

Review Article

Digital Medicine: A Primer on Measurement

Andrea Coravos^{a–c} Jennifer C. Goldsack^{c, d} Daniel R. Karlin^{c, e, f}
Camille Nebeker^{g, h} Eric Perakslis^{i, j} Noah Zimmerman^{k, l} M. Kelley Erb^m

^aElektra Labs, Boston, MA, USA; ^bHarvard-MIT Center for Regulatory Science, Boston, MA, USA; ^cThe Digital Medicine Society (DiMe), Boston, MA, USA; ^dmonARC Bionetworks, San Francisco, CA, USA; ^eHealthMode, New York, NY, USA; ^fTufts University School of Medicine, Boston, MA, USA; ^gDepartment of Family Medicine and Public Health, School of Medicine, University of California, San Diego, La Jolla, CA, USA; ^hCenter for Wireless and Population Health Systems, School of Medicine, University of California, San Diego, La Jolla, CA, USA; ⁱRubenstein Fellow – Duke Forge, Durham, NC, USA; ^jHarvard Medical School, Boston, MA, USA; ^kDepartment of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^lInstitute for Next Generation Healthcare, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ^mDigital Medicine and Translational Imaging, Worldwide Research, Development, and Medical, Pfizer Inc., Cambridge, MA, USA

Keywords

Digital medicine · Digital health · Digital biomarkers · Digital technology tools · FDA · Ethics · Clinical research · Clinical care · Data rights · Data governance · Mobile technology · Sensors · Privacy · Cybersecurity · Regulatory science · Decentralized clinical trials · Research ethics

Abstract

Technology is changing how we practice medicine. Sensors and wearables are getting smaller and cheaper, and algorithms are becoming powerful enough to predict medical outcomes. Yet despite rapid advances, healthcare lags behind other industries in truly putting these technologies to use. A major barrier to entry is the cross-disciplinary approach required to create such tools, requiring knowledge from many people across many fields. We aim to drive the field forward by unpacking that barrier, providing a brief introduction to core concepts and terms that define digital medicine. Specifically, we contrast “clinical research” versus routine “clinical care,” outlining the security, ethical, regulatory, and legal issues developers must consider as digital medicine products go to market. We classify types of digital measurements and how to use and validate these measures in different settings. To make this resource engaging and accessible, we have included illustrations and figures throughout that we hope readers will borrow from liberally. This primer is the first in a series that will accelerate the safe and effective advancement of the field of digital medicine.

© 2019 The Author(s)

Published by S. Karger AG, Basel

Jennifer C. Goldsack
The DiMe Society and monARC Bionetworks
4892 Ancient Marble Drive
Sarasota, FL 34240 (USA)
E-Mail jennifer@dimesociety.org

Introduction

Digital medicine products hold great promise to improve medical measurement, diagnosis, and treatment. While many industries have embraced digital disruption, the healthcare industry has yet to experience the improvements in outcomes, access, and cost-effectiveness long promised by the digital revolution. Healthcare lags behind other industries in part because of the regulatory environment, which tends to slow progress as health authorities strive to minimize adverse outcomes.

Developing effective digital medicine tools is an intensive and challenging process that requires the interdisciplinary efforts of a wide range of experts, from engineers and ethicists to payers and providers. Many of the challenges are compounded by the multidisciplinary nature of this field. The advancement of digital medicine stalls when constituent experts speak different languages and have different standards, experiences, and expectations.

What Is the Goal of This Primer?

We believe that effective communication is essential for turning scientific discoveries into commercial products. Having unclear definitions and inconsistent terminology hinders our abilities to evaluate scientific evidence and, ultimately, develop successful medical products.

Our goals with this primer are to:

- 1 Promote effective collaboration among different stakeholders by providing a common framework of language and ideas within which to collaborate
- 2 Support the advancement of measurement in digital medicine by clarifying core concepts and terms

To achieve these goals, we synthesize the basics of clinical medicine, medical research, regulation, and ethics into an accessible and digestible form and, by clarifying core concepts and terms, we aim to drive the field forward.

This primer focuses specifically on measurement in digital medicine, a foundational component underpinning the decentralization and democratization of clinical care and clinical trials using digital tools. We will also attempt to explicate relationships between digital measurement in research and digital measurement in clinical care. Though these are interrelated concepts, and much technology moves fluidly between research and care, we have chosen the focus on research as this seems to be a logical sequence. The ability to demonstrate reliability and meaningfulness for clinical trials, whether clinic based or otherwise, will ultimately translate into clinical use. Although the research space is fragmented, it is far more cohesive and unitary than clinical care, and we believe that effecting changes in practice across the research domain in a timely manner is a feasible goal, which will benefit patient care both through the translation of new technology and the creation and approval of novel treatments. While our treatment of clinical care may seem sparse, we do attempt to cover a breadth of applicable examples.

This piece is the first in a planned series of primers that will address all aspects of digital medicine. Collectively, the primers will provide a comprehensive introductory resource to digital medicine, equipping all professionals working in the field with the knowledge and language they need to advance the practice of digital medicine and, in turn, patient care.

Who Will This Primer Help?

- **Technology experts**, including software engineers, designers, data scientists, security researchers, and product managers who want to deepen their healthcare knowledge.
- **Academic researchers** and **industry sponsors** of clinical trials, both of whom need to facilitate internal discussions across teams (e.g., data science teams working with protocol designers in the translational medicine teams).

- **Clinicians**, who will be increasingly exposed to digital medicine in their practice.
- **Patients**, who we believe will drive more of their own care as the practice of medicine becomes more personalized and consumer-oriented.

How Should You Use This Primer?

In this primer, we use **bolded** words to indicate important terms or phrases for the field. Where possible, we reference existing definitions. Where we found conflicting definitions, we propose a revised definition. We hope that standardizing terminology will help unify and advance the field. We also recognize that these definitions will evolve over time.

We have organized the primer into three parts:

- **Part I** is an overview of digital medicine, focusing on the software and algorithms that are being used to measure individuals' health and intervene to improve their condition.
- **Part II** is designed for readers newer to the ethical, legal/regulatory and social implications (ELSI) associated with health research and healthcare. We provide an overview of “clinical research” versus routine “clinical care” and the considerations as a product goes to market.
- **Part III** introduces terms that classify types of digital measurements, such as digital biomarkers and electronic clinical outcome assessments. It also describes how to think through developing a digital measure for use in a clinical trial setting versus clinical care – and important considerations to ensure the measures are trustworthy, such as the concepts of verification and validation.

Borrowing from the success of our colleagues in cancer immunotherapy, and research in emergency medicine, we have distilled key clinical ontologies and frameworks into cartoon illustrations [1, 2]. We intentionally published this primer in an open-access journal because it is our hope that readers borrow liberally from this work, both using the concepts and illustrations for internal and external presentations and to spark discussions.

What Does Success Look Like?

As leaders in our field have stated before us, if we are successful in “accelerating the advancement of digital medicine, then soon, we will just be calling it ‘medicine’ [3].”

We share the same vision for the future.

Part I: An Overview of Digital Medicine – Measurements and Interventions

What Is Digital Medicine?

Digital medicine describes a field, concerned with the use of technologies as tools for **measurement**, and **intervention** in the service of human health [4]. Digital medicine products are driven by high-quality hardware and software products that support health research and the **practice of medicine** broadly, including treatment, recovery, disease prevention, and health promotion for individuals and across populations (Fig. 1).

Digital medicine products can be used independently or in concert with pharmaceuticals, biologics, devices, or other products to optimize patient care and health outcomes. Digital medicine empowers patients and healthcare providers with intelligent and accessible tools to address a wide range of conditions through high-quality, safe, and effective measurements and data-driven interventions. Digital products are also used in health research to develop knowledge of the fundamental determinants of health and illness by examining the biological, environmental and lifestyle factors. Digital technologies are increasingly used in observational and interventional research applied to prevention and treatment of disease as well as health promotion.

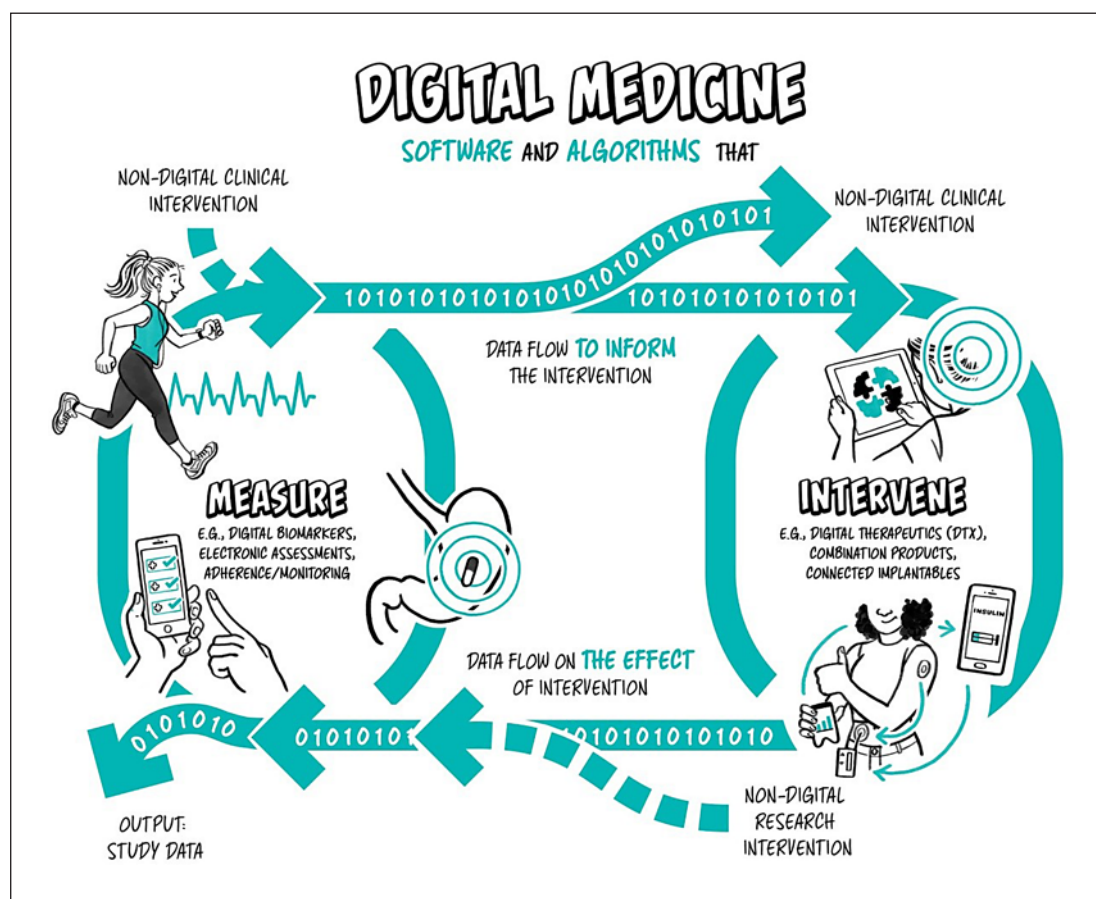


Fig. 1. Digital medicine overview. Digital medicine uses software and algorithmically driven products to measure or intervene to improve human health.

As a discipline, **digital medicine** encapsulates both broad professional expertise and responsibilities concerning the use of these digital tools. Digital medicine focuses on **evidence generation** to support the use of these technologies. **Measurement products** include digital biomarkers (e.g., using a voice biomarker to track change in tremor for a Parkinson’s patient), electronic clinical outcome assessments (e.g., an electronic patient-reported outcome survey), and tools that measure adherence and safety (e.g., a wearable sensor that tracks falls and smart mirrors for passive monitoring in the home) [5]. Digital measurement products are the focus on this primer.

Intervention products include **digital therapeutics** and **connected implantables** (e.g., an insulin pump). Digital therapeutics deliver evidence-based therapeutic interventions to patients that are driven by high-quality software programs to prevent, manage, or treat a medical disorder or disease. They are used independently or in concert with medications, devices, or other therapies to optimize patient care and health outcomes [6]. Notably, digital intervention products are not the primary focus on this primer, although a companion primer on intervention products is planned for future publication.

Combination products both **measure** and **intervene**. For example, continuous glucose monitors (CGMs) for diabetics share patient data automatically with their doctor’s office using a companion app. The level of human involvement may vary in the cycle between measurement and intervention – say, when a doctor diagnoses an abnormal heart condition

from an electrocardiogram (EKG) reading off a smartphone. Over time, this cycle may become more **closed-loop**, with less need for human intervention in response to routine changes. More recently, the development of the “artificial pancreas” has combined the CGM with an insulin pump and a computer-controlled algorithm that allows the system to automatically adjust the delivery of insulin to reduce high blood glucose levels (hyperglycemia) and minimize the incidence of low blood glucose (hypoglycemia) [7].

How Does Digital Medicine Differ from “Digital Wellness” and “Digital Health”?

Similar to the way in which “wellness” products differ from those used in medicine, “**digital wellness**” differs from digital medicine. We use “digital wellness” to describe products that consumers use to measure physical activity or sleep quality – things that might influence their personal well-being. Digital wellness products may include apps or wearable sensors (e.g., Fitbit, Oura ring). Digital wellness products are intended to be consumer-facing rather than used in clinical care as these products often lack evidence necessary to support the medical use of the information they produce.

There are times when it may be appropriate to use “wellness” or “consumer-grade” tools for measurement in clinical research. For example, using an accelerometer manufactured for the consumer market to measure physical activity among research participants enrolled in a clinical trial is common. However, this would require a body of evidence to support this use (see Part III on verification and validation), meaning that high-quality consumer tools can be digital medicine tools as long as they have a reasonable level of evidence behind the measurements instance.

Digital medicine product manufacturers commit to undergo rigorous randomized, controlled clinical studies for their products. Unlike digital products that measure the less well defined concept of wellness, digital medicine products demonstrate success in clinical trials [8]. In this primer, we outline digital products that measure and intervene in all areas of the practice of medicine, extending to and including behavioral health, public health, and population health management.

We have decided not to use the term “digital health” in this primer. While it is one of the buzziest catchphrases in the industry today, it has been so broadly used and misinterpreted that it has no real meaning. Instead, we use **digital medicine** as the term to describe **evidence-based** digital products that measure and intervene, including those intended for health promotion and disease prevention. Digital medicine products are evidence-based tools that support health research and the practice of medicine. Digital medicine describes this broad, evidence-based field and does not refer to the narrow use of the term “medicine,” which is sometimes interpreted as the drug (“medicine”) that is administered to the patient.

Algorithms, Machine Learning, and Artificial Intelligence – Oh My!

The recent explosion of machine learning and artificial intelligence methods, driven in large part by the availability of massive datasets and inexpensive computation, has played an important role in enabling digital medicine products [9]. Whereas traditional health measures represent a snapshot in time – a lab value, a diagnostic image, a blood pressure reading, or a note in a medical record – connected digital devices offer a longitudinal and highly personalized window into human health.

A key component of these systems is the transformation of raw physiological or environmental signals into health indicators that can be used to monitor and predict aspects of health and disease. These data (e.g., from a sensor) are processed, transformed and used to build computational models whose output represent the health indicators of interest. Computational approaches range from simple statistical models like linear regression, to signal processing methods like the Fourier transform, to time series analysis like additive regression

models, or machine learning methods like support vector machines or convolutional neural networks.

For example, an algorithm is required to transform the raw data from a 3-axis accelerometer into the more widely usable health indicator of step counts. There are a variety of different approaches to this task, yielding a variety of different performance characteristics [10]. Importantly, the more examples of real-world walking that the algorithm has access to – by people of different shapes and sizes, under different conditions – the greater the opportunity to improve the accuracy of the model.

How Are Digital Measurements in Medicine Being Used Today?

Some digital measurements are already well-established in routine clinical care, like using ambulatory EKG monitoring to detect arrhythmias in cardiac patients [11]. Similarly, remotely monitoring patients with implanted heart devices allows doctors to better follow their cardiac patients, detecting abnormal heart rhythms and problems with the device sooner.

Digital measures are also used in clinical research to better monitor patients and more efficiently assess safety and efficacy. For example, in-hospital ambulatory cardiac monitoring has existed for many years, enabling real-time monitoring of EKG signals. Similarly, portable EKG technologies have also existed, these recorded signals for later analysis. The digital medicine solutions for cardiac monitoring include non-obtrusive patch-based cardiac monitors that may be worn for days at a time, while ambulatory, and remote from the hospital while sending real time signals.

Across therapeutic areas and technologies, digital medicine solutions can solve weaknesses of existing solutions, and can come to market with more patient-friendly packages.

Examples of Digital Measures in Clinical Care

- **Recovery, performance, and treatment selection:** In patients recovering from orthopedic surgery, app-enabled wearable sensors are increasingly being used during rehabilitation. Digital measurements, such as range of motion and step count, allow remote monitoring of a patient's progress. More sophisticated measurements can monitor in real-time if a patient is doing their rehab exercises.
- **Real-time safety monitoring:** Digital fall detection systems allow for of elderly and frail populations. Such monitoring often relies on either wearable sensors, cameras, motion sensors, microphones, and/or floor sensors.
- **Treatment adherence:** One of the thorniest problems in routine clinical care can now be measured under limited circumstances via an ingestible sensor embedded in a medication that, when it interacts with stomach acid, transmits to a patch sensor worn over the abdomen, monitoring when a pill was taken. Abilify MyCite is the only drug combined with a digital ingestion tracking system to be approved in the US [12]. Other innovations, like “smart packs,” integrate sensors into the packaging of medicines to record when the drug was administered and deliver automatic reminders to take a medication.
- **Multimodal data integration:** By combining remotely captured data into EMRs, personal health records, patient portals, and clinical data repositories, innovators hope to improve clinical decision-making and support data-driven medicine [13].

Examples of Digital Measures in Clinical Research

- Data collected from remote sensors could be used as a **novel endpoint** for hard-to-measure conditions like Parkinson's [14, 15].
- Digital measures are being used to assess **medication adherence** in clinical trials using smart blister packages for investigational drugs.

Box 1. An example of how digital measures can improve screening in clinical trials

In oncology trials, one key inclusion criterion is performance status, an assessment of the extent to which potential participants' disease affects their ability to do activities of daily life, which is considered to be subjective and difficult to assess accurately [17]. A real-time digital measurement of performance status could reduce the variability of this assessment in trials and help ensure enrollment of the intended patient as a research participant.

- Continuous digital measures may allow for the detection of **safety events** that would otherwise go unrecorded. For example, a wearable cardiac monitor can help reveal arrhythmia in research participants during trials of stimulant use in people with ADHD.
- Digital measures may enable more objective and precise **screening for inclusion/exclusion** in a clinical trial. This could expand the pool of eligible research participants, increase diversity of a trial population, and decrease attrition between **evaluation** and **enrollment** by returning information to researchers faster.
- Digital measures may inform better decisions about whether to progress a drug from early phase trials to later, larger, and more costly trials. These are known as “**go/no go**” **decisions**. Digital measures may be particularly important to inform these decisions where current measures are subjective and/or where there is a high failure rate. For example, in Alzheimer's disease, digital cognitive assessments that afford more sensitive and frequent monitoring, but are not endorsed by health authorities yet, could enable better decision-making about which treatments to advance to the next phase of clinical development (Box 1) [16].

Articles reviewing current and prospective wearable technologies and their progress toward clinical application and the use of medical technology in the home provide additional examples of measurement in digital medicine for readers [11, 18].

In summary, the field of digital medicine applies the same rigor to the selection, development, or use of digital technology for measurement and intervention that is applied to other areas of medicine.

Part II: Security, Ethical, Legal, and Regulatory Considerations when Adopting Digital Medicine in Research and Care

In this section, we outline security, ethical, legal, and regulatory considerations when adopting digital medicine technologies in clinical research and routine care (Fig. 2).

Where Does Digital Medicine Fit in within Clinical Research and Care?

Clinical care is familiar to most readers from their own experiences with doctors, hospitals and other parts of the healthcare system. Its primary purpose has historically been to address health problems, and has long been grounded in the interaction between a patient and a healthcare provider. There has been varying progression of healthcare activities toward preventative care and maintenance of wellness. With the introduction of connected technologies, there have also been attempts to move healthcare activities into the home decreasing the need for face-to-face interactions with providers. Clinical care activities include a wide range of diagnostic and treatment processes and procedures such as:

- Real-time monitoring such as the use of continuous blood glucose sensors
- Tools for medical adherence such as smart apps and pill dispensaries
- Physical rehabilitation tools such as digital activity trackers.

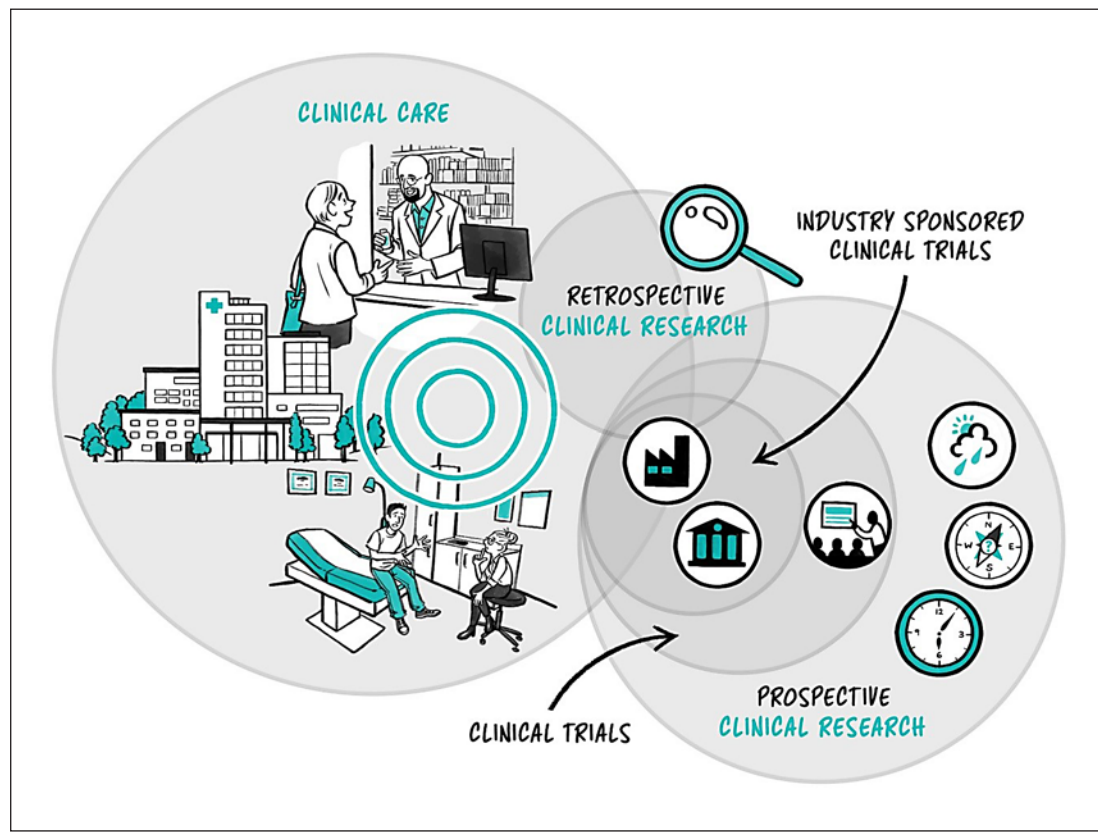


Fig. 2. The clinical landscape. The healthcare landscape can be broadly split into premarket clinical research and postmarket clinical care.

Clinical research may include some of the same activities as clinical care, but the primary purpose of clinical research is to develop a better understanding of factors influencing health and illness in people. The federal regulations define research as a “systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge” [19]. When a person (e.g., patient or healthy individual) volunteers to enroll in clinical research, they are called a **research participant**. There are rules and guidance that must be followed when conducting clinical research to make sure that research participants are protected from undue risks of harm. Clinical research comes in two broad subsets:

- In **interventional studies**, participants receive some form of treatment, education or support (Box 2). **Clinical trials** are a subset of interventional studies designed to evaluate the safety and efficacy of an intervention.
- In **non-interventional studies**, participants do not receive an intervention. Non-interventional studies include observational, exploratory, survey, case-control, cohort, and correlational studies. Computational studies that use existing data sources to build predictive models fall into this category.

Digital Measurement in Clinical Care: Outside the Clinic Walls

Like any other medical tool, at-home monitoring technologies need to prove their worth. Developers, working with researchers and other experts, must demonstrate that these tools produce clinically meaningful information that leads to clinically meaningful improvements in care, processes, and outcomes.

Box 2. Interventional studies

In interventional studies, participants are typically randomized at enrolment to either receive the investigational intervention (experimental arm) or the placebo/current standard of care (control arm). Comparing how participants in these two groups respond allows us to understand the safety and efficacy of the intervention.

Digital measurement in medicine will not replace clinics or clinicians entirely, nor would we want them to. The delivery of clinical care such as intravenous drugs or surgery, and the value that patients place in their relationship with their provider, cannot be replaced by digital tools. Nonetheless, when used appropriately, digital measurements can improve care by giving clinicians more complete information. Also, transferring some practices out of the clinic and into patients' regular lives, for example passively measuring sleep quality with wearables instead of requiring overnight stays in clinics, can enhance access to care and reduce cost.

Continuous at-home monitoring also raises a new set of practical issues: *Who will monitor the data? Who will be responsible for acting on it if it indicates a need for action? How will providers be compensated for these tasks?* Although organizations like Clinical Trials Transformation Initiative (CTTI) have made inroads in addressing the first two questions, the field will need to address these issues and adopt consensus solutions for these tools to be truly integrated into clinical care [20].

A defining moment for any medical product, whether drug or device, is when the product goes to market. From this perspective, the industry splits into “**premarket**” **research activities**, drug and device development in the life sciences and biotech, and “**postmarket**” **commercial activities**, where the products are used in clinical applications like in the hospital. Often, government regulators like the US Food and Drug Administration (FDA) or the Office of Human Research Protection (OHRP) are the gatekeepers between what is considered research (premarket) and what is part of standard of care and commercial activities (postmarket).

National governments are responsible for establishing national medicines and medical device standards and regulatory authorities that determine what claims product manufacturers can make when they go to market in that country [21]. As of 2015, 121 of the 194 members of the World Health Organization had a national regulatory authority responsible for implementing and enforcing product regulations specific to medical devices [22]. For example, in the United States, the FDA serves this function. Across the Atlantic, this oversight is provided by the **European Medicines Agency (EMA)** and in Japan, the **Pharmaceuticals and Medical Device Agency (PMDA)**.

Regulatory Considerations

There are countless articles and books that discuss regulatory considerations associated with medical product development. We will keep this section brief and provide key concepts and frameworks to consider.

It is essential to understand that regulatory agencies like the FDA regulate medical products (like drugs and diagnostic devices) but not the practice of medical care. Structurally, the FDA has six centers, with three most relevant to digital medicine developers:

- 1 Center For Drug Evaluation and Research History (**CDER**)
- 2 Center for Biologics Evaluation and Research (**CBER**)
- 3 Center for Devices and Radiological Health (**CDRH**) [23].

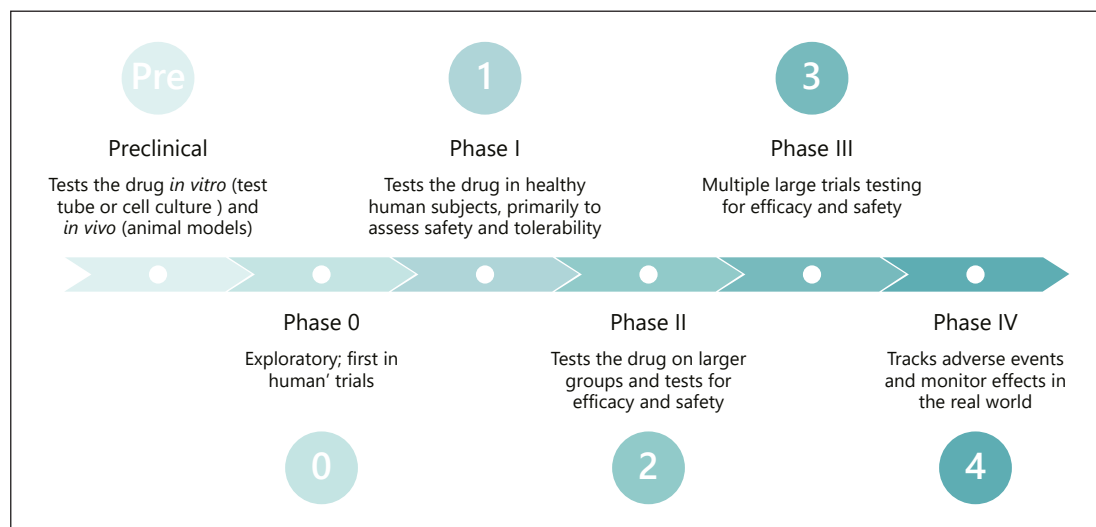


Fig. 3. The phases of clinical trial research. Clinical trials pass through a series of phases as the trial sponsor gains more evidence around the investigational drug or biologic. Preclinical studies are often conducted in cell and animal models (e.g., on mice), and then are slowly expanded into humans. First in healthy humans in small numbers to test safety, and then to larger sets of humans who have the condition in question to test both safety and efficacy.

Drugs and Biologics

Clinical trials for drug and biologic development are organized across phases (Fig. 3), and the results of these trials are reported to CDER or CBER, respectively.

Preclinical studies test the drug *in vitro* (test tube or cell culture) and *in vivo* (animal) models.

Phase 0 started as an informal stage designation that companies used to describe non-drug studies that are exploratory to prepare for the upcoming or ongoing drug research. At this phase, methods of measurement, or specialized techniques may be tested without the risks associated with administering an investigational medicine. In 2006, this phase was formalized by FDA guidance to include studies that use tiny doses of a drug (<1% of expected therapeutic dose) in healthy volunteers to determine if the chemical properties of the drug warrant further development [24].

Phase I “first in human” trials test the drugs in healthy human participants. The goal of this phase is to assess the safety and tolerability of the drug by starting with low doses on a small number of healthy people (e.g., <15 participants) and progressively increasing the doses toward the expected effective concentration. Safety is monitored and measured in all subsequent phases as well. These studies are often conducted in highly controlled and specialized inpatient clinics.

Phase II trials enroll larger groups of patients with the medical condition of interest, and begin to test for efficacy, while attempting to establish dose and frequency schedules for the final drug product.

Phase III is most often the “pivotal” stage of testing in which definitive evidence of efficacy and safety must be developed in multiple large trials. The endpoints of these trials ultimately serve as the evidence for the **label claims** that regulatory agencies allow pharmaceutical companies to use when marketing a drug.

Phase IV trials, also known as **observational** or **postmarketing surveillance** are a hybrid of research and clinical care. In these trials, the drug is already licensed for use and is

Box 3. Real-world data

Clinical trials in all phases are increasingly including data captured during routine clinical care such as electronic healthcare records and claims data. This is driven in large part by the passage of the 21st Century Cures Act which tasked the FDA with considering this type of data during the approval of new medical products. Clinical trial designs leveraging this routinely captured data will be the focus of a future digital medicine primer and will include discussion of how this digital data can improve trial success, yield more generalizable findings, and support novel trial designs such as synthetic control arms.

being prescribed to patients. The purpose of Phase IV trials is to monitor the effects of these new therapies in order to identify and evaluate previously unreported adverse reactions. Phase IV trials allow sponsors (biopharma and device manufacturers) to see how the product is performing in the “real world.” These trials also provide an opportunity to test the therapy in new demographics and find new markets, often resulting in **label expansion**, where the sponsor can make claims that the drug works for additional types of patients/diseases beyond the original use (Box 3).

Novel digital tools are being adopted at different rates in different stages of clinical trials, most likely because different trial stages are associated with different **levels of risk** to the sponsor. Phase III is an unlikely place to see novel measurements of any kind, as disrupting a large complex trial and risking the primary endpoint(s) could be expensive and harmful to the development process. Whereas, implementing exploratory efficacy measures in a small, early stage safety trial may be inexpensive and introduce minimal risk to the primary endpoint(s). Sponsors are now deploying digital tools in Phase I, II, and III trials as digital measurements need to be relatively consistent in early stages of trials to develop the necessary evidence both for internal decision making and regulatory approval.

The regulatory terms that describe tools, methods, materials, or measures that can potentially facilitate the medical product’s development are **drug development tools (DDT)** or **medical device development tools (MDDT)** (“tools” are different from “devices” at the FDA, which we will discuss later) [25, 26]. The FDA has also released a request for comments on **Prescription Drug-Use-Related Software (PDURS)** for software that is developed for use with prescription drugs (including biological drug products), which may include but is not limited to tracking drug ingestion, calculating the appropriate dose, sending reminders to take the drug, or providing information on how to use a drug [27].

There are also **combination products**, which contain both a drug and software that meets the definition of a device because of its function [28]. For example, Abilify MyCite, which is a drug-leg, drug-device combination product comprised of aripiprazole tablets embedded with a software-based ingestible event marker intended to track drug ingestion. Patients can opt to share these data with their healthcare providers or caregivers [27].

Medical Devices

Similar to the approval process in FDA’s Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER), medical devices in the **Center for Devices and Radiological Health (CDRH)** go through a process for clearance or approval. However, in the case of medical devices, regulators generally pay more attention to technical and design aspects of the product when considering the **safety** of medical devices – and particularly those that operate non-invasively may have more predictable effects than introducing a novel chemical to the body. All novel drugs are considered dangerous until proven otherwise, but medical device studies can be adjudicated as posing non-significant risk based on design criteria [29].

CDRH is often the point of contact for digital medicine developers who are building **software-as-a-medical device (SaMD)** products [30]. Although in other sections of this primer, we have used bold to indicate a key phrase, we will break tradition for this one term and use quotes around “device,” because a “device” is a **Term of Art** at the FDA, which means that it has a precise and specialized meaning. CDRH is responsible for regulating “devices” but not tools. As such, we limit our use of the term “device” in this primer to be consistent with the FDA’s definition for a “medical device” (See “Changes to Existing Medical Software Policies Resulting from Section 3060 of 21st Century Cures Act”) [31, 32].

For the US market, it is important to distinguish that the FDA does not regulate what the product actually does, but rather what an organization **claims** the product does.

For instance, let us say Product A and Product B are exactly the same mobile sensor technology, i.e. the same hardware, firmware, and software/algorithm that produce a measurement. If Product A states that the **intended use** of this measurement is for a wellness purpose, it likely is not regulated. If Product B says the intended use of this measurement is to make a diagnosis, then it would be considered a “device” and regulated by the FDA. This means that the exact same product can be developed and marketed either as a “device” (and thus, regulated) or not a “device” (and unregulated) simply through a change of words, and no change in hardware or code. For example, at the time of this publication, a Fitbit is not regulated by the FDA as it does not claim to serve a medical purpose. Therefore, a Fitbit is considered to be a digital measurement tool or a mobile sensor technology but, not a “device.”

Put another way: asking “is my digital product a medical device?” is not the most useful question. A better question would be about the **intended use** of the product (i.e., is the organization making a **medical device claim**?). Generally, answering this question is not easy, which is why many software manufacturers will spend millions of dollars on regulatory consultants. The FDA has an open-door policy where it encourages organizations to come early and often during product development. This is why it is good practice to initiate the conversation about regulatory designation of the product early. A good starting point is with FDA’s Division of Industry and Consumer Education (DICE) [33].

As tools develop **multiple functions** (e.g., can measure step count, and heart rate, and tremor and can be tailored to specific populations), these digital measurement technologies may be used for either medical product development or commercial clinical care activities [34]. Whether the software is a “device” is ultimately determined by a regulatory body and likely will depend upon the software’s intended functions.

Medical devices can presumptively be classified by risk profile, currently defined by the FDA as Class I, II, or III in order of increasing risk; **Class I devices** require little safety testing. Today, around 50% of medical devices fall under this category, and 95% of these are exempt from the regulatory process [35, 36].

Devices that perform a similar function to an existing device on the market (a **predicate**) can be approved simply by demonstrating that it is at least as effective and no more dangerous than the existing device. This pathway is called **510(k) clearance**. As devices increase in risk – meaning they have the potential to cause harm either by a malfunctioning or by providing bad information – there becomes more of a burden on the manufacturer to demonstrate safety and efficacy both from a technical perspective and in controlled human trials.

Of note, many medical device companies, including digital medicine developers, will bypass Class I and strive to get their devices categorized as at least **Class II**, because this category is generally the lowest risk class that is also covered by insurance, enabling greater access. Notably, Apple’s first FDA-cleared product, an ECG **over-the-counter (OTC)** (e.g., non-prescription) app, was categorized as Class II [37].

Class III devices pose a high risk to the patient and/or user (e.g., they sustain or support life, are implanted, or present potential unreasonable risk of illness or injury). These types of

Table 1. Regulatory pathways for device development

| Regulatory pathway | 510k | De novo | Premarket approval |
|----------------------|---|--|---|
| Product risk levels | Class I and II | Class I and II | Class III |
| FDA decision type | Cleared | Granted | Approved |
| Requires a predicate | Yes | No | No |
| Decision criteria | Product demonstrates “substantial equivalence” to a predicate (e.g., no independent assessment of the product required) | Probable benefits of the product outweigh probable risks | Requires independent assessment of the product’s safety and effectiveness |

devices represent around 10% of devices regulated by CDRH [36]. Implantable pacemakers and breast implants are examples of Class III devices.

Historically, the focus of CDRH relied heavily on a concept called a “**predicate**,” which is a legally marketed device (e.g., already on the market) to which a new device would claim equivalence. Whether or not a proposed device has a predicate impacts the regulatory pathway the device can use. For instance, a “**de novo**” **classification** does not need to have an existing device on the market to compare to.

Notably, in the fall of 2018, the FDA made a sweeping announcement that it is evaluating a rehaul to de-emphasize its predicate process and while this is probably a positive move to better treat patients, it will take some years to move away from the current system [38].

Many people do not realize that there is also a difference between **FDA-approved** and **FDA-cleared**. “Approved” indicates that the device successfully completed an **FDA Premarket Approval (PMA)**, which evaluates the safety and effectiveness of Class III high-risk products. “Cleared” indicates that the device successfully completed a **510(k)** pathway, which are for lower risk-level products (Table 1).

It is a herculean task to comb through existing FDA “guidances with digital health content” [39]. A good starting point includes the guidances on General Wellness: Policy for Low Risk Devices, **Mobile Medical Applications (MMA)**, **Software as a Medical Device (SaMD)**, and Clinical and Patient Decision Support (CDS and PDS) Software [40–43]. The FDA has been taking a forward-looking stance on how to handle digital products, including machine learning and algorithms, and streamline the regulatory process.

For instance, the agency is piloting a Software Precertification (**Pre-Cert**) Program with companies like Apple, Fitbit, and Samsung participating [44]. This program would allow software manufacturers a more streamlined review process, making it easier to release new software versions to market if the organization is precertified. The Pre-Cert program draws heavily on the **International Medical Device Regulators Forum (IMDRF)** definitions and categories for **software-as-a-medical device (SaMD)** [45]. SaMDs (e.g., software-only products like apps and algorithms decoupled from a hardware component) may be subject to more flexible regulations than **software-in-a-medical device (SiMD)** (e.g., traditional software contained within a pacemaker).

The **21st Century Cures Act** (Cure’s Act), signed into law on December 13, 2016, amended the definition of “device” in the Food, Drug and Cosmetic Act to exclude certain software functions, including some described in many existing guidance documents. FDA has been assessing how to revise its guidances to represent current thinking on this topic. There has been a recent trend to allow more digital products to go straight to market. For instance, the Cure’s Act made clearer distinctions as to what is considered a regulated medical device, versus a wellness product or a digital technology that is not a “device” (e.g., an **electronic health**

record, EHR) [46]. The distinction between those digital products/technologies that are considered medical devices, and those that are not, is a hazy one.

Determining the nature of a digital product is especially challenging because the FDA has other mechanisms like **enforcement discretion** where the FDA may determine that the product is a “device,” but chooses to not regulate it [47]. As these decisions are continuously evolving, some helpful resources to navigate the area include the FTC Mobile Health Apps Interactive Tool and the FDA Division of Industry and Consumer Education (DICE) [48, 33]. There are also papers that draw a comparison of European and US approval processes (e.g., mapping the EU CE Mark to the FDA framework) [49].

In Europe, a working group led by NHS England has developed “Evidence Standards Framework for Digital Health Technologies” to make it easier for innovators and regulators to define what “good” looks like within digital medicine [50]. There are many groups across the world working towards a more streamlined vision. The Digital Medicine Society is developing a resource on their website (www.dimesociety.org) to keep track of the different standards, papers, and frameworks [4].

Although regulatory authorities have the final say as to whether the digital product is a medical “device,” the organization that develops and markets the product can make many choices that influence the likelihood of being classified as a medical device. For example, organizations choose what claims to make about the product, how much evidence to gather to support those claims, and which markets to enter (and subsequently, which regulatory bodies to be regulated by).

The downstream consequences of these decisions include who can access the product, under what circumstances and for what reasons, and who is likely to pay for such access. Talking with the appropriate regulatory authority early and often is important during the product development process; it will minimize surprises and develop a forward-thinking regulatory strategy.

Put simply: CDRH is primarily concerned with whether a digital medicine tool, including both hardware and software, is safe to use and accurate for measuring what it claims to measure. If the manufacturer of that system does not claim it has a medical use (e.g., diagnostic, monitoring), they will not be regulated by CDRH. Agencies that evaluate new medicines, like CDER and CBER, care about whether the observation being made by a digital tool (concept of interest) is valid for the way it is being used in regulated research (context of use).

Table 2 is a “cheat-sheet” of the primary pathways to market through CDRH for a software product.

Ethical Considerations

As more digital tools are deployed in health research and care settings, new questions emerge about how to use them responsibly and ethically. In this section, we introduce a few terms and describe:

- 1 Ethical principles and how we are collectively responsible for shaping ethical practices in digital medicine.
- 2 Regulatory review processes in place to protect research participants.
- 3 Tips and resources for doing research with digital medicine technologies.

Ethical Principles and Our Responsibilities

Anyone developing and/or testing a digital tool for use in disease prevention and treatment should be aware of the regulatory requirement to obtain **Institutional Review Board (IRB)** approval prospectively when involving people as research participants. The development of the IRB peer review process stemmed from egregious acts whereby researchers disregarded the rights and welfare of research participants. One example is

Table 2. “If my software product is regulated by the FDA, how do I bring it to market?”

| Go-to-market strategy | Risk classification options | Is the product making a “device” claim (e.g., what’s the intended use)? | Can the manufacturer bring the product straight to market? | Does the manufacturer need to “register and list”? | Does the FDA review the product? | Is the manufacturer required to submit postmarket information (e.g., medical device reporting)? | Does the manufacturer need to pass a pre-cert excellence appraisal? | Is this pathway “live”? | Related documents |
|---|--|---|--|--|----------------------------------|---|---|-------------------------|--|
| The product is NOT a device, and is not regulated by the FDA | N/A, not a device | No | Yes | No | No | No | No | Yes | Mobile health apps interactive tool (HHS, ONC, OCR, and FDA): https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool |
| The product IS a device, but the FDA will exercise “ enforcement discretion ” and will not regulate it | Class I or II | Yes | Yes | No | No | No | No | Yes | Examples of mobile apps for which the FDA will exercise enforcement discretion: https://www.fda.gov/medicaldevices/digitalhealth/mobilemedicalapplications/ucm368744.htm |
| The product IS a device and is exempt from review by the FDA | Class I or II | Yes | Yes | Yes | No | Yes | No | Yes | Class I/II exemptions: https://www.fda.gov/medicaldevices/deviceregulationandguidance/overview/classifyyourdevice/ucm051549.htm |
| The product IS a device and will be reviewed under the FDA Pre-Cert program | Class I, II, or III (or will fall under newly proposed IMDRP risk determination) | Yes | No | Yes | Yes | Yes | Yes | No (In pilot) | Digital health software precertification (Pre-Cert) program: https://www.fda.gov/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/default.htm |
| The product IS a device and will be reviewed under one of the traditional CDRH pathways (e.g., 510k, de novo or PMA) | Class I, II, or III | Yes | No | Yes | Yes | Yes | No | Yes | FDA Division of Industry and Consumer Education (DICE): https://www.fda.gov/medicaldevices/deviceregulationandguidance/contact-divisionofindustryandconsumereducation/default.htm |

known as the “Tuskegee Study of Untreated Syphilis in the Negro Male” which was an observational study of the natural progression of syphilis initiated by the Public Health Services in 1932. At that time, there was no treatment for syphilis; however, after penicillin was developed, the study participants were not treated and the study continued for nearly 40 years. The **National Research Act** was passed in 1974 which involved creating a **National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research** with a goal of preventing future atrocities. It was this commission that required the formation of IRBs and also wrote the **Belmont Report**. The Belmont Report, which was published in 1979, describes three guiding principles of ethical biomedical and behavioral research: respect for persons, beneficence, and justice. In addition to the Belmont Report, federal regulations were introduced for research participant protections and adopted by several federal agencies and institutes. The regulations speak to basic protections in subpart A, now known as the **Common Rule**, as well as additional protections for vulnerable populations.

Here, we describe the ethical principles and follow with the how the regulations and Common Rule are implemented in practice. The three core principles of biomedical ethics described in the **Belmont Report** are at the core of research ethics and should be carefully considered during the study design phase and ethics review process [51].

Respect for Persons

This principle is demonstrated through the **informed consent** process, which occurs when a person is given the information needed to make a sound decision about whether to volunteer to willingly participate. How this information is conveyed is important because volunteering to be a research participant is different from, say, accepting **terms of service (ToS)** to access an app, or signing a consent form to obtain medical care. In the latter, a person will not be able to access the app if they do not accept the ToS nor will they receive medical care if they do not sign the medical consent form. Consent to participate in research is a choice that an individual can only make if presented with information in a setting conducive to good decision making. There can be no coercive actions (e.g., high incentive payments, free services) that may compromise an individual’s ability to volunteer. The informed consent process involves more than signing a form to document voluntary participation – it is the first of what may be many interactions between a participant and the research team and is part of developing a trusted relationship.

Another important consideration is a person’s technological literacy. For informed consent to be meaningful, participants will need to be “tech-literate” enough to understand the specifics of how their data will be obtained and used. Likewise, concerns about **privacy** are often raised when discussing the passive and ubiquitous nature of the tools used in digital medicine. Attitudes and preferences also vary across generations, with older adults preferring more privacy control compared to teens and young adults [52].

All these concerns suggest a need to better educate prospective participants – and yet, integrating these concepts into the consent process is not easy. Moving forward, this charge will require a commitment from the medical community to provide accessible public-facing educational modules. For example, one way to improve tech literacy might be to include a brief animation describing the difference between de-identification and anonymity when describing data sharing practices, or an illustration of what it means to store data in a cloud. A participant may also think that if the study team has access to their health data in real-time, 24/7, then that means someone is paying attention to them (which may not be the case). Clarifying these concepts is important and how best to do this will require experts in instructional design who can deliver creative educational content.

Beneficence

Beneficence is where an evaluation of probability and magnitude of potential harms are weighed against the possible benefits to a participant, the people they represent, and society. Determining risk of harm is a somewhat subjective process, yet worth breaking down. We need to consider potential sources of harm and try to quantify the likelihood of something going wrong as well as the consequences. For instance, if a technology collects and then transmits a study participant's location data to a publicly accessible or non-secure website, the likelihood of a loss of privacy is 100% for all users – yet the consequences will vary. For most people, these will be negligible, but for domestic abuse survivors or undocumented migrants, consequences might be severe. Thus, the same hazard presents a low risk for most, but high risk for some important others. Thinking about how to safeguard data and manage data sharing protocols is an important consideration when applying the principle of beneficence and one that researchers, IRBs and research participants need to think about carefully. When using third party commercial apps or measurement tools, it is critical that ToS and **End User License Agreements (EULAs)** be reviewed to ensure they do not introduce unnecessary risks to the end user be it a research participant or patient.

Other factors specific to risk assessment include the type of potential harm (e.g., physical, psychological, economic, social) as well as the duration and severity of harm to research participants. Research is inherently risky because we are learning something that is not yet known. Research participants are often told about risks as an odds ratio. For example, in studies that include a test for maximum oxygen uptake, participants are required to exercise to exhaustion. There is a 1 in 12,000 chance that a healthy individual doing this study will have a cardiac event that may lead to death. Because of this particular risk of harm, the research team can mitigate risk by having access to personnel and equipment used to treat a cardiac event. Having this information, an individual can decide whether they want to take that chance or not.

Within the domain of beneficence is the need for the digital measurement tools to be valid and reliable (see Part III). There is no potential benefit of knowledge gain if the study is poorly designed and the tools are not trustworthy. The old adage “**garbage in, garbage out**” (**GIGO**) is a serious concern and one that must be addressed by doing the appropriate studies early to ensure the products, regardless of whether there is a medical claim, are safe and produce useful data.

Justice

This principle focuses on the fair distribution of the benefits and burdens of research and recruitment protocols that are inclusive of those most likely to benefit from knowledge gained. With digital tools, we have the opportunity to reach a more diverse audience, including those in communities where health disparities are most prevalent. To do that requires that we design technologies that are accessible and, in some cases culturally tailored. With that in mind, including end users in the development process who represent a wide cross-section of our society is one way we can be responsive to the principle of justice.

For example, in a study designed to increase physical activity in refugee women, the researcher decided to use a wrist worn accelerometer to assess daily movement. The participants were given the sensor and shown how to use it. One week later, the researcher returned to gather the measurement tools and found that no data had been collected. Turns out a wrist-worn mobile technology was culturally unacceptable and drew unwanted attention to the women, so they did not wear it [53]. This story sheds light on the fact that while digital tools should improve access to health research and health care, they can also perpetuate disparities and prevent access if not well designed and deployed.

When Is an IRB/REB Needed?

Because of past harms associated with research involving human participants, there is an expectation, and in many cases, a regulatory requirement that an ethics committee review will take place in advance of the research commencing. In research supported by the US Department of **Health and Human Services (HHS)** or under FDA oversight, this review process is carried out by an **Institutional Review Board (IRB)** that is registered with the federal **Office for Human Research Protections (OHRP)** [54]. These regulations were initiated in 1974 as part of the National Research Act.

The involvement of an IRB in behavioral and biomedical research is common globally, though often by other names, such as a **research ethics board (REB)** or “research ethics committee.” In the US, an IRB is required to have a minimum of five people, including scientists, non-scientists and someone who is unaffiliated with the organization. An IRB can be a part of the organization conducting the research (i.e., medical center or university), or operate as an independent fee-for-service entity.

The IRB is responsible for reviewing **research** that involves **human participants** to evaluate the probability and magnitude of potential harms to research participants and weigh these risks against the potential benefits of knowledge to be gained. The IRB also reviews the proposed research to make sure that participants selected to participate represent those most likely to benefit from its results. Moreover, the IRB wants to make sure that people who are invited to participate in research have a good understanding of the study purpose and what they will be asked to do. This process of sharing study information with a prospective participant is called **informed consent** and is a central tenet of biomedical research.

Federal regulations and accepted ethical principles are in place to guide the conduct of “research” so that the science is rigorous and the participants are protected. In any research, an important step is to determine if people involved in the testing phase are considered to be human participants in the research. The federal regulations include definitions for what qualifies as “research” and “human subject” and addresses the responsibilities of the organization and research team.

Rather than go into detail here, we suggest that you contact the IRB affiliated with your organization to discuss the process for getting approval to test a product on humans. The IRB review and approval is usually needed if the activity is considered to be research and, the people involved with testing meet the definition of a human subject (e.g., clinical or non-clinical research). This is true regardless of whether the product is seeking FDA clearance or approval.

IRB Review Criteria and Pathways

Depending on the risk level (e.g., minimal or greater than minimal risk of harm) and type (e.g., psychological, physical, economic), there are three review pathways:

Exempt from the Common Rule

The exempt classification is appropriate if the study procedures pose no more than a minimal risk of harm (e.g., observation of public activities, survey of adults, analysis of existing data). The concept of minimal risk is defined in the federal regulations and means that the risk to a participant, whether it be physical or psychological, is no greater than what they encounter in normal daily life [19]. When a study is exempt, it means that the Common Rule does not apply to the research. Normally, the IRB makes the decision about whether a study meets the criteria for exemption.

Expedited Review

To qualify for an expedited review, the study procedures may not exceed minimal risk of harm and must align with one of the criteria described in the regulations [55]. For example, if the research involves collection of biosamples, noninvasive clinical testing (e.g., sensory acuity, moderate exercise by healthy volunteers), or examination of existing data like EHRs, it may be eligible for an expedited review. However, studies that are designed to carry out safety and efficacy testing of a medical device are probably not eligible for an expedited review and will be reviewed by a convened gathering of IRB members. The only difference between an Expedited and Convened Committee review is the number of people involved. An Expedited review can be conducted by a subset of the IRB membership which is usually the IRB chair and one other member.

Convened Committee

Any study that does not qualify for Exempt or Expedited review is evaluated by a convened group of IRB members. For research covered by the **Common Rule**, documentation of informed consent is required; however, sometimes that requirement can be waived.

Once the type of review is known, an IRB application is developed by the research team that includes a detailed research protocol and a draft of the informed consent document. The protocol will briefly describe the scientific literature that the study is building from, as well as the study aims, procedures, participant inclusion criteria, risks, benefits, risk management, data management, investigator qualifications, and informed consent details. The IRB will review this protocol application to evaluate whether the risks are appropriate in relation to the potential contribution to science and benefits to people like those who participate in the study.

Application of Ethical Principles

Researchers have applied these principles and relied on IRBs to help shape ethical research practices for nearly half a century. However, as digital products are increasingly used in health research and clinical care, all relevant stakeholders have a collective responsibility to think proactively about how to conduct digital health research ethically and responsibly. While IRB approval is an important step in the process for identifying and mitigating risk in studies, it is truly the responsibility of developers, researchers, and clinicians to be a part of the ethical decision-making process. Simply stated, we cannot outsource ethics and hope for the best.

Digital Medicine Ethics Resources

Of course, these regulations and ethical principles are sometimes difficult to put into practice. Because the use of digital methods is relatively new, accessing resources at the protocol development phase is important. Over the past few years, several initiatives have begun to address the **ethical, legal, and social implications (ELSI)** of emerging technologies. A few focus specifically on AI broadly (e.g., autonomous vehicles, facial recognition, city planning, future of work). AI initiatives presently underway (e.g., AI Now, A-100) are well-funded and global collaborative programs. Others addressing digital medicine technologies more specifically include the Connected and Open Research Ethics (CORE) initiative, MobileELSI research project, Sage Bionetworks and the CTTI, which are described below.

The CORE initiative, launched in 2015 at UC San Diego, is a learning “ethics” commons developed to support the digital medicine research community, including researchers and IRBs. The CORE features a Q&A Forum and a Resource Library with over 100 IRB-approved protocols and consent snippets that have been shared by 700+ members of the CORE Network. All resources are tagged for ease of access. For example, you can search the library to find

| | |
|---|---|
| <p>Knowme Networks Parent Permission/Youth Assent Form</p> <p> </p> <p>From Metzlab</p> <p>This study is about developing mobile sensors that can measure physical activity, sleep, heart rate, and other physical and emotional reactions. The s... (continued)</p> | <p>Category: Consent Form Resource Type: Document Institution Approved: USC IRB Posted on March 30, 2017</p> |
| <p>SMART: A Social and Mobile Weight Control Program for Young Adults</p> <p> </p> <p>Added by Lindsay Dillon</p> | <p>Category: Research Protocol (<i>Specific Aims</i>) Resource Type: Text Snippet Institution Approved: UCSD Posted on March 24, 2017</p> |

Fig. 4. Screenshot from the Connected and Open Research Ethics (CORE) Q&A Forum.

protocols that have used a digital tool in clinical research involving Latino middle-schoolers or breast cancer survivors (see Fig. 4).

In addition, the CORE is creating checklists to assist the community in proactive decision making. One checklist was inspired by a psychiatrist who had recommended to a patient that he use a mobile app to help with managing his daily patterns and mood. Upon closer inspection of the app’s ToS and privacy policy, the clinician realized she was inadvertently putting her patient at increased risk because the app was sharing their personal information with third parties. The checklist prompts researchers to think about ethics, privacy, risks and benefits, access and usability, and data management (see <http://thecore.ucsd.edu/dmchecklist/>) (Table 3) [56].

The MobileELSI project is led by investigators from Sage Bionetworks and the University of Louisville with a goal of understanding the scope of unregulated mobile health research to inform the development of a governance model. The increase in public access to technology has led to everyday citizens becoming involved self-experimentation, a form of “Citizen Science,” which is largely unregulated as it falls outside of traditional regulatory requirements. In addition, technology companies are increasingly involved in biomedical research. Neither are obligated to apply the federal regulations to protect research participants unless, of course, they are developing an FDA-covered product or are conducting federally funded research. The MobileELSI project will develop recommendations to guide the conduct of unregulated digital medicine research.

Sage Bionetworks and its governance team have led the charge in creating accessible informed consent templates for use on smartphone that enable digital medicine research. For example, the “Elements of Informed Consent” toolkit (Fig. 5) is available to researchers to help them think through developing an effective informed consent process on a mobile device [57].

Another source for guidance is the CTTI. CTTI has developed recommendations, resources, and practical solutions to facilitate responsible practices in mobile clinical trials [58].

Ethics when an IRB Review Is Not Required

When research involves retrospective analysis of existing data or prospective observation, testing or experimenting with people to generate “generalizable” knowledge, an IRB review is needed. Generalizable typically means that the results will be shared through peer-reviewed publication or presentations. The need for an IRB review is relatively clear in the world of premarket clinical trials, but the lines defining human research in the postmarket, commercial world have been less obvious. Some instances of **A/B testing** may be considered

Table 3. Excerpt from the Digital Health/Medicine Decision Making Checklist

| Domains | Ethical Principles | | | Researcher Responsibility | |
|--|--|---|--|------------------------------------|---|
| | Autonomy | Beneficence | Justice | Addressed in the research protocol | Addressed during the informed consent process |
| Ethical Principles Privacy Risks & Benefits Access & Usability Data Management | Actions demonstrate respect for the person | Actions involve comprehensive risk and benefit assessment | Actions demonstrate access to those who may benefit most | | |
| Privacy (respect for participants) | | | | | |
| Personal information collected is clearly stated | | | | Yes No Unsure | Yes No Unsure |
| What data are shared is specified | | | | Yes No Unsure | Yes No Unsure |
| With whom data are shared is stated | | | | Yes No Unsure | Yes No Unsure |
| Privacy Agreement – When a commercial product is used: | | | | | |

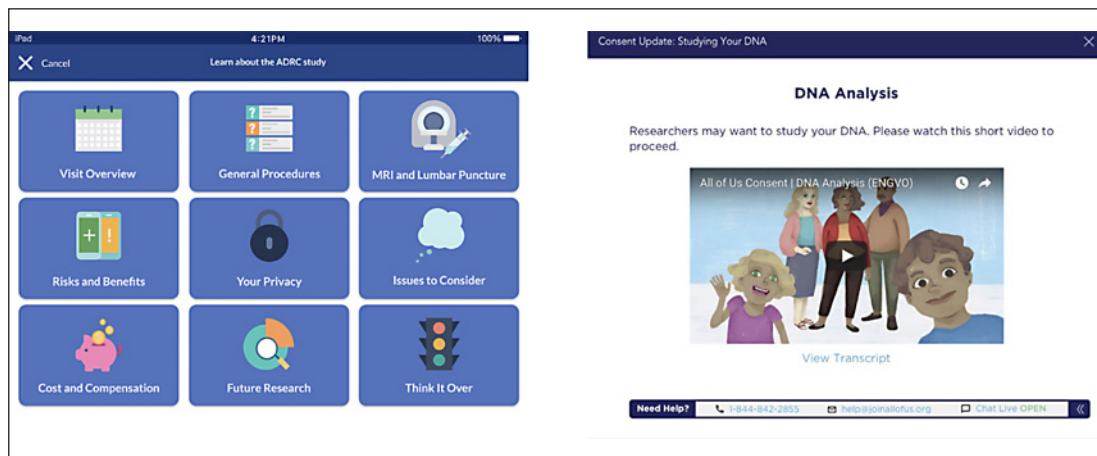


Fig. 5. Screenshot from Sage Bionetwork’s “Elements of Informed Consent” toolkit.

human subject research, depending on whether an organization intends to share knowledge broadly or use it internally to improve their product or service. For example, Facebook found itself in hot water in 2014 after testing different versions of its Newsfeed with users to study emotional contagion. Had the results been kept internal to Facebook for product improvement, it would have flown under the radar for needing an IRB consult. However, they published the study results to share knowledge produced with the public.

Sharing knowledge is believed by research ethicists and the scientific community to be a responsibility to society – which is certainly a good thing. In this case, though, many users

were outraged about being involved in research that they did not consent to. In effect, more than 800,000 Facebook users had become inadvertent research participants [59, 60].

The takeaway message here is that ToS and EULAs are not a substitute for informed consent [61]. People want the right to opt-in to being involved in biomedical research, and that is a clear call for respecting the ethical principle of “respect for persons.” Yet, when we are doing work that is technically not research, what is our ethical obligation? In software development, the way user data has been treated has an emerging history of malfeasance. This practice is likely due to the lack of universally agreed upon guidelines and standards. We strongly recommend that those involved in the developing digital medicine field adopt ethical principles to guide responsible practice when guidelines are lacking.

In response to the lack of guidelines and exploitation of consumer data, new regulations have emerged that speak to consent and privacy concerns. The **General Data Protection Regulations (GDPR)** was passed by the EU parliament in 2016 and took effect April 2018. The GDPR was designed to harmonize EU privacy laws, protect EU citizens’ data privacy, and change how organizations, regardless of where they are located, process and manage EU citizen data. An important change that the GDPR introduced was the need for companies to obtain explicit informed consent separate from a ToS or EULA. This shift from consumers being helpless data subjects to empowered actors in the digital data economy is moving to the US. In 2018, California passed the **California Consumer Privacy Act (CCPA)** which, when implemented in 2020, gives consumers control over their data and requires that companies like Facebook and Google explain what data they collect, what they do with it, and who it is shared with [62].

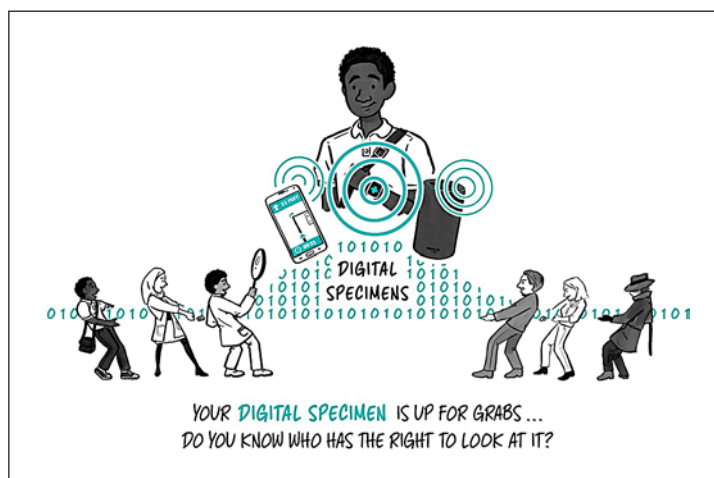
Prioritizing Data Rights and Governance

New digital tools, such as digital biomarkers, can capture an unprecedented amount of information about users, including fine-grained behavioral and physiological states. Many of these tools are non-invasive and collect data passively, which is certainly more convenient but also runs the risk that people do not understand how much of their digital footprint is being collected or shared [63]. Recently granted patents include a shopping cart that monitors your heart rate and Alexa’s new ability to apparently diagnose your cough [64]. Data collected from such technologies could be used by a doctor to make a clinical decision or by an insurer to approve or deny a claim [63]. Society needs to decide how to create systems that will deliver real benefits while protecting citizen privacy and safety [65].

For example, there is a lot of excitement in the healthcare community to use these tools in postmarket **monitoring**, or **surveillance**, to track metrics like safety monitoring and efficacy. Although many of these surveillance techniques in healthcare are still early, security researchers in the tech world are understandably cautious. Put simply: personalized medicine holds great promise for humanity, but it is not possible to have personalized medicine without some amount of “surveillance” – indeed, they go hand in hand. Now that de-identification gets more difficult with the vast amount of data generated for an individual, it is critical to understand who, what, and when does an entity get access to our data [66]. Health insurers and data brokers have been vacuuming up personal details on individuals, to create predictions on health costs based on race, marital status, whether you pay your bills on time or even buy plus-size clothing [67]. Similar data has been used to create “health risk scores” for the opioid crisis, determining who gets access to what types of care [68]. The biases in these types of algorithms have been well documented, exacerbating health disparities – and yet our society lacks clear regulatory interventions or punishment for misuse [69, 70].

We have all heard about the “**Internet of Things**” (IoT). What is coming next is the “**Internet of Bodies**” (IoB) – a network of smart devices that are attached to or inside our bodies, as defined by Prof. Andrea Matwyshyn [71]. Most of the digital tools we have discussed

Fig. 6. Digital specimens and social contracts. Our healthcare system has strong protections for a patients' biological specimens, such as a blood sample, but what about our "digital specimens"? With the increase in biometric surveillance from these tools, data rights and governance – who gets access to what data and when – becomes critical.



in this primer fit within the IoB paradigm. Matwyshyn argues that using the human body as a technology platform raises a number of legal and policy questions that regulators and judges need to prepare for.

Our healthcare system has strong protections for how to store and share a patient's biological specimens, such as a blood or genomic sample – but what about our **digital specimens**? With the increase in biometric surveillance from these tools, data rights and governance, who gets access to what data and when becomes critical [65].

EULAs and ToS, which gain consent one time upon sign-up, are not sufficient as a method to actually inform a person about how their health data – in the form of a digital specimen – will be protected. Our society needs better **social contracts** with tech platforms that have accessible and meaningful informed consent processes baked into the product itself and can be tailored to adapt as user preferences change over time (Fig. 6).

As the field of digital medicine advances, and indeed as medicine as a whole advances a process for creation and promotion of policies, we need standards to ensure that people are protected from known and unknown harms due to myriad computational and human factors, including both failures of knowledge and failures of intent on the part of developers.

Security and Compliance Considerations

Any organization that works with **personally identifiable information (PII)**, **personal health information (PHI)** and direct access to patients is at a high risk for cyber threats. Today with the rise of connected products, even a few vulnerable lines of code can have a profound impact on human life. Healthcare has seen a proliferation of vulnerabilities, particularly in connected technologies, many of which are life-critical: Johnson & Johnson's insulin pumps, St. Jude Medical's implantable cardiac devices, and the WannaCry ransomware attack, which infected 200,000 computers, many part of critical hospital infrastructure, across 150 countries [72–74]. Vulnerabilities in connected technologies can be exposed by either **black hat** or **white hat** hackers. Black hat refers to a style of breaking into networks for personal or financial gain, often illegally without the owner's permission. White hat hackers, or **security researchers**, perform a style of ethical hacking on mission-critical networks and will employ the policies of **coordinated disclosure** to the network owner if vulnerabilities are found [75].

A number of organizations have created protocols to prioritize risk levels of known vulnerabilities. For instance, MITRE, a nonprofit that operates research and development

centers sponsored by the federal government, created the **Common Vulnerabilities and Exposures (CVE)** program, to identify and catalog vulnerabilities in software or firmware into a free “dictionary” for organizations to improve their security [76]. Major agencies that are addressing healthcare cybersecurity include **National Institute of Standards and Technology (NIST)**, which has published a number of well-documented methodologies on how to quantitatively and qualitatively assess cyber risks, and the FDA, which has released a number of both premarket and postmarket guidances on cyber security best practices.

Researchers and developers should not count on others to implement critical basic protections, but have knowledge of their organizations policies and important contacts, such as the **chief information security officer (CISO)**. For those embarking on this journey, check out “A Cybersecurity Primer for Translational Research” [77]. Newcomers to the field often confuse the concepts of **compliance** and **security**. From the Cybersecurity Primer, “**Security** is the application of protections and management of risk posed by cyber threats. **Compliance** is typically a top-down mandate based on federal guidelines or law, whereas security is often managed bottom-up and is decentralized in most organizations” [77]. Compliance typically relates to documentation (e.g., for the **Health Insurance Portability and Accountability Act, HIPAA**), whereas security relates to how the technologies are updated, assessed, and used [77].

Most modern software is not written completely from scratch and includes common, off-the-shelf (COTS) components, modules, and libraries from both open-source and commercial teams. A tool to help manage potential vulnerabilities is called a “**software bill of materials (SBOM)**”, which is analogous to an ingredients list on food packaging and contains all the components in a given piece of software. The FDA has been investing more time and guidance around sharing SBOMs in both pre- and postmarket settings, and so have medical device makers like Philips and Siemens, and healthcare providers like NY Presbyterian and the Mayo Clinic [78–80].

As monitoring and surveillance tools become mainstream, it is critical to have secure and ethical checks and balances. For example, upon graduation from medical school, soon-to-be physicians take the Hippocratic Oath, a symbolic promise to provide care in the best interest of patients. As connected tools increasingly augment clinicians, a critical question emerges: should the manufacturers and adopters of these connected technologies be governed by the symbolic spirit of the Hippocratic Oath? [81].

Inspired by the traditional Hippocratic Oath, a number of security researchers from **I Am The Cavalry**, a grassroots organization with ties to **DEF CON**, an underground hacking conference, drafted “**The Hippocratic Oath for Connected Medical Devices (HOCMD)**” [82]. The Oath outlines a number of security and ethical principles, including “secure by design” and “resilience and containment” [81].

While the FDA has not called out the HOCMD by name, in the 3 years since the Oath was published, the FDA has incorporated elements from the five principles across the pre- and postmarket cybersecurity guidelines [83, 84]. The FDA has supported further collaboration between security researchers and connected device manufacturers through the agency-led **#WeHeartHackers** initiative, which launched in early 2019 [85].

Many government agencies support initiatives to improve security for medical connected technologies and **healthcare delivery organizations (HDOs)**. For instance, **Health and Human Services (HHS)**, sponsors the **Healthcare and Public Health Sector Coordinating Council (HSCC)** joint **Cybersecurity Working Group (CWG)**. The mission of the HSCC CWG is to collaborate with the HHS and other federal agencies by crafting and promoting the adoption of recommendations and guidance for policy, regulatory and market-driven strategies to facilitate collective mitigation of cybersecurity threats to the sector that affect patient safety, security, and privacy, and consequently, national confidence in the healthcare system [86].

Box 4. Types of measurements in clinical trials

| |
|--|
| There are multiple types of clinical measurements, like biomarkers, endpoints, and outcomes, which are related but take on different meanings in contexts outside of clinical trials. |
| Biomarker – a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. |
| Endpoint – an event or outcome that can be measured objectively to determine whether the intervention being studied is beneficial. |
| Clinical outcome – describes or reflects how an individual feels, functions or survives. |

In summary, as advances in technology enable digital tools to gather ever larger amounts of high-resolution personal health information, core principles of medical and research ethics must be integrated at every step, beginning in the design phase. Methods common to the consumer technology industry for obtaining agreement to corporate ToS, including privacy policies, are not sufficient or appropriate for obtaining informed consent from users, be they patients receiving care or participants in health research. The field of Digital Medicine must develop innovative ways of ensuring that the values of respect, privacy and trust are not lost in the pursuit of better data.

Part III: Categorizing Types of Digital Measurements

Are Digital Measures the Same Thing as “Digital Biomarkers”?

In some cases, yes. But not always. Determining the best term boils down to what you are measuring and for what purpose. Although we worked hard to minimize jargon in this paper, this section will have more technical terms because we strive to use the same language that regulators use to categorize types of measurements, and we want to arm developers with the right language and frameworks to work with regulatory bodies.

The FDA and NIH came together in 2016 to create the BEST (**B**iomarkers, **E**ndpoint**S**, and other **T**ools) glossary resource to help clarify terms in this confusing space [87]. **BEST** defines an **outcome** as any “measurable characteristic... that is influenced or affected by an individual’s baseline state or an intervention as in a clinical trial or other exposure.”

Let us start by reviewing differences between **biomarkers** and **clinical outcomes** (Box 4). The purpose of medicine is to improve health and reduce the risk of an early death. Outcomes are essential measures to determine whether the practice of medicine is working. Outcomes can be clinical outcomes or biomarkers.

A **clinical outcome** “describes or reflects how an individual **feels, functions or survives**” [87]. For example, “average gait speed” for walking bouts greater than a certain number of steps measured over multiple days outside the clinic may be a direct measure of mobility, which is one component of physical functioning.

A **biomarker**, on the other hand, is “a defined characteristic that is measured as an indicator of **normal biological processes**, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions” [87]. For example, a biomarker could be measuring hemoglobin A1c in diabetics to determine an individual’s average blood sugar level over the past several months. For biomarkers, think “signs,” not “symptoms,” which are more likely a clinical outcome. Most of the time, biomarker measures are not the things that are necessarily meaningful to patients – but the measurements might be accurate representations, predictors, or monitors of underlying biology and disease.

Table 4. Biomarker examples

| Biomarker type | Example |
|--|--|
| Susceptibility to or the risk of developing a condition | Gait characteristics (step-length during everyday walking) may be valid biomarkers to calculate fall risk in certain populations Using a combination of genetic variants to predict the onset of age-related macular degeneration |
| Diagnosis or health status monitoring | Presence or absence of P waves on an ECG trace obtained from a wearable/portable sensor as one of the inputs to a software algorithm to detect AFib that comes and goes (paroxysmal AFib) Tremor detected in a limb at rest with a wrist-worn wearable may be one marker among several to help detect and monitor early Parkinson's disease |
| Predicting whether an individual is more likely than others to experience a future change or clinical outcome | A breast cancer patient's HER2 status can be a predictive biomarker, when it is used to assess whether to treat with Herceptin The number of threshold-crossing events on an intrathoracic impedance signal measured by an implanted device may be a useful biomarker to enrich heart failure clinical studies with people more likely to experience clinical endpoints like hospitalization [107, 108] |
| Evaluating an individual's prognosis , or the likelihood that a disease recurs, progresses, or is cured, or that some other clinical event may happen in the future | Using a composite mobility measure to predict future need for long-term care |
| Monitoring the molecular effects of an intervention | Monitoring an antigen released by pancreatic cancer cells (CA 19-9) to judge a treatments efficacy Using a continuous glucose monitor to detect how a patient is reacting to insulin |
| Detecting a safety signal or adverse event | Using a simple accelerometer or gyroscope for fall detection |
| Precisely quantifying the time-course over which a dose of medicine has an effect on the body (its pharmacodynamics) | Sweat chloride may be used when evaluating patients with cystic fibrosis, to assess response to cystic fibrosis transmembrane regulator-potentiating agents |

In many circumstances, meaningful clinical outcomes such as survival or the occurrence a cardiovascular event may occur so infrequently that they are difficult or unethical to study directly. Biomarkers provide researchers interim evidence about the safety and efficacy of interventions while more definitive clinical data is collected. In some cases, it may be preferable to use established biomarkers as surrogate endpoints to reduce the risk of harm to patients [88]. Though biomarkers have intrinsic characteristics, they also have contextual ones. Depending on how a biomarker is used, it could have a relationship to a number of other medical concepts as shown in Table 4 (Fig. 7) [87].

Although traditionally most biomarkers fall into the modalities of molecular, fluid or imaging categories, more **digital biomarkers** are being developed [89]. A digital biomarker could be any of the seven BEST biomarker types [89]. The term **digital** refers to the method of collection as using sensors and computational tools, generally across multiple layers (e.g., a full stack) of hardware and software [90] (Fig. 8).

Clinical outcome assessments (COAs) are the instruments used to measure clinical outcomes, and include the instructions to participants, scoring models, and protocols for administration.

There are currently four recognized types of COAs (Table 5) [91].

Fig. 7. The BEST (Biomarkers, EndpointS, and other Tools) Framework. In 2016, the FDA and NIH collaborated to draft “Bio-markers, EndpointS, and other Tools (BEST),” which contains a description of seven types of bio-markers. All of these biomarkers could be measured using digital tools, which results in a digital biomarker.

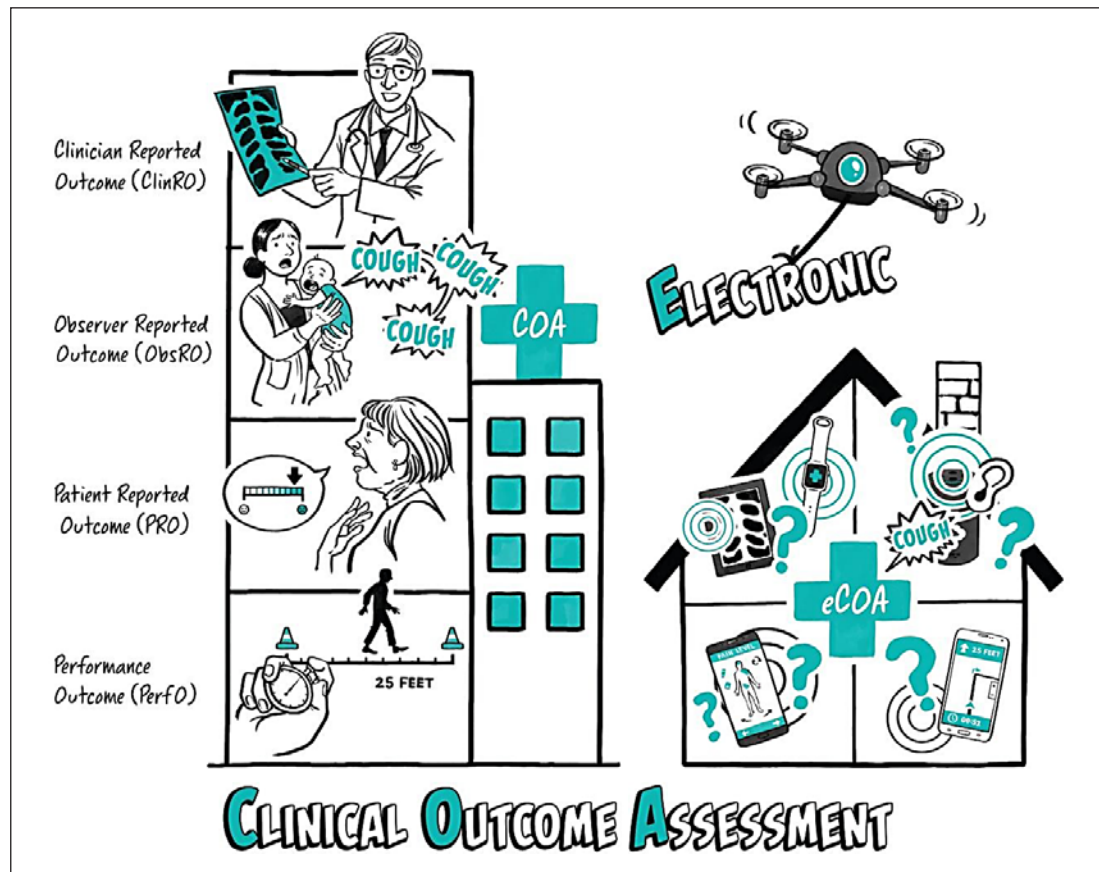
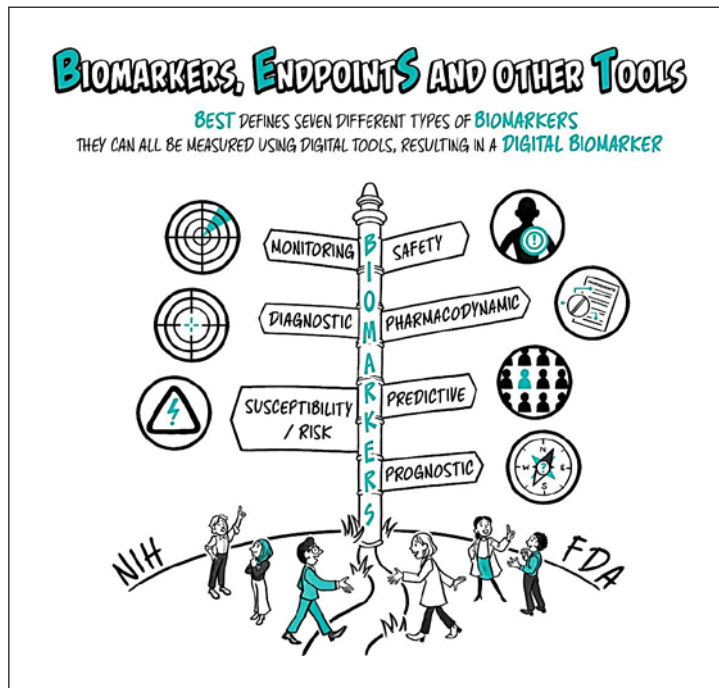


Fig. 8. Clinical outcome assessment. Clinical outcome assessments (COAs) are the instruments used to measure clinical outcomes. The FDA recognizes four types: ClinRO, ObsRO, PRO, and PerfO. If digitized, they are considered electronic clinical outcome assessments (eCOA).

Table 5. Clinical outcome assessment examples

| Clinical outcome assessment (COA) | Example |
|--|--|
| Clinician-reported outcome (ClinRO) – ClinROs are those COAs requiring clinical expertise | “Readings” are clearly defined results that are observed and reported in a dichotomous manner on the basis of clinicians’ judgment like the presence or absence of clinician-identified radiographic vertebral fractures [109] “Ratings” are categorical (either ordered or not) or continuous measures like those in Part III of the Unified Parkinson’s Disease Rating Scale or the Brief Psychiatric Rating Scale in mental disorders [109] “Clinician Global Assessments” (CGAs) are assessments based on a clinician’s overall judgment like the “clinician global impression” (CGI) or “clinician global impression of change” [109] |
| Observer-reported outcome (ObsRO) – ObsROs are assessments of how patients feel or function in their daily lives made by a non-expert third party (spouse, caregiver, parent, sibling, etc.) ObsROs are useful when the patient themselves may struggle to reliably assess their own symptoms and experiences (i.e. children or cognitively impaired patients) | A parent’s report of a child’s vomiting episodes A caregiver reporting a patient wincing through pain during activities when they cannot report this themselves |
| Patient-reported outcome (PRO) – PROs are assessments about how patients feel or function in their daily lives where the information is reported by the patient themselves, without interpretation or modification by someone else | Gastrointestinal Quality of Life instrument (GQLI) European Organization for Research and Treatment of Cancer QLQ-C30 (EORTC QLQ-C30) is a questionnaire that assesses the quality of life of patients with cancer The Impact of Weight on Quality of Life (IWQOL-Lite) assesses obesity-specific quality of life measures |
| Performance outcome (PerfO) – PerfOs are assessments of a task(s) performed by a patient following instructions given by a healthcare professional; performance outcomes require patient cooperation and motivation | Timed 25 foot walk test as a measure of gait speed Severe Impairment Battery as a measure of cognitive function |

The industry generally makes a distinction between a digitally collected COA and a non-digital one (e.g., a paper questionnaire) by putting an “e” for “electronic” in front of the acronym (e.g., PRO to ePRO, COA to eCOA). Interestingly, though, the lines between technology-based assessments and questionnaires have become blurry. If a wearable monitored your sleep overnight and then in the morning asks when you went to bed – was that evidence from a sensor or a questionnaire? As technologies continue to advance, there is increasing discussion around where digital measurement tools fit within this framework.

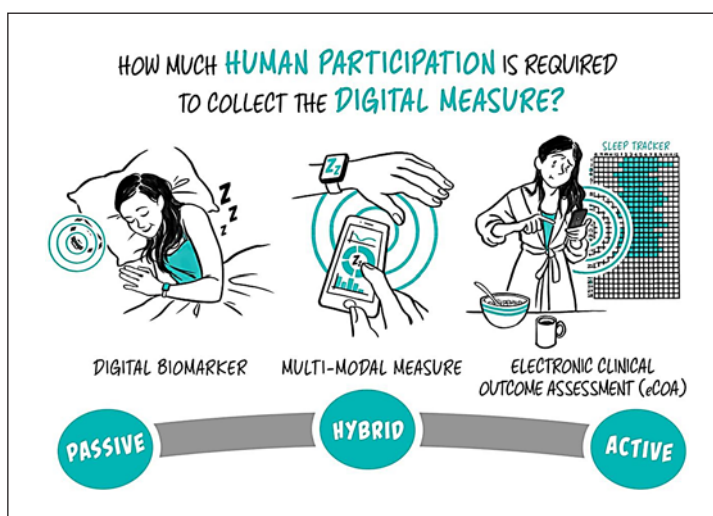
What if some of the human raters in the COA examples above were replaced with technology? For example, a medically informed algorithm that processes movement data from a wearable to rate ataxia.

The possibility of a fifth COA to describe digital clinical outcome assessments, and specifically those measurements made using technology, was proposed during a summer 2018 Public Workshop at the FDA, though no firm next step was defined [92].

Making human assessments digital takes far more work than translating a paper questionnaire into an app and putting an “e” in front of the tool (e.g., PRO to ePRO). eCOAs have unique properties that offer new ways to measure outcomes. What matters is whether the concept being measured is directly meaningful to patients.

Take for example, multiple sclerosis (MS). “Ability to go about my daily activities” is a meaningful aspect of health to MS patients.

Fig. 9. Human participation. Connected technologies can collect physiological and behavioral data. Some do so passively via sensors, while others (like sleep tracker surveys) may be more active in nature. Some are a hybrid where they will track some metrics algorithmically and then ask for confirmation.



- A **PRO** (or even **ePRO!**) could measure a person’s ability to perform activities of daily living through a self-reported survey. Self-reported measures require the person to reflect on and assess their own ability, leading to data that are meaningful to that individual, but potentially fraught with person-to-person variability, and subject to recall bias.
- A **PerfO** might measure the amount of time it takes a person to walk 25 yards. This is easy to measure and may be correlated with the capacity to perform typical daily activities. This approach has the benefit of not relying on patient recall, but it is an imperfect measure of the real outcome of interest – ability to perform daily living tasks in a natural environment.
- A **multimodal** product may continuously and passively assess total daily activity (mins), average daily walking speed (m/s), and number of sit-stand transitions per day (n). This is a much, much closer approximation of the meaningful aspect of health than the PerfO or even the “ePerfO,” but is not accompanied by the bias and subjectivity of the PRO or ePRO.

We can also combine **multimodal data** from sensors, questionnaires, and other clinical data (e.g., lab test, genomic tests) to create **composite measures** or “complex biomarkers” [91]. A composite measure consists of several individual measures that are combined to reach a single interpretive readout. For example, you could use sensor, keyboard, voice and speech data from a smartphone to construct a composite measure for cognition, and augment that measure over time with genetic data to make it more multimodal [93] (Fig. 9).

Over time, a useful distinction between metric types will be whether the measurement required action by a **human intermediary** to gather the data. At the **passive** end of the spectrum, human participation is minimal, and sensors simply capture data as individuals engage in daily activities like eating and sleeping. At the **active** end of the spectrum, more action by the intermediary is required. This could include a patient entering information into an electronic sleep diary or performing a task like a cognitive test. **Hybrids** of these two measures that use multimodal assessments and combine active and passive measures will also be valuable. An example is using actigraphy to passively measure when an individual fell asleep, coupled with an ePRO asking the individual to self-report the time.

To summarize, depending on the context in which the measurements are made and the data are interpreted, digital measures likely fit into many of the existing categories for medical



Fig. 10. Fit-for-purpose digital measure decision tree. Decisions about what to measure in a clinical trial are rooted in the research question being posed. Whether a fit-for-purpose digital measure exists to help to answer that question depends on both clinical and operational considerations that involve many stakeholders.

measurements (as biomarkers, clinical outcomes, etc.). In addition, technological advancement will present opportunities for new categories.

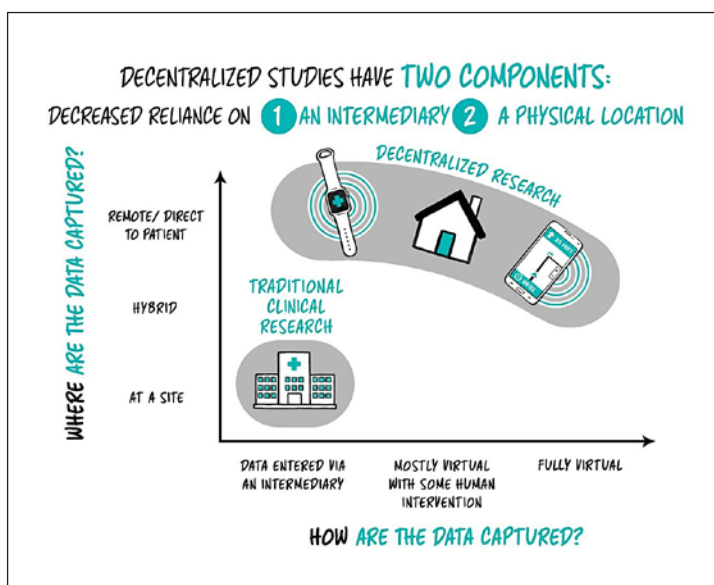
Why Is Measurement Important in a Clinical Trial?

The goal of any trial is to determine both the **safety** and **efficacy** of a new medical product. Measures within the trial process must demonstrate the product's safety and efficacy to regulators before it is allowed to be labeled and marketed for use by patients.

However, measurement in clinical trials does not only inform regulatory decision making. Early in the development of a new drug or novel medical device, a company will make business decisions about whether to advance their new product for further testing. Early intelligence is extremely valuable to biopharma companies, which face a USD 2 million revenue opportunity per day the drug is on the market (or not) [59]. Digital measures allow for the collection of data outside of the clinic, providing a more continuous stream of data points on whether the drug or device is working or not.

Additionally, measurement data from clinical trials informs reimbursement decisions, which impacts the value of the market. In countries like the US that depend on **third-party** (e.g., non-government) **payers**, insurance companies need evidence to decide *whether* to reimburse the manufacturer for their approved product and *at what price*. In countries with a **single-payer system**, often the decision about pricing coincides with the regulatory approvals process.

Fig. 11. Decentralized clinical trials. There are two components to consider to determine the level of decentralization in a clinical trial: whether the data are captured at a site or near the patient (e.g., location), and how the data are captured (e.g., manually or digitally).



Trial success is not correlated with number of measures. It is much more important to select the *right* measures – those that are the most informative regarding the product’s safety and efficacy – rather than the *most* measures. In fact, medical product manufacturers, regulators, patients, and ethics review boards all worry about burdening participants with excessive tasks, activities, and technologies. The ability of a measure to effectively and accurately operate **in the wild** (e.g., out of a patient’s home and across many geographies and languages) is also a concern. Medical product manufacturers are often reticent to assume even more risk – not just in their new product, but also in a novel measure – without a substantial body of evidence.

Decisions regarding the inclusion of digital measurement tools in clinical trials are complex, affect many stakeholders, and require extensive consideration of factors related to the **clinical implications** of the measure itself, along with considering the **operational aspects** of the measure and the potential effects on the trial design and on the participants (Fig. 10).

How Can Digital Measurement Change Clinical Trials?

Traditional clinical trials collect snippets of data when a participant visits the study site and represent a tiny snapshot of patients’ lived experience with a disease or condition. Yet researchers, industry sponsors, and regulators rely on this limited information to make life-or-death decisions and multibillion dollar investments.

Digital measurements will convert that snapshot into a movie with the ability to collect near continuous data outside the physical confines of the clinical environment, such as in a person’s home, using connected products, including smartphones, wearables, implantables, and ingestible devices and sensors. Digital tools enable new forms of research like **decentralized clinical trials (DCTs)**, which are conducted outside of the clinic to capture data about a study participant in their day to day life (Fig. 11).

DCTs have a number of potential benefits, including faster participant recruitment, improved participant retention in the trial, greater control and conveniences for participants, increased diversity (e.g., because it is easier to enroll in the first place), and trial results that are more generalizable [94, 95].

DCTs offer a way to make better informed decisions about the efficacy of new therapies. More sensitive, objective measures from digital technologies coupled with a greater density of information – continuously sampling multiple times a day, not just one a quarter – will help the industry fail faster and win more efficiently.

Two features of data collection determine how “decentralized” a clinical trial is [96]:

1 **Where are data collected?** (i.e., how dependent is the trial on a “site,” such as a clinic or hospital center?)

In traditional clinical trials, drugs, devices and therapies are administered in a clinic or research hospital. In newer **direct-to-patient** or **remote** trials, participant data is collected in the home or in the study participant’s natural environment.

2 **How are the data collected?** (i.e., does collection require an intermediary like a study team or phlebotomist?)

In the past, most data were collected *via an intermediary* – someone from the team would record information in a custom software system and/or case report form. As digital tools advance, we can collect more endpoint-supporting data at home via digital surveys and sensors, and study teams can “visit” patients at home via telemedicine conference calls. This means that more of the data is **participant-generated** and collected “virtually,” without an intermediary.

Putting these words into context, a doctor who remote teleconferences into a patient would be conducting a “remote trial,” but they might collect the data manually through a survey, so the study would not be considered a “virtual trial.” In contrast, a study team might collect all the data passively from a smartwatch in a clinic, and this study would be “virtual trial” from a data collection perspective, but not a “remote trial” because the patient is in a centralized location.

Notably, the industry has not yet settled on language around these types of trials, which is not unexpected. As a new field emerges, so does a new vocabulary. Historically, some researchers, primarily behavioral scientists, have referred to this style as a **community-based clinical trial (CBCT)** as a clinic may not be involved in the intervention being tested [97]. The FDA has been using the term DCT more often in the past few years to describe trials taking place at home or in the community [98].

What Is Real-World Data?

Real-world is a word we have so far avoided in this primer and is important to define as it is often misused. The regulatory definition of real-world data is the data collected outside of a traditional clinical study, such as a **randomized controlled trial (RCT)** [99]. These data sources include EHRs, claims and billing activities, product and disease registries, patient-generated data including in home-use settings, and data gathered from other sources that can inform on health status, such as mobile technologies [99].

Real-world **evidence** is the evidence derived from real-world **data** [100]. In the context of a traditional RCT, if study participants contribute to some measurements at home, such as pain measurement via ePRO or step count from a wearable sensor, many often mistakenly believe this would constitute “real world data.” However, these measures would not constitute real-world data because the participants have been preselected for study entry by the inclusion and exclusion criteria of a given trial. They therefore do not represent the overall population in a certain indication. Therefore, when working with clinical research, it is important to focus on the benefits of health-related data collected in natural settings– which may not be classified as “real world” by a strict regulatory definition.

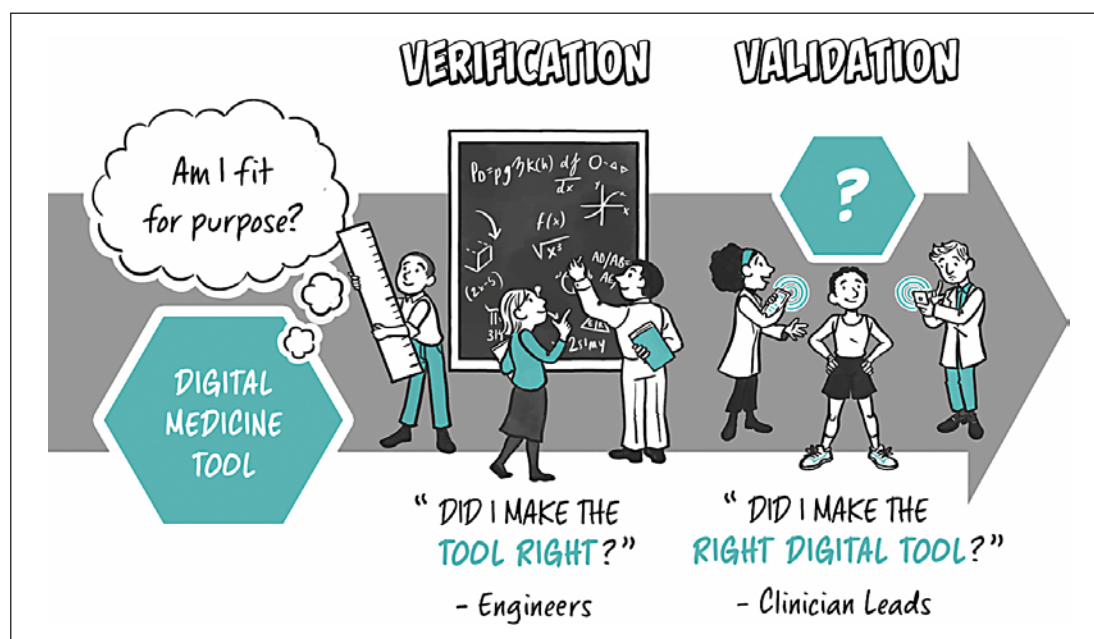


Fig. 12. Verification and validation. Verification and validation are two processes that, together, indicate whether a digital measurement tool is fit-for-purpose.

Verification and Validation: How Do I Assure Myself and Others that My Digital Measurement Tool Generates Good, Trustworthy Data? (Fig. 12)

Once you have arrived at a construct to measure and you have some ideas about the way a digital tool could be used to measure it, there are a number of ways to assure yourself and others that the tool generates high quality, meaningful, and trustworthy data. **Verification and validation (V&V)** are terms that are probably familiar to you if you come from an engineering or product development background. They apply to digital medicine, too.

- **Verification** is the assessment of sensor accuracy (which describes the agreement between the measurement made by a single sensor vs. a ground truth), precision (which describes the agreement between multiple measurements made by a single sensor back-to-back), consistency (which describes the agreement between multiple measurements made by a single sensor over longer time periods), and/or uniformity which describes the agreement across measurements made by multiple sensors simultaneously). A sensor that is accurate, precise, consistent, and uniform will “give the right answer every time.” By undertaking verification assessments, the investigator can also be assured that the relevant firmware/software that generates processed data is also accurate, precise, consistent, and uniform [101]. Verification answers the question: “**did I make the tool right?**” Verification is an *engineering assessment*, and is entirely separate from data collection on human subjects.
- **Validation** is the process of ensuring that the digital measurement tool is meeting its intended use by generating objective data that accurately represents the concept of interest – the specific way in which the patient feels, functions, or survives – that it purports to be measuring. Validation answers the question, “**Did I build the right tool?**” The concept of validation can be broken down as follows:
 - **Analytic validation:** Is the algorithm processing the data to report the measurement of interest? For example, is the algorithm accurately processing raw accelerometry data to calculate gait speed in a particular patient population?

- **Clinical validation:** Is the measurement of interest reflecting the concept of interest – the specific way in which the patient feels, functions, or survives? For example, is gait speed a meaningful measure that reflects the ability of how a particular patient populations feels, functions, or survives?
- To answer whether the technology is measuring what its intended to measure (clinical validation) and is correct (analytic validation), developers should work with researchers to ensure that validation studies are well-designed.

The most important concept to remember when considering verification and validation is the separation of the *sensor* from the *endpoint*. All sensors can be boiled down to the physical construct that is measured, such as acceleration, temperature, or pressure. Verification testing can be completed at the bench, and is the process of ensuring that the sensor works appropriately. Validation testing almost always involves human subject testing, and is the process of ensuring that the output data from the technology is accurate against a gold-standard (analytic validation) and an appropriate reflection of the clinical concept of interest (clinical validation). Answering the latter questions will often involve testing the technology with human participants, which may require an ethics committee review prior to testing (see Part II). Typically, verification studies happen at the bench without human subjects and so do not require ethics committee review.

Clinical validation has multiple dimensions. We will not go into all of them here, but one example is whether a change in your new measurement is regarded as **meaningful** by people with the disease. Do the results generated by the tool capture all aspects of the concept you are measuring? Do changes to measurement values predict certain clinical outcomes down the road? Does the measurement respond to an intervention that is well understood to have an effect on the property you are measuring in that population? Can the measure correctly identify those patients with and without the condition (sensitivity and **specificity**)? Likewise, does that property remain unchanged in circumstances when it should not change? (e.g., in a different population, or when there is no intervention). Of particular note in the realm of digital medicine is the reliance on computational algorithms, whose performance can improve over time given access to more representative datasets (see Part I). The regulatory framework to deal with such systems is under active development at the FDA [102].

Often, researchers will ask how a novel digital measure compares to a gold standard assessment. We believe “gold standard” is often a misnomer because many gold standards are not necessarily high-quality measures. If the most widely used existing measurement is suboptimal, more appropriate terms include **legacy standard**, where a new and better measurement has been developed, or **current standard**, in cases where the standard is acknowledged to be inadequate but no alternative yet exists.

For example, the legacy assessment in Duchenne’s muscular dystrophy (DMD) is the six-minute walk test. This is a poor endpoint for a number of reasons. First, it does not apply to the roughly 60% of that patient population who are confined to a wheelchair and cannot participate in a walk test. Second, DMD trial participants typically range in age from 7 years (though this has been noted to be lowering in recent years) to mid-teens, so bias may be introduced in a number of directions; in some cases due to issues of diminished patient volition and in others due to coaching or gaming by parents of young participants. These biases may lead to noise that obscures and signal in the endpoint.

Finally, many consider that the six-minute walk test fails to demonstrate strong **ecological validity**, i.e. that it is a poor measure of how test performance predicts behaviors in real-world settings. Stellman et al. [103] offer an excellent exploration of the ecological validity of mobility outcomes, including the six-minute walk test, in MS and consider the opportunities for digital tools to improve these measures.

Regardless of the quality of a legacy standard, it is unlikely that a digital measurement will agree perfectly with existing standard. In fact, for this reason, digital measures garner a lot of excitement: they may turn out to be more sensitive than traditional measures or be capable of measuring something researchers have never been able to measure before. Traditional measurements not only provide only a tiny snapshot of information about a patient's experience of their disease, but are also fraught with confounders such as **white coat syndrome**, where a patient's feeling of anxiety in a medical environment results in an abnormally high reading when assessing blood pressure [15].

In other instances, a digital assessment may measure an aspect of disease that has been inaccessible with traditional measures. For the DMD example, a more inclusive digital measurement could be of upper limb mobility, which can be applied to a much broader population of DMD patients. However, there would be little utility in trying to tether this new digital measure against the legacy standard of a six-minute walk test.

To summarize simply, if the test can be performed by a good engineer or physicist who flunked biology, then it is **verification**. If it requires medical knowledge, it is **validation**.

When Is Something “Validated” Enough?

This depends on the specific application. Tools need to be **fit-for-purpose**. This means that the level of validation associated with a medical product development tool is sufficient to support its **context of use**, a regulatory term that refers to a description of how the tool is used and where it is applied [104].

The CTTI, a public private partnership co-founded by Duke University and the FDA, has developed comprehensive recommendations and resources on developing digital measurement tools for use as clinical trial endpoints [105]. This may be a valuable resource to anyone looking to understand the body of evidence that is required to support the use of a digital medicine tool in a clinical trial.

The Increasing Overlap between Clinical Research and Care

Historically, measures that support research (endpoints) and care (outcomes) were siloed. But that is changing: many clinically validated endpoints used in research will likely transition into clinical care. A number of companies are working towards a universal vision of human digital measurement across the continuum of research and clinical care. Clinical research provides a practical use-case to link (or validate) everyday behaviors and outcomes. Companies that develop digital biomarkers validate these tools through clinical research as a first step towards what may eventually become a validated digital diagnostic or a digital therapeutic.

Similarly, a number of big tech companies are also developing digital measures for clinical settings. In 2018, the FDA cleared a “software as a medical device” (SaMD) for the Apple Watch, which can determine the presence of atrial fibrillation, an abnormal heart condition. This clearance provides a regulatory pathway for companies to create more advanced diagnostics and interventions for the patient, at home, decentralized.

There is crossover in use cases between research and routine care, such as the safety monitoring of trial participants during the course of a clinical study. However, even in such crossover cases, the clinical trial environment is by its nature more structured and controlled than the variety of settings and scenarios where clinical medicine is practiced.

Conclusion

Our society is hitting an inflection point in digital medicine – more high-quality data than ever can be collected and remotely that can influence high-stakes measurements and intervention in near real-time – and now is the time to shape the healthcare community. There are a number of ontologies, frameworks, and decisions that we need to make to ensure that the advent of digital medicine creates a better healthcare system. Clarifying language and establishing a standard lexicon will advance the field faster, together, and with more trust. This primer is a first stake in the ground, and we expect and hope a common framework and understanding will evolve over time.

Our communities face challenging decisions, particularly ethical ones around surveillance, convenience, personalization, and privacy – and it is important to remember that all systems are first built by humans who design the incentives. Let us build an intentional future that we want to live in, and not an accidental one. We need to ensure that we adopt healthcare technologies that are worthy of the trust we place in them [106].

Acknowledgements

For refinement of ideas, frameworks, and terms relating to digital medicine, we gratefully acknowledge Adam Conner-Simons, Beau Woods, Jessie Bakker, and Emily Singer. For editing and formatting work, we acknowledge Rebecca B. Malcolm and Christine J. Manta. We also acknowledge the team at COGNI+IVE (wearecognitive.com) for bringing our ideas to life as illustrations. Andrea Coravos acknowledges support from Harvard-MIT Center for Regulatory Science.

Statement of Ethics

The authors have no ethical conflicts to disclose.

Disclosure Statement

Andrea Coravos is a founder and owns a significant interest in Elektra Labs Inc. *Jennifer C. Goldsack* is employed by monARC Bionetworks Inc. *Daniel Karlin* is a paid employee and/or owns a significant interest in HealthMode, CEAI, NightWare, Humanity.co, and 4YouandMe. *Camille Nebeker* is employed by UC San Diego School of Medicine and directs the ReCODE center. *Eric Perakslis* is affiliated with Harvard Medical School, Duke University, Elektra Labs Inc., and NuMedii Inc. *Noah Zimmerman* holds equity in Picnic Health. *Kelley Erb* is a paid employee of Pfizer Inc.

Funding Sources

The illustrations prepared by Cogni+ive were funded by Pfizer Inc.

Author Contributions

A.C., J.C.G., D.R.K., C.N., E.P., N.Z., and M.K.E. all made substantial contributions to the conception or design of the work. *A.C., J.C.G., D.R.K., C.N., N.Z., and M.K.E.* all drafted this work and revised it critically for important intellectual content; *E.P.* revised this work critically for important intellectual content. *A.C., J.C.G., D.R.K., C.N., E.P., N.Z., and M.K.E.* all gave final approval of the version to be published. *A.C., J.C.G., D.R.K., C.N., E.P., N.Z., and M.K.E.* all agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

- 1 Canavan N. Cartoons offer a peek into cancer immunotherapy — and scientists' minds. *STAT* [Internet]. 2018 Oct 17 [cited 2019 April 13]. Available from: <https://www.statnews.com/2018/10/17/cancer-immunotherapy-scientists-cartoons/>
- 2 Delp C, Jones J. Communicating information to patients: the use of cartoon illustrations to improve comprehension of instructions. *Acad Emerg Med*. 1996 Mar;3(3):264–70.
- 3 Steinhubl SR, Topol EJ. Digital medicine, on its way to being just plain medicine. *Npj Digit Med*. 2018;1:5.
- 4 The Digital Medicine Society. The Digital Medicine (DiMe) Society [Internet]. 2019 [cited 2019 Apr 13]. Available from: www.dimesociety.org
- 5 Miotto R, Danieletto M, Scelza JR, Kidd BA, Dudley JT. Reflecting health: smart mirrors for personalized medicine. *Npj Digit Med*. 2018 Nov 8;1(1):62.
- 6 Digital Therapeutics Alliance. Digital Therapeutics: Combining Technology and Evidence-based Medicine To Transform Personalized Patient Care [Internet] [updated 31 Oct 2018; cited 2019 April 13]. Available from: https://www.dtxalliance.org/wp-content/uploads/2018/09/DTA-Report_DTx-Industry-Foundations.pdf
- 7 Linebaugh K. Citizen Hackers Tinker with Medical Devices. *Wall Street Journal* [Internet]. 2014 Sept 26 [cited 2019 April 13]. Available from: <https://www.wsj.com/articles/citizen-hackers-concoct-upgrades-for-medical-devices-1411762843?tesla=y>
- 8 Coravos A. The Doctor Prescribes Video Games and Virtual Reality Rehab [Internet]. *Wired*. 2018 Nov 20 [cited 2019 April 13]. Available from: <https://www.wired.com/story/prescription-video-games-and-vr-rehab/>
- 9 Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med*. 2019 Jan;25(1):24–9.
- 10 Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. *BMC Sports Sci Med Rehabil*. 2015 Oct;7(1):24.
- 11 Dunn J, Runge R, Snyder M. Wearables and the medical revolution. *Per Med*. 2018 Sep;15(5):429–48.
- 12 U.S. Food and Drug Administration. FDA approves pill with sensor that digitally tracks if patients have ingested their medication [Internet]. 2017 [updated 2018 Feb 22; cited 2019 Apr 13]. Available from: <https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm584933.htm>
- 13 Shameer K, Badgeley MA, Miotto R, Glicksberg BS, Morgan JW, Dudley JT. Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams. *Brief Bioinform*. 2017 Jan;18(1):105–24.
- 14 Clinical Trials Transformation Initiative. CTTI Recommendations: Developing Novel Endpoints Generated by Mobile Technologies For Use in Clinical Trials [Internet]. 2017 [updated 2017 June; cited 2019 April 13]. Available from: <https://www.ctti-clinicaltrials.org/files/novelendpoints-recs.pdf>
- 15 Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. *Clin Pharmacol Ther*. 2018 Jul;104(1):42–52.
- 16 Fillit HM. FDA removes an unnecessary barrier to testing Alzheimer's drugs. *STAT* [Internet]. 2018 Mar 5 [cited 2019 Apr 13]. Available from <https://www.statnews.com/2018/03/05/alzheimers-disease-drugs-fda/>
- 17 Kosov M, Kolesnik V, Riefler J, Belotserkovskiy M. Implication of Murphy's Law in Clinical Trials. *J Neoplasms*. 2017;2(3):24.
- 18 Ten Haken I, Ben Allouch S, van Harten WH. The use of advanced medical technologies at home: a systematic review of the literature. *BMC Public Health*. 2018 Feb;18(1):284.
- 19 Basic HHS Policy for Protection of Human Research. Definitions, 45 CFR 46.102 (2009) [regulation on Intranet] [cited 2019 Apr 13]. Available from: <https://www.hhs.gov/ohrp/sites/default/files/ohrp/policy/ohrpregulations.pdf>
- 20 Clinical Trials Transformation Initiative. Recommendations Executive Summary: Advancing the Use of Mobile Technologies for Data Capture & Improved Clinical Trials [Internet] [cited 2019 Apr 13]. Available from: <https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/mobile-technologies-executive-summary.pdf>
- 21 World Health Organization. Medicines Regulatory Support [Internet] [cited 2019 Apr 13]. Available from: https://www.who.int/medicines/areas/quality_safety/regulation_legislation/en/

- 22 World Health Organization. Global Atlas at Medical Devices. Geneva: World Health Organization; 2017. 480 p. Available from: https://www.who.int/medical_devices/publications/global_atlas_meddev2017/en/
- 23 U.S. Food and Drug Administration. History of FDA's Centers and Offices [Internet] [updated 2018 Feb 1; cited 2019 Apr 13]. Available from: <https://www.fda.gov/aboutfda/history/forgshistory/historyoffdascenter-sandoffices/default.htm>
- 24 Fromer MJ. FDA introduces new phase 0 for clinical trials: some enthusiastic, some skeptical. *Oncol Times*. 2006 Aug;28(15):18–9.
- 25 U.S. Food and Drug Administration. Drug Development Tools (DDTs) [Internet] [updated 2018 Oct 26; cited 2019 Apr 13]. Available from: <https://www.fda.gov/drugs/developmentapprovalprocess/drugdevelopmenttoolsqualificationprogram/>
- 26 U.S. Food and Drug Administration. Medical Device Development Tools (MDDT) [Internet] [updated 2019 Mar 15; cited 2019 April 13]. Available from: <https://www.fda.gov/medicaldevices/scienceandresearch/medicaldevicedevelopmenttoolsmddt>
- 27 Prescription Drug Use Related Software; Establishment of Public Docket. Request for Comments. *Fed Regist*. 2018 Nov;83(244):58574–82.
- 28 U.S. Food and Drug Administration. About Combination Products [Internet] [updated 2019 Feb 26; cited 2019 Apr 13]. Available from: <https://www.fda.gov/CombinationProducts/>
- 29 U.S. Food and Drug Administration. Information Sheet Guidance For IRBs, Clinical Investigators, and Sponsors. Silver Spring, MD: U.S. Food and Drug Administration. January 2006 [cited 2019 Apr 13]. 15 p. Available from: <https://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126418.pdf>
- 30 U.S. Food and Drug Administration. Software as a Medical Device (SaMD) [Internet] [updated 2019 Apr 2; cited 2019 Apr 13]. Available from: <https://www.fda.gov/MedicalDevices/DigitalHealth/SoftwareasMedicalDevice/ucm20086412.htm>
- 31 U.S. Food and Drug Administration. Is The Product a Medical Device? [Internet] [updated 2018 Mar 22; cited 2013 Apr 13]. Available from: <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm051512.htm>
- 32 U.S. Food and Drug Administration. Changes to Existing Medical Software Policies Resulting from Section 3060 of the 21st Century Cures Act. Silver Spring: U.S. Food and Drug Administration. 2017 Dec 3 [cited 2019 Apr 13]. Available from: <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM587820.pdf>
- 33 U.S. Food and Drug Administration. Contact Us – Division of Industry and Consumer Education (DICE) [Internet] [updated 2019 Jan 28; cited 2019 Apr 13]. Available from: <https://www.fda.gov/medicaldevices/deviceregulationandguidance/contactdivisionofindustryandconsumereducation/default.htm>
- 34 U.S. Food and Drug Administration. Multiple Function Device Products. Policy and Considerations. Silver Spring: U.S. Food and Drug Administration. 2018 Apr 27 [cited 2019 Apr 13]. 18 p. Available from: <https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm605683.pdf>
- 35 U.S. Food and Drug Administration. Reclassification [Internet] [updated 2018 Dec 13; cited 2019 Apr 13]. Available from: <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm080412.htm>
- 36 What's the Difference Between the FDA Medical Device Classes? [Internet]. 2018 Feb 2 [cited 2019 Apr 13] Available from: <https://www.bmpmedical.com/blog/whats-difference-fda-medical-device-classes-2/>
- 37 U.S. Food and Drug Administration. Device Classification under Section 513(f)(2)(de novo) [Internet] [updated 2019 Apr 15; cited 2019 Apr 13] Available from: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/denovo.cfm?ID=DEN180044>
- 38 Burton TM. FDA Is Revamping Clearance Procedures for Medical Devices. *Wall Street Journal* [Internet]. 26 Nov 2018 [cited 2019 Apr 13]. Available from: <https://www.wsj.com/articles/fda-is-revamping-clearance-procedures-for-medical-devices-1543234015>
- 39 U.S. Food and Drug Administration. Guidance with Digital Health Content [Internet] [updated 2018 Aug 30; cited 2019 Apr 13]. Available from: <https://www.fda.gov/MedicalDevices/DigitalHealth/ucm562577.htm>
- 40 U.S. Food and Drug Administration. General Wellness: Policy for Low Risk Devices. Silver Spring: U.S. Food and Drug Administration. 2016 July 20 [cited 2019 Apr 13]. 10 p. Available from: <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm429674.pdf>
- 41 U.S. Food and Drug Administration. Mobile Medical Applications. Silver Spring: U.S. Food and Drug Administration. 2015 Feb 9 [cited 2019 Apr 13]. 44 p. Available from: <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf>
- 42 U.S. Food and Drug Administration. Software as a Medical Device (SAMd). Silver Spring: U.S. Food and Drug Administration. 2017 Dec 8 [cited 2019 Apr 13]. 30 p. Available from: <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM524904.pdf>
- 43 U.S. Food and Drug Administration. Clinical and Patient Decision Software. Silver Spring: U.S. Food and Drug Administration. 2017 Dec 8 [cited 2019 Apr 13] 13 p. Available from: <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm587819.pdf>
- 44 U.S. Food and Drug Administration. Digital Health Software Precertification (Pre-Cert) Program [Internet] [updated 2019 Feb 15; cited 2019 Apr 13]. Available from: <https://www.fda.gov/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/default.htm>

- 45 International Medical Device Regulators Forum. Software as a Medical Device (SaMD) [cited 2019 Apr 13]. Available from: <http://www.imdrf.org/workitems/wi-samd.asp>
- 46 U.S. Food and Drug Administration. 21st Century Cures Act [Internet] [updated 2018 Mar 29; cited 2019 Apr 13]. Available from: <https://www.fda.gov/regulatoryinformation/lawsenforcedbyfda/significantamendmentstothefdcact/21stcenturycuresact/default.htm>
- 47 U.S. Food and Drug Administration. Examples of Mobile Apps for Which the FDA Will Exercise Enforcement Discretion [Internet] [updated 2016 Aug 1; cited 2019 Apr 13]. Available from: <https://www.fda.gov/medicaldevices/digitalhealth/mobilemedicalapplications/ucm368744.htm>
- 48 Federal Trade Commission. Mobile Health Apps Interactive Tool. In: Federal Trade Commission [Internet]. 2016 Apr 2016 [cited 2019 Apr 13]. Available from: <https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool>
- 49 Van Norman GA. Drugs and devices: comparison of European and U.S. approval processes. *JACC Basic Transl Sci*. 2016 Aug;1(5):399–412.
- 50 National Institute for Health and Care Excellence. Evidence Standards Framework for digital Health Technologies. 2019 March [cited 2019 Apr 13]. Available from: <https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf>
- 51 National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Department of Health, Education and Welfare (DHEW). The Belmont Report [Internet]. Washington: United States Government Printing Office. 1979 September 30 [cited 2019 Apr 13]. Available from: https://videocast.nih.gov/pdf/ohrp_belmont_report.pdf
- 52 Mao W, Raju A, Nebeker CA. Privacy Attitudes and Generational Differences. Poster session presented at *AI and Health Living Network Conference*, Cambridge MA.
- 53 Nebeker C, Murray K, Holub C, Haughton J, Arredondo EM. Acceptance of mobile health in communities under-represented in biomedical research: barriers and ethical considerations for scientists. *JMIR Mhealth Uhealth*. 2017 Jun;5(6):e87.
- 54 Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, et al. Reimagining human research protections for 21st century science. *J Med Internet Res*. 2016 Dec;18(12):e329.
- 55 Basic HHS Policy for Protection of Human Research. Expedited review procedures for certain kinds of research involving no more than minimal risk and for minor changes in approved research, 45 CFR 46.110 (2009) [regulation on Intranet] [cited 2019 Apr 13]. Available from: <https://www.hhs.gov/ohrp/sites/default/files/ohrp/policy/ohrpregulations.pdf>
- 56 Nebeker C, Bartlett Ellis R, Torous J. Digital Health Decision Making Checklist: Designed for Researchers [Internet]. 2018 [cited 2019 Apr 13]. Available from: <http://thecore.ucsd.edu/dmchecklist/>
- 57 Sage Bionetworks. Elements of Informed Consent [Internet]. 2018 Nov 5 [cited 2019 Apr 13]. Available from: <http://sagebionetworks.org/in-the-news/elements-informed-consent/>
- 58 Clinical Trials Transformation Initiative. Project: Mobile Technologies [Internet] [cited 2019 Apr 13] Available from: <https://www.ctti-clinicaltrials.org/projects/mobile-technologies>
- 59 Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov*. 2010 Mar;9(3):203–14.
- 60 Clinical Trials Transformation Initiative. Developing Novel Endpoints Generated by Mobile Technology for Use in Clinical Trials [Internet]. 15 May 2017 [cited 2019 Apr 13]. Available from: <https://www.ctti-clinicaltrials.org/briefing-room/recommendations/developing-novel-endpoints-generated-mobile-technology-use-clinical>
- 61 The Editorial Board. How Silicon Valley Puts the ‘Con’ in Consent. *New York Times* [Internet]. 2019 Feb 2 [cited 2019 Apr 13]. Available from: <https://www.nytimes.com/2019/02/02/opinion/internet-facebook-google-consent.html>
- 62 State of California Department of Justice. California Consumer Privacy Act (CCPA) [Internet]. 2019 March [cited 2019 Apr 13]. Available from: <https://oag.ca.gov/privacy/ccpa>
- 63 Allen M. You Snooze, You Lose: Insurers Make the Old Adage Literally True. *Pro Publica* [Internet]. 2018 Nov 21 [cited 2019 Apr 13]. Available from: <https://www.propublica.org/article/you-snooze-you-lose-insurers-make-the-old-adage-literally-true>
- 64 Mehta I. Amazon's new patent will allow Alexa to detect a cough or a cold. *Artificial Intelligence* [Internet]. 2018 Oct 15 [cited 2019 Apr 13]. Available from: <https://thenextweb.com/artificial-intelligence/2018/10/15/amazons-new-patent-will-allow-alexa-to-detect-your-illness/>
- 65 Perks E. Protecting patient privacy and security while exploiting utility of next generation digital health wearables [Internet]. *BMJ Opinion*. 2019 Jan 18 [cited 2019 Apr 13]. Available from: <https://blogs.bmj.com/bmj/2019/01/18/protecting-patient-privacy-and-security-while-exploiting-the-utility-of-next-generation-digital-health-wearables/>
- 66 Berinato S. There is No Such Thing as Anonymous Data. *Harvard Business Review* [Internet]. 2015 Feb 9 [cited 2019 Apr 13]. Available from: <https://hbr.org/2015/02/theres-no-such-thing-as-anonymous-data>
- 67 Allen M. Health Insurers Are Vacuuming Up Details About You—And it Could Raise Your Rates. *ProPublica* [Internet]. 2018 July 17 [cited 2019 Apr 13]. Available from: <https://www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates>

- 68 Ravindranath M. How Your Health Information Is Sold and Turned into Risk Scores. *Politico* [Internet]. 2019 Feb 3 [cited 2019 Apr 13]. Available from: https://www.politico.com/story/2019/02/03/health-risk-scores-opioid-abuse-1139978?mc_cid=e83d0f0d3e&mc_eid=4d57c573ca
- 69 Khullar DA. Could Worsen Health Disparities. *New York Times* [Internet]. 2019 Jan 31 [cited 2019 Apr 13]. Available from: <https://www.nytimes.com/2019/01/31/opinion/ai-bias-healthcare.html>
- 70 Coravos A, Chen I, Gordhandas A, Stern, AD. We Should Treat Algorithms Like Prescription Drugs. 2019 Feb 14 [cited 2109 Apr 13]. Available from: <https://qz.com/1540594/treating-algorithms-like-prescription-drugs-could-reduce-ai-bias/>
- 71 Matwyshyn AM. Internet of Bodies. 61 *William & Mary. Law Rev.* 2019 Oct.
- 72 Rockoff JD. J&J Warned Insulin Pump Vulnerable to Cyber Hacking. *Wall Street Journal* [Internet]. 2016 Oct 4 [cited 2019 Apr 13]. Available from: <https://www.wsj.com/articles/j-j-warns-insulin-pump-vulnerable-to-cyber-hacking-1475610989?mg=prod/accounts-wsj>
- 73 Larson S. FDA Confirms that St. Jude's Cardiac devices can be hacked [Internet]. *CNN Business*. 2017 Jan 9 [cited 2019 Apr 13]. Available from: <https://money.cnn.com/2017/01/09/technology/fda-st-jude-cardiac-hack/index.html>
- 74 Brewster T. Medical Devices Hit by Ransomware for the First Time in US Hospitals. *Forbes* [Internet]. 2017 May 17 [cited 2019 Apr 13]. Available from: <https://www.forbes.com/sites/thomasbrewster/2017/05/17/wannacry-ransomware-hit-real-medical-devices/#dad2b18425cf>
- 75 I am the Cavalry. I am the Cavalry Position on Disclosure [Internet]. 2018 Aug 24 [cited 2019 Apr 13]. Available from: <https://www.iamthecavalry.org/about/disclosure/>
- 76 Common Vulnerabilities and Exposures. Common Vulnerabilities and Exposures [Internet] [updated 2019 Apr 18; cited 2019 Apr 13]. Available from: <https://cve.mitre.org>
- 77 Perakslis ED, Stanley M. A cybersecurity primer for translational research. *Sci Transl Med*. 2016 Jan;8(322):322ps2.
- 78 Philips. Committed to proactively addressing our customers security and privacy concerns [Internet]. 2018 Jan [cited 2019 Apr 13]. Available from: https://www.philips.com/cdam/corporate/newscenter/global/standard/resources/healthcare/2018/cybersecurity/Cybersecurity_position_paper.pdf
- 79 Ask J, Jacobson J. NTIA Software Component Transparency [Internet]. 2018 Nov 6 [cited 2019 Apr 13]. Available from: https://www.ntia.doc.gov/files/ntia/publications/healthcare_wg_-_proof_of_concept_overview.pdf
- 80 Mayo Clinic. Medical and Research Device Risk Assessment Vendor Packet Instructions [Internet]. 2019 Jan 1 [cited 2019 Apr 13]. Available from: <http://www.mayoclinic.org/documents/medical-device-vendor-instructions/doc-20389647>
- 81 Woods B, Coravos A, Corman JD. The Case for a Hippocratic Oath for Connected Medical Devices: viewpoint. *J Med Internet Res*. 2019 Mar;21(3):e12568.
- 82 I Am the Cavalry. Hippocratic Oath for Connected Medical Devices [cited 2019 Apr 13]. Available from: <https://www.iamthecavalry.org/domains/medical/oath/>
- 83 U.S. Food and Drug Administration. Content of Premarket Submissions Management in Cybersecurity in Medical Devices. Silver Spring: U.S. Food and Drug Administration. October 2018 [cited 2019 Apr 13]. 24 p. Available from <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM623529.pdf>
- 84 U.S. Food and Drug Administration. Post market Management of Cybersecurity in Medical Devices. Silver Spring: U.S. Food and Drug Administration. January 2016 [cited 2019 Apr 13]. 30 p. Available from: <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm482022.pdf>
- 85 #WeHeartHackers. #WeHeartHackers [Internet] [cited Apr 13]. Available from: <http://wehearthackers.org>
- 86 Health Care Council. HealthCare and Public Health Sector Critical Infrastructure Security and Resilience Partnership [cited 2019 Apr 13]. Available from: <https://healthsectorcouncil.org>
- 87 FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource Glossary [Internet]. Silver Spring: U.S. Food and Drug Administration; 2016 Jan 28 [updated 2018 May 2; cited 2019 Apr 13]. Available from <https://www.ncbi.nlm.nih.gov/books/NBK338448/>
- 88 Strimbu K, Tavel JA. What are biomarkers? *Curr Opin HIV AIDS*. 2010 Nov;5(6):463–6.
- 89 Coravos A, Khozin S, Mandl KD. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *Npj Digit Med*. 2019 Mar 11;2(1):14. <https://doi.org/10.1038/s41746-019-0090-4>.
- 90 Califf RM. Biomarker definitions and their applications. *Exp Biol Med (Maywood)*. 2018 Feb;243(3):213–21.
- 91 FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Silver Spring: U.S. Food and Drug Administration. 2016 [cited 2019 Apr 13]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/27010052>
- 92 U.S. Food and Drug Administration. FDA Public Workshop: 2018 Clinical Outcome Assessments in Cancer Clinical Trials [updated 2018 Oct 19; cited 2019 Apr 13]. Available from: <https://www.fda.gov/newsevents/meetingsconferencesworkshops/ucm602540.htm>
- 93 Insel TR. Digital phenotyping: technology for a new science of behavior. *JAMA*. 2017 Oct;318(13):1215–6.
- 94 Khozin S, Coravos A. Decentralized Trials in the Age of Real-World Evidence and Inclusivity in Clinical Investigations. *Clin Pharmacol Ther*. <https://doi.org/10.1002/cpt.1441>.

- 95 Clinical Trials Transformation Initiative. Decentralized Clinical Trials [Internet]. 27 May 2016 [cited 2019 Apr 13]. Available from: <https://www.ctti-clinicaltrials.org/projects/decentralized-clinical-trials>
- 96 Coravos A. Decentralized clinical trials. Medium [Internet]. 2018 Nov 12 [cited 2019 Apr 13]. Available from: <https://medium.com/elektra-labs/decentralized-clinical-trials-f12fffb72610>
- 97 King AC, Haskell WL, Taylor CB, Kraemer HC, DeBusk RF. Group- vs home-based exercise training in healthy older men and women. A community-based clinical trial. *JAMA*. 1991 Sep;266(11):1535–42.
- 98 U.S. Food and Drug Administration. Breaking Down Barriers Between Clinical Trials and Clinical Care: Incorporating Real World Evidence into Regulatory Decision Making [Internet] [updated 2019 Jan 28; cited 2019 Apr 13]. Available from: <https://www.fda.gov/NewsEvents/Speeches/ucm629942.htm>
- 99 U.S. Food and Drug Administration. Real World Evidence [Internet] [updated 2019 Feb 6; cited 2019 Apr 13]. Available from: <https://www.fda.gov/scienceresearch/specialtopics/realworldevidence/default.htm>
- 100 Makady A, de Boer A, Hillege H, Klungel O, Goettsch W; on behalf of GetReal Work Package 1. What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health*. 2017 Jul - Aug;20(7):858–65.
- 101 Clinical Trials Transformative Initiative. Figure 2. Data Processes and Information to Provide to FDA [Internet] [cited 2019 Apr 13]. Available from: <https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/figure-2-data-process-fda-submission.pdf>
- 102 U.S. Food and Drug Administration. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as Medical Device (SaMD) [Internet] [cited 2019 Apr 13]. Available from: <https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/SoftwareasaMedicalDevice/UCM635052.pdf>
- 103 Stellmann JP, Neuhaus A, Götze N, Briken S, Lederer C, Schimpl M, et al. Ecological validity of walking capacity tests in multiple sclerosis. *PLoS One*. 2015 Apr;10(4):e0123822.
- 104 Banegas JR, Ruilope LM, de la Sierra A, Vinyoles E, Gorostidi M, de la Cruz JJ, et al. Relationship between clinic and ambulatory blood-pressure measurements and mortality. *N Engl J Med*. 2018 Apr;378(16):1509–20.
- 105 Meyer R. Everything We Know About Facebook’s Secret Mood Manipulation Experiment. The Atlantic [Internet]. 2014 Jun 28 [2019 Apr 13]. Available from: <https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/>
- 106 I Am The Cavalry. I Am The Cavalry Cyber Safety Outreach [updated 2019 Mar 18; cited 2019 Apr 19]. Available from: <https://www.iamthecavalry.org>
- 107 Kramer F, Dinh W. Molecular and digital biomarker supported decision making in clinical studies in cardiovascular indications. *Arch Pharm*. 2016 Jun;349(6):399-409.
- 108 Miyoshi A, Nishii N, Kubo M, Okamoto Y, Fujii S, Watanabe A, et al. An improved algorithm calculated from intrathoracic impedance can precisely diagnose preclinical heart failure events: Sub-analysis of a multicenter MOMOTARO (Monitoring and Management of OptiVol Alert to Reduce Heart Failure Hospitalization) trial study. *J Cardiol*. 2017 Nov 1;70(5):425-31.
- 109 Powers III JH, Patrick DL, Walton MK, Marquis P, Cano S, Hobart J, et al. Clinician-reported outcome assessments of treatment benefit: report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. *Value Health*. 2017 Jan 1;20(1):2-14.