dexterity/speed, and memory) were analyzed with corresponding standard clinical scales.[11]

The goal was to develop a method to monitor symptoms of Parkinson's disease (PD) in patients who are levodopa responsive with levodopa-induced peak-dose dyskinesias followed by re-emergence of tremor, rigidity, and bradykinesia (i.e., on/off symptoms). The benefit of such a measurement is noninvasive data collection using a smartphone app with a minimal burden for the participants. The risk of deploying such an application remotely included potential biases in the long-term data collection and downstream data processing errors. However, the more frequent measurements collected by smart apps compared with Unified Parkinson's Disease Rating Scale-based assessments during the clinical visits provide more objective measurements and more completely reflect the patient experience.

The study results demonstrated high intra-person variability in selected measurements, reflecting good days and bad days of living with PD as patients with PD have more variations over time. The challenge posed by the need to compare a new digital tool to the existing disease assessments, such as Unified Parkinson's Disease Rating Scale, were discussed as they present a formidable challenge for developing a new measurement benchmarking it to an instrument with serious limitations. The lessons learned included the need for addressing biases in participant recruitment and long-term retention in fully remote studies.[12]

Case study #2: Cardiac monitoring in phase I clinical trials

The current paradigm of safety monitoring foresees data collection and predefined time points during study subject confinement to the clinical pharmacology unit (CPU). The limitations of this approach include sporadic data that can miss safety signals collected in a somewhat artificial environment that does not reflect how people feel and function in their normal daily living environment. Additionally, the data collected after the CPU discharge is solely based on patient recall, which may not be reliable.

This case study examined the results from two clinical trials that evaluated single-lead echocardiogram devices, measuring heart rate, respiratory rate and skin temperature, and wrist-worn actigraphy measurements, assessing physical activity, and sleep parameters, as exploratory end points.[13,14] Continuous safety monitoring in early stage clinical trials has multiple benefits, including dense data sets allowing an early detection of a potential safety signal, optimization of dose selection, and moving safety monitoring from CPU to participant's homes, enabling easier participation in clinical trials and more feedback from participants. Risks included missed safety signals, data loss, and a large amount of false-positive signals that require time-consuming follow-ups.

This case study reinforced the notion that having 510(k) regulatory clearance does not necessarily render a device fit-for-purpose for use in clinical trials. Validation in the COU is needed. A need for reference ranges and reference intervals for continuous measurements, as well as novel statistical approaches to analyze and interpret large amounts of data, was highlighted.

Case study #3: Digital remote continuous glucose monitoring

Despite availability of multiple drug classes to treat diabetes, a large portion of the patient population remains inadequately controlled, highlighting continued unmet clinical need that merits new drug development. HbA1c is a biomarker used in the late stage clinical trials to assess efficacy of antidiabetic investigational drugs. However, early drug development for diabetes requires dynamic characterization of the acute glucose response to pharmacologic intervention with a direct measurement. Frequent direct analyzer-based glucose measures are only possible in the very early phase I inpatient setting. Self-monitoring of (capillary) blood glucose (SMBG) levels is used but is not convenient for participants and limits data volume.

Continuous glucose monitoring (CGM) can provide more data that includes more frequent measurements and data captured during time periods not normally monitored with SMBG (e.g., sleep and exercise). If automated remote glucose sensing is effectively integrated with prompt automated transmission of data from capture to database, it can also shorten cycles for new data integration in decision making. One early example of integration of remote glucose-monitoring technology with real-time data streaming is the VERKKO trial. This was a pilot virtual clinical study that enrolled 51 participants with type 2 diabetes with the objective of studying the performance of a 3G-enabled wireless glucose monitor (SMBG) that streamed data to a central database in real time as captured by the study participants.[15]

The VERKKO study was used as the platform for the discussion of implementation of a rich daily glucose profile collected through CGM to measure parameters, such as Time in Range, Time Above Range, and Time Below Range that could be used as pharmacodynamic/response biomarkers for dose selection in adult participants with diabetes receiving an experimental antidiabetic agent in early drug development studies. Potential benefits afforded by CGM with real-time data streaming include the ability to provide a more comprehensive characterization of each subject's glucose profile than discrete time measurements by SMBG, and rapid integration of collected data into pharmacokinetic/pharmacodynamic models utilized for dose optimization. Hurdles to implementation include technical (multiple sources of variability) and clinical validity (direct correlation to outcomes) gaps.

Case study #4: Stride velocity 95th centile secondary end point in Duchenne's muscular dystrophy

The 6 minute walk test in Duchenne's Muscular Dystrophy is a functional test, measuring submaximal exercise capacity as the total distance a patient can walk on a 30 minute straight and unimpeded track and is used to predict how long a patient will remain ambulatory. It is often used as a primary end point in clinical trials. Using this test in clinical trials presents several challenges: the population in clinical trials can be skewed by patients that train to participate; shorter duration tests are influenced by patient reflexes and longer tests by patient motivation. Patients with rare diseases may have to travel to participate in clinical trials; consequently, the baseline testing may not accurately reflect disease activity due to fatigue and other factors.

The 95th percentile stride velocity can be seen as a disease response or pharmacodynamic marker of treatment in ambulatory patients diagnosed with different neuromuscular diseases, including Duchenne's Muscular Dystrophy.[16] This case study considered the data from wrist-or-ankle-worn wearable sensor that was constructed to continuously identify and quantify movements. The 95th percentile stride velocity correlated with the established 6-minute walk test. This measure could be deployed to assess an early efficacy signal in a drug trial. High reliability of data collected in natural settings can lead to smaller clinical trials and reduce patients' burden of participating in clinical trials.

This case study emphasized the need of getting patient input early in the process of technology development.

Clinical development in rare diseases can be particularly challenging because very few centers will conduct clinical studies; however, partnership with patient advocacy groups can be beneficial. The regulatory approval from the European Medicines Agency (EMA) paved the ground for the future regulatory submissions, which should be less time-consuming (e.g., qualification of real-world gait speed or upper limb movements in other indications).[17]

Case study #5: Remote assessment of disease and relapse in major depressive disorder

Clinical research in major depressive disorder (MDD) is hindered by several issues: there is no valid diagnostic biomarker and diagnosis is syndromal; symptoms are self-reported at infrequent intervals, leading to recall bias. MDD is usually treated as a single entity, whereas diagnosis requires definition of multiple subtypes, which might be used to stratify participants into therapeutic trials.

This case study focuses on a 2-year study of remote assessment of ~ 600 patients with MDD using a technology platform that collects, processes, and manages passive data from wearable devices, smartphone sensors, and active remote measurements (e.g., performing tasks and answering questionnaires).[18,19]

The benefits of this methodology include a pathway for novel drugs with novel mechanisms targeting patients' unmet needs in defined subpopulations. However, this digital assessment approach may be less prone to the placebo effect and allow for smaller, faster clinical trials. The risks include incorrect patient stratification leading to suboptimal real-world treatment benefits and limited patient participation due to the requirement for a digital device. Additionally, the patient voice may not be adequately captured if there is over-reliance on objective passive measures without incorporating task-based measures that incorporate the patient perspective.

The case study focused on the ability to identify an MDD subtype characterized by hyperarousal, which manifests itself as sleep disturbances and ruminative thoughts. This subtype of MDD is inadequately treated. Sufficient evidence is available that sleep disturbance can be identified via actigraphy-based digital measures. Plausible evidence is available that hyperarousal can be identified by a combination of self-reports and physiological measures from wearable sensors.

This case study highlighted the challenges of developing new digital end points of efficacy in the context of limitations of a reliable gold-standard. There is a distinct possibility that passive digital assessments of efficacy may indeed be more reliable and valid than subjective clinical assessment scales, which are widely used.

NIH PIPELINE IN DIGITAL TECHNOLOGIES

The current state of health care is impoverished for information. Blood, imaging, and even genomic tests are performed annually or every few years and only provide a static snapshot of a condition or disease. Continuous monitoring could optimize health through reductions in disease and hospitalizations and reduce costs by empowering patients to manage their own access to health care. The National Institute of Biomedical Imaging and Bioengineering (NIBIB) aims to use modeling, computation, and machine intelligence, together with therapeutic devices, imaging modalities, engineered biology, and sensors to better understand and apply these technologies at the point of care. The Institute has developed several strategies to reach these aims: (i) Point of Care Technologies Research Network centers focused on developing point-of-care technologies and sensors; (ii) workshops on integrating machine learning with multiscale modeling[20]; and (iii) convening a strategic planning group across several fields (e.g., data science, imaging, engineered biology, etc.) to discuss the current state of each field. Through these programs, the NIBIB looks to lead the transition of future healthcare systems away from episodic snapshots to continuous monitoring approaches and technologies that improve patient healthspans and lifespans.

CONCLUSIONS

Harmonized vocabulary

The need for a common lexicon for digital technologies is essential. Confusion over terms and definitions of devices, measures, and biomarkers hinder medical practice and drug development, often leading to a misinterpretation of evidence and a misunderstanding of evidentiary requirements and regulations. The BEST glossary is meant to be a "living" resource and needs updated language and definitions that address remote or digital monitoring. Fostering consistent usage of terms and resources will help to accelerate medical product development and improve health outcomes.

Regulatory alignment

The evidentiary criteria framework provided a very useful tool to organize the information about certain DDTs and identify gaps to be filled before these tools can be used in clinical trials. The concept of digital monitoring is not about specific tools; it is about establishing the purpose, validating tools for their intended use, and deploying digital measurements to get to interpretable results. It is important to evaluate potential digital tools considering their intended COU, as the evidentiary requirements may differ between uses. For example, a DDT intended to be used as a surrogate end point is likely to have higher standards for clinical and analytical validity than a tool used for prognostic enrichment in a clinical trial. Regulators also highlighted the importance of early engagement and frequent communication. Last, harmonization across regulatory agencies and regions is key and remains a work in progress.

Patient centricity

Patient engagement and partnership with physicians and healthcare professions is crucial to align these important stakeholders to drive the adoption of digital monitoring forward. Engagement with patients and patient groups early and throughout the development cycle is critical. These coordinated approaches will help patients better engage payers to gain access to devices. Digital medicine has the

promise to enable greater patient engagement and participation in trials.

Power of digital assessments

Digital monitoring tools provide an opportunity to measure disease in a way that is both meaningful to patients and clinically meaningful to providers, drug developers, and regulators. It is also an opportunity to develop objective and continual quality of life measurements that go beyond subjective and episodic self-reports.

Precompetitive collaborations

Data accessibility and precompetitive collaborations are key for success, but it is easier said than done. Establishing collaborations between different parties may take a long time, and aligning on goals can be challenging. The workshop participants challenged the notion of proprietary value of digital monitoring data and emphasized the role of cross-sector partnerships in advancing regulatory acceptance of tools that are of broad relevance.

Although substantial progress of digital monitoring and application of evolving regulatory frameworks for evidentiary criteria was demonstrated, a number of issues were expressed: (i) the digital monitoring community needs to establish and share real-life examples of the validation parameters. The issue of conventionally accepted comparators or legacy standards, also called "gold standards," was discussed as they are essential for establishing analytical validity by comparing new measurements to a conventional one. Legacy standards are not uniformly the ideal measurements. The limitations of legacy standards are well understood; however, there is no immediate solution to replace these measurements with alternative methods: (ii) the data types and volume of data require completely new approaches for signal recognition and data interpretation. Co-evolution of devices, validation methodology, and data science are crucial: (iii) the current evidentiary framework lacks definitions and clarity for the validation and use of remote measures as biomarkers and needs to evolve based on the needs and characteristics of digital monitoring.[21]

FUTURE DIRECTIONS

The digital monitoring community needs to harmonize semantics, vocabularies, and definitions in order to build common methodologies and practices. Collaborations and leveraging experience from other life science disciplines can be highly beneficial. Digital monitoring data need to be integrated with the totality of the data in clinical trials, which includes real-world data/real-world evidence, conventional molecular or imaging biomarkers, and imaging end points to create a holistic picture of the health care of the future. This community needs to pool resources and expertise and seek expanded opportunities to share experiences and debate the most pressing issues in order to bring those back to their respective organizations. Success in bringing these digital technologies to the forefront of health care will require a multidisciplinary group comprised of a new breed of scientists, entrepreneurs, physicians, data analysts, regulators, and patient advocacy groups to name a few. This cross-stakeholder group needs to work together to accumulate enough information that can be shared on a precompetitive basis to drive a paradigm shift on how clinical trials are planned, conducted, and results brought to regulatory reviews.

Acknowledgments. The content of this article rests with the authors, who sought to provide highlights from presentations and comments made by individual participants at the workshop. The authors gratefully acknowledge Carolyn Shore and Eeshan Khandekar (National Academy of Sciences, Engineering, and Medicine) and Denyse Niemerov (Abbott) for their expertise and refinement of the workshop aims and outcomes. We also acknowledge Lynn Anderson (Prescott Medical Communications Group) for her formatting and editorial assistance. Webcast video and slide presentations of the workshop plenary sessions and panel discussions are available at https://fnih.org/what-we-do/biomarkers-consortium/programs/digitalmonitoring.

Funding. The FNIH Biomarkers Consortium, which hosted this workshop, is a major public-private biomedical research partnership supported by stakeholder membership, including government, industry, academia, and patient advocacy, and other not-for-profit organizations.

Conflict of Interest. E.S.I. is an employee of Koneksa Health and owns stock. J.S.G. is a part-time employee of HealthMode. J.M. and S.C.H. are employees of the Foundation for the National Institutes of Health (FNIH). P.M.A.G. is an employee of Idorsia Pharmaceuticals and owns stock. J.A.W. is an employee of Foresite Capital and a member of the FNIH Biomarkers Consortium executive committee. I.C. is an employee of Evidation Health, Inc., owns stock and is an editor for Karger Digital Biomarkers. M.H. is a PI of the RADAR-CNS program grant funded by a pre-competitive private public partnership with members of the European Federation of Pharmaceutical Industries and Associations, which provided research funds (Janssen, Biogen, UCB, MSD, and Lundbeck). D.R.K. is an employee and shareholder of HealthMode, Inc. B.P.-L. is a consultant for Bayer. All other authors declared no competing interests for this work.

As Editor-in-Chief of Clinical and Translational Science, John Wagner was not involved in the review or decision process for this paper.

Disclaimers. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services. This article reflects the views of the authors and should not be construed to represent FDA's or EMA's views or policies.

Word Count Overage Justification. This succinct review provides an expedited accounting of the key content and outcomes from the 2-day Workshop that included a series of plenary presentations, five case studies, and open panel sessions. Especially with the real-time and pressing need for guidance and regulatory alignment for the application of remote digital tools in clinical trials, the authors justify the necessary extended word count to fully contextualize the Workshop's content in an abbreviated but accelerated manor. Further reduction in content would dilute the instructive steps and case study relevance to your readers and the multisector stakeholders in this field.

1. Office-based physician electronic health record adoption <https://dashboard.healthit.gov/quickstats/pages/physician-ehr-adoption-trends.php>. Accessed April 24, 2020.
2. Notice of proposed rulemaking to improve the interoperability of health information <https://www.healthit.gov/topic/laws-regulation-and-policy/notice-proposed-rulemaking-improve-interoperability-health>. Accessed April 24, 2020.
3. FDA's Technology Modernization Action Plan (TMAP) <https://www.fda.gov/media/130883/download>. Accessed April 24, 2020.
4. Health IT in health care settings <https://www.healthit.gov/topic/health-it-health-care-settings/health-it-health-care-settings>. Accessed April 24, 2020.
5. DiMe: Calling all who serve in digital medicine <https://www.statnews.com/2019/06/05/new-digital-medicine-society/>. Accessed April 24, 2020.
6. Qualification Process for Drug Development Tools Guidance for Industry and FDA Staff <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/qualification-process-drug-development-tools-guidance-industry-and-fda-staff>. Accessed April 25, 2020.
7. Leptak, C. *et al*. What evidence do we need for biomarker qualification? *Sci. Transl. Med.* **9**, eaal4599 (2017).
8. Goldsack, J.C. *et al*. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *NPJ Digit. Med.* **3**, 55 (2020).
9. Coravos, A. *et al*. Modernizing and designing evaluation frameworks for connected sensor technologies in medicine. *NPJ Digit. Med.* **3**, 37 (2020).
10. BEST (Biomarkers, EndpointS, and other Tools) Resource: Monitoring Biomarker <https://www.ncbi.nlm.nih.gov/books/NBK402282/>. Accessed April 25, 2020.
11. Bot, B. *et al*. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci. Data* **3**, 160011 (2016).
12. Pratap, A. *et al*. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit. Med.* **3**, 21 (2020).
13. Izmailova, E.S. *et al*. Continuous monitoring using a wearable device detects activity-induced heart rate changes after administration of amphetamine. *Clin. Transl. Sci.* **12**, 677–686 (2019).
14. Izmailova, E.S. *et al*. Evaluation of wearable digital devices in a phase I clinical trial. *Clin. Transl. Sci.* **12**, 247–256 (2019).
15. Schrangl, P. *et al*. Limits to the evaluation of the accuracy of continuous glucose monitoring systems by clinical trials. *Biosensors (Basel)* **8**, 50 (2018).
16. Lilien, C. *et al*. Home-based monitor for gait and activity analysis. *J. Vis. Exp.* **8**, 150 (2019).
17. Haberkamp, M. *et al*. European regulators' views on a wearable-derived performance measurement of ambulation for Duchenne muscular dystrophy regulatory trials. *Neuromuscul. Disord.* **29**, 514–516 (2019).
18. Yatharth, R. *et al*. RADAR-Base: Open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. *JMIR Mhealth Uhealth* **7**, e11734 (2019).
19. Matcham, F. *et al*. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. *BMC Psychiatry* **19**, 72 (2019).
20. Integrating machine learning with multiscale modeling for biomedical, biological, and behavioral systems (2019 ML-MSM) <https://www.nibib.nih.gov/news-events/meetings-events/integrating-machine-learning-multiscale-modeling-biomedical-biological-and-behavioral-systems-2019-ml-msm-0>. Accessed April 25, 2020.
21. Goldsack, J.C., Izmailova, E.S., Menetski, J.P., Hoffmann, S.C., Groenen, P.M.A. & Wagner, J.A. Remote digital monitoring in clinical trials in the time of COVID-19. *Nat. Rev. Drug Discov.* **6**, 378–379 (2020).