

EXECUTIVE SUMMARY FOR THE PATIENT ENGAGEMENT ADVISORY COMMITTEE MEETING

Connected and Empowered Patients: e Platforms Potentially Expanding the
Definition of Scientific Evidence

November 15, 2018

Table of Contents

INTRODUCTION.....	2
Regulation of Medical Devices.....	4
Overview of Patient-Generated Health Data	12
SOCIAL MEDIA LISTENING.....	14
SENSOR-BASED DIGITAL HEALTH TECHNOLOGIES.....	22
PATIENT-DRIVEN REGISTRIES.....	25
Challenges with Patient-Generated Health Data.....	31
Conclusion.....	37

Disclaimer: This Executive Summary is for discussion purposes only and does not represent draft or final guidance. It is not intended to propose or implement policy changes regarding postmarket actions or the use of real world evidence including patient-generated health data for regulatory activities or the applicable statutory and regulatory requirements for supporting a regulatory decision.

INTRODUCTION

FDA's commitment to patient safety. Providing patients with access to high quality, safe and effective medical devices that meet their health care needs is a top priority at the Food and Drug Administration (FDA). As part of FDA's public health responsibilities, FDA strives to permit marketing of devices that have a favorable benefit-risk profile that spans the total product life cycle. FDA has published several guidance documents that demonstrate a flexible, patient-centric, benefit-risk approach to regulatory decisions, including consideration of patient preferences and uncertainty.^{1,2,3,4} However, new information about a device's safety, such as reports of unexpected adverse events, may only become available when the device reaches the market and is used under real-world conditions in a broader patient population. In April 2018, CDRH published the Medical Device Safety Action Plan which details how FDA will encourage innovation to improve safety, detect safety risks earlier, and keep doctors and patients better informed.⁵ The plan builds on steps FDA has taken over the past several years to enhance patient safety. A key tenet of this plan includes

¹FDA. *Factors to Consider When Making Benefit-Risk Determinations for Medical Device Investigational Device Exemptions*. Available at <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM451440> (accessed October 17, 2018).

² FDA. *Patient Preference Information – Voluntary Submission, Review in Premarket Approval Applications, Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in Decision Summaries and Device Labeling*. Available at <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM446680.pdf> (accessed October 17, 2018).

³ FDA. *Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications*. Available at <https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm517504.pdf> (accessed October 25, 2018)

⁴ FDA. *Factors to Consider Regarding Benefit-Risk in Medical Device Product Availability, Compliance, and Enforcement Decisions*. Available at <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM506679.pdf> (accessed October 25, 2018).

⁵ FDA. *Medical Device Safety Action Plan: Protecting Patients, Promoting Public Health*. Available at <https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM604690.pdf> (accessed October 17, 2018).

establishing a robust medical device patient safety net in the United States. One approach to accomplishing this goal is the use of high quality real-world evidence that FDA could use to detect emerging safety signals and take early appropriate actions to promote public health. Real-world data that could be used as clinical evidence can be derived from a number of sources, including but not limited to disease and device registries, electronic health records, medical billing and claims activities, and patient-generated health data.⁶ In 2017, FDA issued final guidance entitled *Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices*, which detailed how FDA evaluates real-world data (data relating to patient health status and/or delivery of healthcare collected from a variety of sources) and determines whether the quality of the data could support FDA medical device decision-making at different stages of the device life cycle.⁷

Why focus on Patient-Generated Health Data? With the evolution of the internet and the expansion of data sharing capabilities, patient-generated data are poised to change how healthcare is delivered and evaluated. The proliferation of smart phones and digital health technologies, including online questionnaires, mobile applications, wearable devices, and social networking sites, has increased the amount and types of patient-generated data available.⁸ Patient-generated health data is defined by the National Coordinator for Health Information Technology as “health-related data created, recorded, or gathered by or from patients (or family members or other caregivers) to help address a health concern.”⁹ This data can be derived from a variety of sources, including social media, sensors, and patient-driven registries. These technologies are poised to complement existing public health approaches to monitoring the safety of medical

⁶ FDA. Real World Evidence. Available at <https://www.fda.gov/ScienceResearch/SpecialTopics/RealWorldEvidence/default.htm> (accessed October 17, 2018).

⁷ FDA. *Use of real-world evidence to support regulatory decision-making for medical devices. Guidance for industry and Food and Drug Administration staff.* Available at <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf> (accessed October 17, 2018).

⁸ M Shapiro, D Johnston, J Wald, et al. Patient-Generated Health Data. Research Triangle Park, NC: RTI International, April 2012.

⁹ Patient-Generated Health Data | Policy Researchers & Implementers, September 30, 2015, <https://www.healthit.gov/policy-researchers-implementers/patient-generated-health-data> (accessed October 17, 2018).

devices; however, there are potential challenges to incorporating this data in postmarket regulatory activities.

Regulation of Medical Devices

The 1976 Medical Device Amendments (Public Law 94-295) to the Federal Food, Drug, and Cosmetic Act (FD&C Act) described a risk-based framework for regulating medical devices. The three-tiered risk classification system is based on the risk posed to patients should the device fail to perform as intended, leading to more regulatory controls and requirements for devices that pose a greater risk to patients. Specifically, general controls are sufficient to provide reasonable assurance of a class I device's safety and effectiveness, whereas special controls are utilized for class II devices for which general controls alone are insufficient to provide reasonable assurance of device safety and effectiveness. Most low-risk devices, such as bandages or scalpels, are exempt from FDA review before marketing, although manufacturers are still subject to certain requirements. Manufacturers of many moderate-risk devices may obtain marketing authorization by demonstrating that their devices are substantially equivalent to a legally marketed "predicate" device (in other words, a device already cleared by FDA), which can often be achieved through non-clinical testing. Class III devices are those intended to be used in supporting or sustaining human life, or for a use which is of substantial importance in preventing impairment of human health, or that which may present a potential unreasonable risk of illness or injury, and for which insufficient information exists to determine that general controls and special controls are sufficient to provide reasonable assurance of the safety and effectiveness of a device. This highest-risk class of devices is subject to premarket approval (PMA) to demonstrate a reasonable assurance of safety and effectiveness.¹⁰

For all devices, including the highest-risk class of devices, the evidence FDA requires for premarket approval has long been flexible, varying according to the characteristics of the device, its conditions of use, the existence and adequacy of warning and other restrictions, among other factors.¹¹ This flexibility in the amount of clinical evidence is different for devices compared to drugs and biological products for a variety of reasons, including: devices are subject to different statutory criteria, the mechanism of action and modes of failure are generally more predictable and better understood for

¹⁰ 21 USC § 360(a)(1) section 513.

¹¹ 21 CFR 860.7.

devices than for drugs and biological products, and the design process for a device is more often an iterative process based largely on rational design and non-clinical testing rather than clinical studies.

Mobile Medical Applications. While many commercially available mobile applications are designed to support patients in maintaining health, not all such mobile applications (apps) are medical devices. Software-based digital health technologies that are run on mobile platforms and meet the definition of a medical device are considered to be Mobile Medical Apps (MMA), and are regulated by FDA. Of note, recently passed legislation has changed the definition of a medical device to exclude certain software functions, including software intended for maintaining or encouraging a healthy lifestyle, and those intended to transfer, store, convert formats, or display data from medical devices.^{12,13}

MMA products may be designed for primary use by a patient, or for primary use by a clinician, or for patients and healthcare providers to share information with one another. Many patient-facing MMA products achieve their intended medical purpose by using algorithms to analyze sensor data, which may be derived from a wearable medical device, or from general-purpose sensors including cameras, accelerometers, or electrodes. Examples of these types of mobile medical devices include tremor transducers, audiometers, electronic stethoscopes, nystagmographs, EEG recorders, and skin imaging products.¹⁴

FDA uses the same risk-based approach to assure the safety and effectiveness for mobile medical apps as it does for other moderate-risk (class II) and high-risk (class III) medical devices. FDA has focused its oversight on higher-risk devices, while choosing not to actively regulate low-risk MMA products, such as those that help patients self-manage a diagnosed clinical condition without providing specific treatment suggestions.

¹² FDA. *Changes to Existing Medical Software Policies Resulting from Section 3060 of the 21st Century Cures Act Draft Guidance*. Available at <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM587820.pdf> (accessed October 26, 2018).

¹³ Public Law No: 114-255 (12/13/2016) 21st Century Cures Act. Available at <https://www.congress.gov/bill/114th-congress/house-bill/34> (accessed October 26, 2018)

¹⁴ FDA. *Examples of MMAs the FDA regulates*. <https://www.fda.gov/MedicalDevices/DigitalHealth/MobileMedicalApplications/ucm368743.htm> (accessed October 18, 2018).

How does FDA monitor ongoing safety of products in widespread use on the U.S. market? Medical device postmarket surveillance includes the ongoing, systematic collection, analysis, interpretation, and dissemination of health-related data to improve public health and reduce morbidity and mortality. Medical devices present challenges that are different than those encountered when performing surveillance of drugs and biologics due to the great diversity and complexity of medical devices, the learning curve associated with technology adoption, and the iterative development process leading to relatively short product life cycles. Once a medical device is available on the US market, the FDA continues to evaluate the safety and effectiveness under its existing regulatory authorities. The FDA uses multifaceted postmarket surveillance methods and approaches, including:

- *Medical device reports (MDRs)*, which are reports of certain adverse events and product problems, including device malfunctions. The MDR regulation (21 CFR 803) contains mandatory requirements for manufacturers, importers, and device user facilities (e.g., hospitals and ambulatory surgical facilities, nursing homes, outpatient treatment and diagnostic facilities) to report certain device-related adverse events and product problems to the FDA.¹⁵ Each year the FDA receives several hundred thousand reports from mandatory reporters which includes manufacturers, device user facilities, and importers and voluntary reporters (for example, healthcare professionals, patients, caregivers, and consumers). MDR reports should provide information regarding the condition of the patient, the suspected device, and an adverse experience suspected to be due to the device, including malfunctions that would be likely to cause or contribute to a death or serious injury if it were to recur.¹⁶ In addition, MDRs rely on an individual to identify that a problem occurred, to realize that the problem may have been associated with use of the device, and to take the time to report the incident to FDA and/or the manufacturer. While MDRs are a valuable source of information, there is significant underreporting of events, lack of denominator (exposure) data, lack of report timeliness, and the potential submission of incomplete or inaccurate data.

¹⁵ FD&C Act § 519; 21 C.F.R. Part 803.

¹⁶ FDA. Medical Device Reporting (MDR).

<https://www.fda.gov/MedicalDevices/Safety/ReportaProblem/default.htm> (accessed on October 17, 2018).¹⁷FDA. MedSun: Shining a light on medical product safety.

<https://www.fda.gov/MedicalDevices/Safety/MedSunMedicalProductSafetyNetwork/default.htm> (accessed October 18, 2018).

- *Medical Product Safety Network (MedSun)* is an adverse event reporting program launched with the goal to work collaboratively with the clinical community to identify, understand, and solve problems with the use of medical devices.¹⁷ When a problem is identified with a medical device, MedSun analysts work with each facility's representatives to clarify and understand the problem. The lessons learned are shared with the clinical community and the public so that clinicians nationwide may take necessary preventive actions. Similar to MDRs, this process requires the voluntary participation of hospitals and related healthcare facilities to identify and report an adverse event which is subject to underreporting.
- *Post-approval studies (PAS)* are studies that may be required as a condition of approval of a PMA, Humanitarian Device Exemption (HDE), or product development protocol (PDP) application to help assure continued safety and effectiveness (or continued probable benefit in the case of an HDE) of the approved device.^{18,19} A sponsor's failure to comply with the PAS requirement may be grounds for withdrawing approval. PAS can be challenging to conduct if the data are not otherwise being collected in the routine provision of healthcare. The absence of incentives for patients to enroll and remain in a clinical study once a device has been approved may lead to inadequate sample sizes and missing data at the time of data analysis.
- *Postmarket surveillance under Section 522* are studies that FDA may require for a class II or class III device under Section 522 of the FD&C Act (21 U.S.C. § 360I) in the following instances²⁰:
 - failure of the device would be reasonably likely to have a serious adverse health consequence;

¹⁷FDA. MedSun: Shining a light on medical product safety. <https://www.fda.gov/MedicalDevices/Safety/MedSunMedicalProductSafetyNetwork/default.htm> (accessed October 18, 2018).

¹⁸ 21 C.F.R. § 814.82.

¹⁹ FDA. Post-Approval Studies. <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/PostmarketRequirements/PostApprovalStudies/default.htm> (accessed October 18, 2018).

²⁰FDA. *Postmarket Surveillance Under 522 of the Federal Food, Drug, and Cosmetic Act. Guidance for Industry and Food and Drug Administration Staff*. Issued May 16, 2016. Available at <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM268141.pdf> (accessed October 18, 2018).

- expected to have significant use in pediatric populations;
- intended to be implanted for more than one year; or
- intended to be life-sustaining or life-supporting devices used outside a device user facility.

These studies are typically ordered to help address a potential safety issue noted in the postmarket period (so called “for cause” studies). They may also be required at the time of device clearance or approval and as a condition of clearance or approval for devices expected to have significant use in pediatric populations. Of note, these studies are often challenging for reasons similar to those listed for PAS.

- *Premarket approval application annual reports* are submitted to FDA as one tool to assure the continuing safety and effectiveness of a medical device after it is distributed. These reports contain information about device distribution, manufacturing changes, design changes, and labeling changes made during the preceding year for the product reviewed under the PMA.²¹ While this is a viable approach to collecting safety information, it only represents a small fraction of all medical devices on the US market. In addition, this information is provided annually, potentially leading to delays in detecting safety concerns.
- *Review of the scientific literature* is part of the signal evaluation efforts undertaken by FDA. A signal represents a new potentially causal association or a new aspect of a known association between a medical device and an adverse event or set of adverse events. Scientific literature, while a rich source of information, can be a delayed process subject to publication cycles and the motivation of clinicians to write the manuscript. CDRH’s Signal Management Program is discussed in more detail below.
- *Inspection of device establishments for compliance with quality system and other applicable requirements.* The FDA conducts inspections of foreign and domestic medical device manufacturers, to ensure they are complying with medical device regulatory requirements, including having established methods for collecting MDRs and appropriate correction and removal procedures in place. When there are observations that these processes are inadequate, FDA responds with tools that include warning letters and recalls. In most cases, the company (manufacturer, distributor, or other responsible party) will

²¹ 21 CFR 814.82(a)(7) and 814.84(b). Section 814.84(b)

take voluntary corrective actions to address the violations identified by the FDA.²²

- *Manufacturer reports of corrections and removals* are issued based on the potential threat to the public, also known as recalls.
- *Complaints and allegations made by members of the public* may include misleading promotion of the device, failure to follow quality system requirements, marketing uncleared or unapproved devices, or failure to register and list a medical device.²³ This process helps FDA identify potential risks to the public and determine whether further investigation is warranted. Oftentimes these complaints are made by competitor companies.

While useful, these tools have limitations. Therefore, FDA has taken several important steps to enhance our ability to assure the safety of medical devices.

One of these modern enhancements includes the CDRH Signal Management Program. This program established in October 2012 helps to ensure consistency, efficiency, accountability, and transparency in how CDRH evaluates and addresses signals related to marketed medical devices. Signal management also provides an avenue to transfer new postmarket information to the premarket review process, so that device safety concerns are considered before similar devices reach the marketplace. Following identification of a signal, a team of multi-disciplinary subject matter experts is convened to refine, research and understand the issue, then determine the appropriate public health and/or regulatory actions to mitigate the identified risks. Since the program's inception, more than 150 signals have been evaluated, resulting in numerous public health and/or regulatory

²² FDA. *Medical Device premarket approval and postmarket inspections—Part III: Inspectional*. <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/MedicalDeviceQualityandCompliance/ucm296095.htm> (accessed October 18, 2018).

²³ FDA. *Allegations of regulatory misconduct form*. <https://www.fda.gov/MedicalDevices/Safety/ReportingAllegationsofRegulatoryMisconduct/ucm526129.htm>

actions taken, the most common being the issuance of a public safety communication.²⁴

Why Use Real-World Evidence in Regulatory Decision-Making? Real-world evidence—derived from multiple sources outside typical clinical research settings (such as electronic health records, claims and billing activities, product and disease registries, or health-monitoring devices)—provides an immense new set of information about medical devices, and it plays an increasing role in health care decisions. Under the right conditions, real-world evidence may be suitable to support clearance or approval of a new device, or the expansion of indications for the use of devices that are already on the market. FDA has over 50 recent examples of such use. In addition, aggregation of real-world data such as that obtained from medical device registries is already proving useful for ongoing device safety surveillance and additional evidence for effectiveness. In particular, leveraging real world data sources has helped address the current challenges with patient enrollment in post-approval and other postmarket studies. To realize the full promise of real-world evidence, FDA has sought to clarify what it is, what it can reveal, and how it can be used most effectively at various stages of the device life cycle. Not only may this information lead to more effective regulatory decision-making, but it may also foster innovation in the medical device ecosystem.²⁵

How can FDA use NEST? Recognizing the need to optimize postmarket data collection, quality, completeness, and analysis, FDA has championed the creation of the National Evaluation System for health Technology (NEST). NEST is operated by the NEST Coordinating Center which resides within the Medical Device Innovation Consortium (MDIC).²⁶ In fact, FDA has invested in projects and related efforts to strengthen the system's capabilities through partnership with the public and private sectors.²⁷ The

²⁴ FDA. *Medical Device Safety Action Plan*.

<https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM604690.pdf> (accessed October 18, 2018).

²⁵ FDA. *Medical Device Safety Action Plan*.

<https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM604690.pdf> (accessed October 18, 2018).

²⁶ MDIC. *NEST coordinating center, who we are*. <https://nestcc.org/about/who-we-are/> (accessed October 18, 2018).

²⁷ FDA. *Medical Device Safety Action Plan*.

<https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM604690.pdf> (accessed October 18, 2018).

role of NEST is to build a network of collaborators and stakeholders across the medical device ecosystem (regulators, industry, patient groups, payors and health systems) committed to advancing the use of high quality real world evidence generated in the routine course of patient care through active “real time” device surveillance and evaluation. This network system complements the passive surveillance approaches currently used by FDA to more efficiently leverage real-world evidence for medical device evaluation and regulatory decision-making. **Table 1** lists some of the anticipated benefits of NEST to the various healthcare stakeholders.

Table 1. Anticipated Benefit of NEST to Augment Medical Device Evaluation and Postmarket Monitoring

Stakeholder Group	Value
Healthcare providers, patients, & caregivers	Enable more informed decisions based on better information about evolving benefit-risk profile of devices on the market
Researchers & Industry	Provide another source of information to assess safety and effectiveness of their devices over the total product lifecycle, and support innovative improvements (device modifications) and expanded uses (label changes)
Regulators	Improve quality of real-world evidence and capabilities to detect emerging safety signals quickly
Payors	Access to additional data sources for coverage and value analysis along with other decision-making purposes

Derived from FDA Medical Device Safety Action Plan.

This system is intended to link data from various registries, electronic health records, and billing claims, such that the data can be applied to decision making across the product lifecycle. NEST recently issued a call for test cases using patient-generated health data, highlighting the importance of patient-generated health data to not only strengthen the patients’ voice in the regulatory context, but also to provide an opportunity for augmented active surveillance of medical devices.

Hence, FDA has tools available to ensure that medical devices are safe once on the US market and is actively exploring additional novel tools such as patient-generated health data to potentially enhance postmarket surveillance.

The following sections of this Executive Summary provide a summary of the different types of patient-generated health data and the challenges with using this type of data to inform regulatory actions as provided in the published literature. It is intended for discussion purposes only and does not represent draft or final guidance. It is not intended to propose or implement policy changes regarding the use of patient-generated health data for regulatory activities or the applicable statutory and regulatory requirements for supporting a regulatory decision.

Overview of Patient-Generated Health Data

Patient-generated health data is defined as health-related data (such as health history, symptoms, biometric data, treatment history, lifestyle choices, and other information) that is created, recorded, gathered, or inferred by or from patients or their designees (which are often caregivers or those who assist them) to help address a health concern.²⁸ It is distinct from traditional data-gathering mechanisms where data is generated in clinical settings, through encounters with providers, and in healthcare provider-driven research in two important ways:

1. Patients, not providers, are primarily responsible for capturing or recording these data; and
2. Patients decide how to share or distribute these data to health care providers and other stakeholders.²⁹

The data may be written, audiovisual, or may be physiological and environmental data recorded on a monitoring device. This data that is generated, recorded, and collected by patients outside of the clinical setting can be characterized in the following ways:

Format of the Data

Structured data is comprised of clearly defined data that can be easily sorted, queried, recalled, analyzed, and manipulated by machines.³⁰

²⁸ M Shapiro, D Johnston, J Wald, et al. *Patient-Generated Health Data*. Research Triangle Park, NC: RTI International, April 2012.

²⁹ M Shapiro, D Johnston, J Wald, D Mon. *Patient-Generated Health Data*. Research Triangle Park, NC: RTI International, April 2012.

Examples include data in standardized formats, such as weight, steps, and blood glucose.

Unstructured data may also be human- or machine- generated, but it is not as easily searchable. It includes formats like audio, video, narrative/free text fields and social media postings. Search and subsequent analysis is more challenging.

Data Collection Method

Data can be collected and reported via passive or active means. **Passive data collection** occurs without any patient interaction and is often acquired from sensors (e.g., heart rate). **Active data collection** requires some action on the patients' part such as data entry, survey completion, or uploading documents.

Data Accessibility

Some sites aggregate patients' uploaded data. Patients may choose to share their data for research purposes before uploading or may choose to opt in or out at a later date.

Intended use

Certain platforms aggregate data for research purposes and actively seek out partnerships with academia and industry. Other platforms are designed to allow the patient community to connect with one another, or to allow individuals to track their personal progress.

Integrity

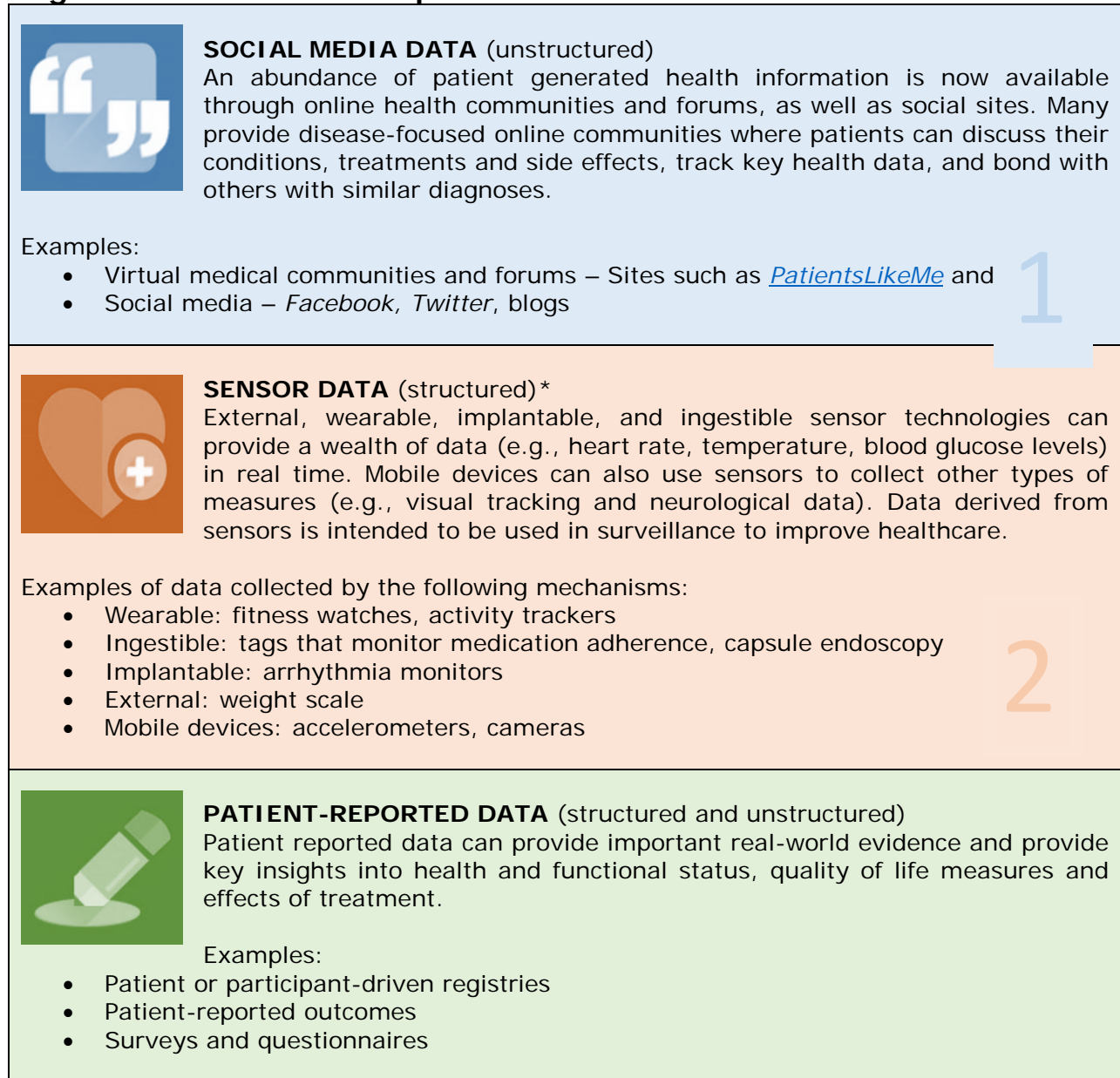
Integrity is the property of data to be accurate, complete and consistent, not being improperly modified.³¹ This characteristic reflects the validity of the data.

Some examples of patient-generated health data are shown in **Figure 1**. The following sections will discuss ways in which this data can be analyzed and potentially used for postmarket surveillance purposes.

³⁰ CTTI. *CTTI Recommendations: advancing the use of mobile technologies for data capture & improved clinical trials*. <https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/mobile-devices-recommendations.pdf> (accessed October 18, 2018)

³¹ FDA. *Content of Premarket Submissions for Management of Cybersecurity in Medical Devices Draft Guidance*. Available at <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM623529.pdf> (accessed October 25, 2018).

Figure 1. Potential Examples of Patient-Generated Health Data



*Not all devices that collect sensor data are regulated as medical devices.

SOCIAL MEDIA LISTENING

The internet has increased access to health-related information. With the advent of Web 2.0, internet users have been able to share, co-create, discuss, and modify user-generated content (such as texts, images, audios, videos, games) employing mobile and web-based technologies.³²

Usage of social media in the U.S. In this context, social media adoption has exponentially increased over time (**Figure 2**) with 69% of the public using some type of social media site.³³ The demographic of social media adopters has also begun to better reflect the general population; however, there are still differences with respect to gender, income, education, and type of community (urban, suburban, or rural).³⁴ The most frequently used social media sites are Facebook and YouTube (**Figure 3**). Social media are used by patients, consumers, and health care providers to share health issues, with 72% of adult internet users reporting searching online for health information, 26-29% reading someone's commentary about a health or medical issue, and 32% posting about health experiences of friends or family members.^{35,36} User-generated content such as blogs, microblogs (e.g., Twitter), forums, message boards, wikis, and podcasts are commonly used by patients for emotional support, information, self-esteem support, network support, social comparison, and emotional expression (**Table 2**).³⁷

³² JH Kietzmann, K Hermkens, IP McCarthy et al. Social media? Get serious! Understanding the functional building blocks of social media. *Bus Horiz* 2011;54:241-51.

³³ Pew Research Center. *Social media fact sheet*. February 5, 2018. <http://www.pewinternet.org/fact-sheet/social-media/> (accessed October 18, 2018).

³⁴ Pew Research Center. *Social media fact sheet*. February 5, 2018. <http://www.pewinternet.org/fact-sheet/social-media/> (accessed October 18, 2018).

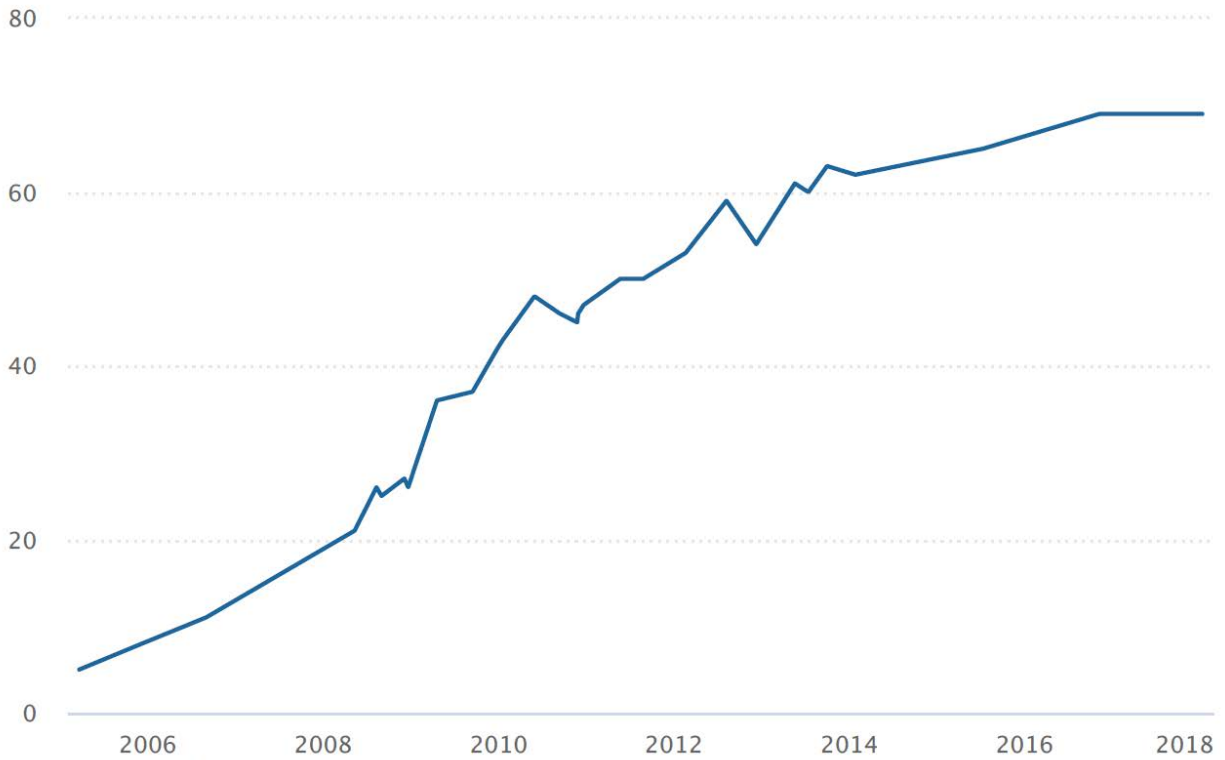
³⁵ S Fox. *The social life of health information*. Washington: Pew Research Center Internet, Science and Technology. May 12, 2011. <http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/> (accessed October 18, 2018).

³⁶ ReferralMD. *30 Facts & statistics on social media and healthcare*. <https://getreferralmd.com/2017/01/30-facts-statistics-on-social-media-and-healthcare/> (accessed October 18, 2018).

³⁷ E Smailhodzic, W Hooijsma, A Boonstra, et al. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. *BMC Health Serv Res* 2016;16:442.

Figure 2. Use of Social Media Over Time in the United States³⁸

% of U.S. adults who use at least one social media site



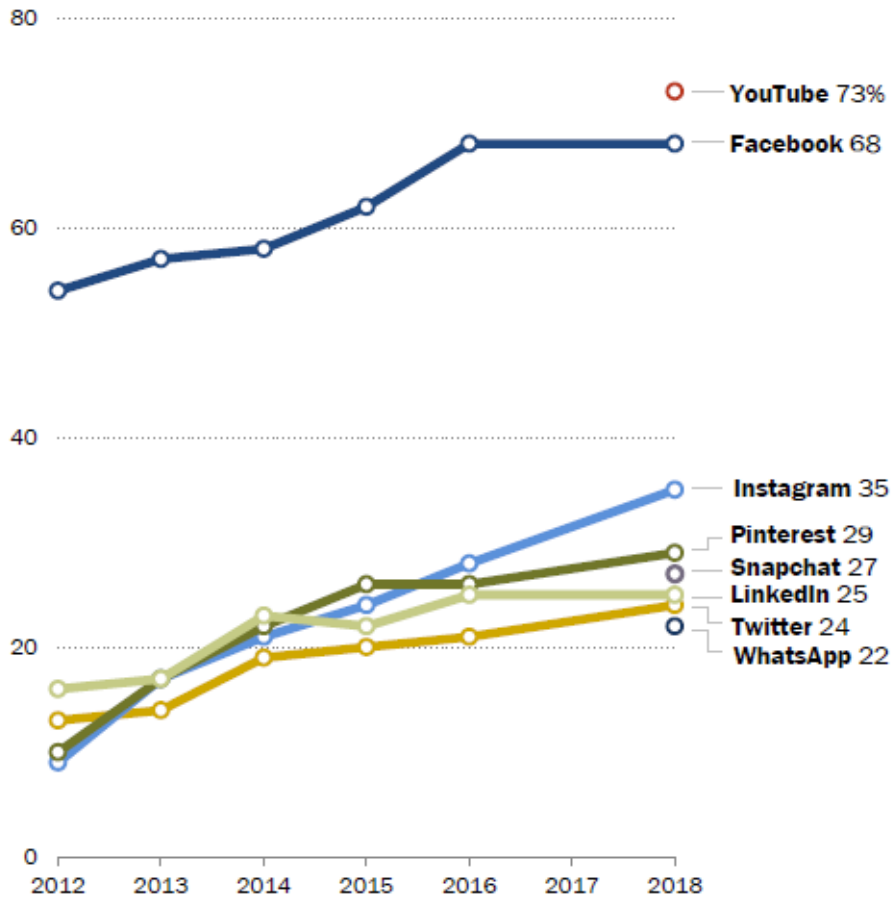
Source: Surveys conducted 2005-2018.
PEW RESEARCH CENTER

³⁸ Pew Research Center. *Social media fact sheet*. February 5, 2018. <http://www.pewinternet.org/fact-sheet/social-media/> (accessed October 18, 2018).

Figure 3. Most Commonly Used Social Media Sites in the US³⁹

Majority of Americans now use Facebook, YouTube

% of U.S. adults who say they use the following social media sites online or on their cellphone



Note: Pre-2018 telephone poll data is not available for YouTube, Snapchat or WhatsApp.
Source: Survey conducted Jan. 3-10, 2018. Trend data from previous Pew Research Center surveys.

"Social Media Use in 2018"

PEW RESEARCH CENTER

³⁹ Pew Research Center. *Social media fact sheet*. February 5, 2018. <http://www.pewinternet.org/fact-sheet/social-media/> (accessed October 18, 2018).

Table 2. Common Social Media Platforms

Type	Characteristics	Examples
Social networks	Website that allows the user to build a Web page and connect with a friend or other acquaintance in order to share user-generated content	Facebook, LinkedIn, Google Plus+, Meetup
Blogs and blog comments	An online journal where the user (“blogger”) can create the content and display it in reverse chronological order. Blogs are generally maintained by a person or a community	Huffington Post, Business Insider
Microblogs	Similar to a blog but with content size limitations	Twitter, Tumblr
Forums	Place for members to discuss a topic by posting messages	Online discussion communities, Raising Children Forum
Social bookmarks	Services that allow users to save, organize, and search links to various websites, and to share their bookmarks of Web pages	Pintrest, Google Bookmarks
Wikis	Collaborative sites where users can add or edit content on a community-based database	Wikipedia, Wikihow, Wikitravel
Social news	Service that encourages their community to submit news stories or to vote on the content and share it	Reddit, Digg, Slashdot
Media Sharing	Site that enables users to capture videos and pictures or upload and share with others	YouTube, Flickr, Instagram, Vine

Adapted from A Farzindar and D Inkpen.⁴⁰

⁴⁰ A Farzindar and D Inkpen. Natural language processing for social media In Synthesis Lectures on Human Language Technologies #30: Editor G Hirst. San Rafael, CA: Morgan & Claypool Publishers, 2015.

Processing social media data for analysis. As people increasingly use social media for health information, these platforms may potentially be used as sources of observational data for postmarket surveillance of medical products. To effectively curate the information and explore meaningful analyses, social media data typically undergoes processing.

- Data Acquisition: The social media data used for postmarket surveillance activities could be stored textual information, dynamic online data collection processed in real time, or retrospective data collection. The data includes not only the text but also the rich metadata that accompanies it. The metadata may help localize the users because it contains the time and date stamps. Often publicly available data (such as Twitter) are obtained using the sites' public Application Programming Interfaces (APIs) which are commercially available from authorized data resellers. Webcrawler, a metasearch engine that blends the top search results from Google Search and Yahoo! Search, is another approach that has been used to assemble the relevant webpages, after which a web scraper has been used to obtain the unstructured text embedded in the webpages.⁴¹
- Translation, Filtration & Extraction of Data. The volume of social media data and the exponential rate at which it is created makes human evaluation and analysis of the data untenable. Social media is much noisier than traditional print media, plagued with spam, ads and unsolicited, irrelevant and distracting content. Additional challenges that complicate analysis of the data include the inconsistent or absent punctuation and capitalization which impede differentiation of sentences; use of emoticons; incorrect or non-standard spelling; abbreviations; and the informal conversational "stream of consciousness" tone.⁴²

In order to overcome some of these challenges, Tricco et al. suggest a variety of data processing approaches be applied to create analyzable data sets including the following:⁴³

⁴¹ J Lardon, R Abdellaoui, F Bellet, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res* 2015; 17:e171.

⁴² A Farzindar and D Inkpen. *Natural language processing for social media* In Synthesis Lectures on Human Language Technologies #30. Editor G Hirst. San Rafael, CA: Morgan & Claypool Publishers, 2015.

⁴³ AC Tricco, W Zarin, E Lillie, et al. Utility of social media and crowd-intelligence data for pharmacovigilance: a scoping review. *BMC Med Inform Decis Mak* 2018;18:38.

- Supervised learning, which is a machine learning approach by which an algorithm is trained from a set of labeled data that has been coded by humans. The choice of which database is appropriate for training the algorithm will depend heavily on the media site targeted; whereas the Medical Dictionary for Regulatory Activities (MedDRA) may be used in one context, dictionaries of colloquial terminology and slang (e.g., Consumer Health Vocabulary, urban dictionary) may be more appropriate in another context;
- Semi-supervised learning, which is a machine learning method that typically supplements a small amount of labeled data with a larger volume of unlabeled data for the purposes of training an algorithm; and
- Unsupervised learning, which is a machine learning approach in which an algorithm identifies patterns in data that have not been labeled or classified and uses those patterns to establish associations between data points.

Regardless of the processing approach used, Tricco et al. emphasizes that the text also has to be pre-processed, which includes removing punctuation and stop words, correcting spelling errors, reducing words to the root, tagging the parts-of-speech, breaking the text into words, phrases, and symbols or tokens. The data is also deidentified (removal of personally identifying information such as screen names, user names, first and last names, and addresses), de-duplicated (removal of retweets), and de-spammed (removal of advertisements and vulgar content).⁴⁴

Benefits and challenges of using social media data for monitoring postmarket safety. Researchers increasingly use social media data to better understand trends in healthcare and outcomes with medical products. One review article cited the exponential increase in publications examining adverse events being captured in social media, noting only 9 publications in the literature during the years 2002 to 2010 rapidly increasing to 23 publications in 2014 alone.⁴⁵ In general, adverse events mentioned in social

⁴⁴ AC Tricco, W Zarin, E Lillie, et al. Utility of social media and crowd-intelligence data for pharmacovigilance: a scoping review. *BMC Med Inform Decis Mak* 2018;18:38.

⁴⁵ S Golder, G Norman, YK Loke. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *British Journal of Clinical Pharmacology* 2015;80:878-88.

media were concordant with those already documented in other sources such as published trials and medical product labels; however, the results of these analyses often concluded that the level of reporting for the same adverse events was much higher in social media. Unlike traditional studies, social media data can offer insights into the performance of medical products in a shorter time frame.⁴⁶ One study has compared the timeliness of identifying adverse events from social media with other sources and found that the events would have been detected earlier if social media listening was part of the medical product surveillance process.⁴⁷

Some authors have suggested that social media could be used to supplement traditional reporting systems, to uncover adverse events which are less frequently reported in traditional reporting systems, to communicate risk to the public and to generate hypotheses.⁴⁸ Golder et al. suggested that social media may be a better source for “symptom-related” or less “serious” adverse events.⁴⁹ However, challenges exist with the use of social media, such as difficulties interpreting relationships between the devices and adverse events (in other words, there are inadequate data to draw causality), potential lack of representativeness between social media users and the general population, and the resource-intensive process of using social media data for medical device surveillance.

Validity of social media data may be expanded by the development of formal ontologies that account for slang and other vernacular terminology and how it is used to represent medical concepts, how this language may vary from region to region, and how it is influenced by each type of social media site.⁵⁰ In addition, the terms used to describe adverse events may also be used to describe the condition being treated, beneficial effects (for example, sleepiness is beneficial for someone with insomnia), or may not have been experienced by the patient (such as talk about fears of developing

⁴⁶ Y Kim, J Huang S Emery. Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology, and digital disease detection. *J Med Internet Res* 2016;18: e41.

⁴⁷ CC Yang, H Yang, L Jiang. Postmarketing drug safety surveillance using publicly available health-consumer-contributed content in social media. *ACM Trans Manag Inf Syst* 2014;5: 1-21.

⁴⁸ AC Tricco, W Zarin, E Lillie, et al. Utility of social media and crowd-intelligence data for pharmacovigilance: a scoping review. *BMC Med Inform Decis Mak* 2018;18:38.

⁴⁹ S Golder, G Norman, YK Loke. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *British Journal of Clinical Pharmacology* 2015;80:878-88.

⁵⁰ GE Powell, HA Seifert, T Reblin et al. Social media listening for routine post-marketing safety surveillance. *Drug Saf* 2016;39:443-54.

diabetes).⁵¹ While there is much enthusiasm for this potentially rich source of data, Kim et al. warn that a large quantity of data does not assure valid and reliable results and may lead to overpredicting or underpredicting the outcome of interest.^{52,53,54} The ability to leverage and incorporate social media into the medical device surveillance strategy is largely unexplored and additional evaluation and validation studies are needed.

SENSOR-BASED DIGITAL HEALTH TECHNOLOGIES

Digital health technologies can empower consumers in making health decisions and provide new opportunities to impact disease prevention, early diagnosis, and effective management of chronic conditions.⁵⁵ While the broad scope of digital health technologies includes mobile medical applications, wearable devices, and telehealth, digital health technologies most directly involved in the generation of patient-generated health data often rely on data derived from sensors. The miniaturization of sensors and their accompanying circuitry, together with the decreased cost of production, has led to the proliferation of novel medical devices and consumer products that are designed to improve patients' health and overall wellness. Sensor data can be gathered from external, wearable, implantable, and ingestible technologies to provide a wealth of information. Sensors can identify environmental exposure (such as indoor smoke), location, physical activity (via accelerometry), sleep, social interactions, images, visual stimuli, glucose levels, and heart rhythms, with many more measures in development. This technology may provide new ways to assess clinical outcomes at a higher frequency, outside of structured research settings, and

⁵¹ S Golder, G Norman, YK Loke. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *British Journal of Clinical Pharmacology* 2015;80:878-88.

⁵² Y Kim, J Huang S Emery. Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology, and digital disease detection. *J Med Internet Res* 2016;18: e41.

⁵³ D Lazer, R Kennedy, G King, et al. Big data. The parable of Google Flu: traps in big data analysis. *Science* 2014; 343:1203-5

⁵⁴ IC Fung, ZTH Tse, C Cheung, et al. Ebola and the social media. *Lancet* 2014;384:2207.

⁵⁵ FDA. *Digital Health Innovation Action Plan*. Available at <https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/UCM568735.pdf> (accessed on October 18, 2018).

during activities of daily living. The wearable and implantable sensor market is predicted to grow ten-fold over the next five years (**Table 3**).⁵⁶

Table 3. Digital Connections

	2010	2015*	2020*
World Population, billion	6.8	7.2	7.6
Number connected			
Devices, billion	12.5	25	50
Devices per person	1.8	3.5	6.6
Number of smartphone subscriptions, billions	0.5	3.0	6.0
Number of sensors	20 million	10 billion	1 trillion

*Adapted from EJ Topol et al.*⁵⁷

**Numbers are estimates based on Moore's law.*

Opportunities for using sensor-based digital health technologies in research and monitoring. Given the recent emergence and widespread public adoption of these technologies, researchers are exploring ways to integrate it into clinical trials, real-world data platforms, and healthcare modules. Perry et al. conducted a systematic literature review on the use of sensors and clinical outcomes measured by sensors (mobile outcomes) in observational and interventional clinical research.⁵⁸ The most commonly used sensors were wearable inertial sensors/accelerometers, continuous glucose monitors, ingestible pH monitors, pressure sensors and instrumented walkways, and medication adherence monitors. Similar to other tools used to capture a given concept (e.g., visual acuity measured by a Snellen acuity chart), Byrom et al. have asserted that the concept that is captured by the sensors should be clearly defined, reliability and validity assessed, and be fit for the purpose for which it is integrated in the research

⁵⁶ JA Levine. The Baetylus theorem—the central disconnect driving consumer behavior and investment returns in wearable technologies. *Technology Investments* 2016;7:59-65.

⁵⁷ EJ Topol, SR Steinhubl, A Torkamani. Digital medical tools and sensors. *JAMA* 2015;313:353-4

⁵⁸ B Perry, W Herrington, JC Goldsack, et al. Use of mobile devices to measure outcomes in clinical research, 2010-2016: a systematic literature review. *Digit Biomark* 2018;2:11-30.

protocol.⁵⁹ Byrom et al. considers the following to be the minimum evidence to show that the sensor data has suitable measurement properties:

1. *Content validity*. The sensor demonstrates that it is capturing a meaningful aspect of the disease/condition or treatment from the patient's perspective. For example, if the concept of interest is improved physical function and the sensor used is an activity monitor, then the measure derived from the activity monitor data should be viewed by patients as relevant to their condition and to an improvement in physical function.
2. *Reliability*. Intra-device and inter-device agreement including calibration methods.
3. *Concurrent (criterion-related) validity*. Evidence that the sensor data correlates with another measure that is viewed as being more accurate in patients with similar characteristics to the intended patient population.
4. *Responsiveness*. Ability to detect change when a change exists.

Sensor design considerations. When determining which sensor to use or whether a sensor is appropriate to address a given safety concern, Byrom et al. encourages researchers to consider the following factors⁶⁰:

- Wear location, appearance, design comfort, duration of wearing, and ease of use all may impact the acceptability of the sensor;
- Measurement period since some devices may be inconvenient if worn for long period;
- Battery length and storage capacity coupled with the need to charge the sensor could impact data quality and lead to missing data;
- Visibility of the data to the wearer could alter behaviors, essentially amounting to a health, potentially affecting conclusions drawn from the data;

⁵⁹ B Perry, W Herrington, JC Goldsack, et al. Use of mobile devices to measure outcomes in clinical research, 2010-2016: a systematic literature review. *Digit Biomark* 2018;2:11-30.

⁶⁰ B Byrom, C Watson, H Doll, et al. Selection of and evidentiary considerations for wearable devices and their measurements for use in regulatory decision making: recommendations from the ePRO consortium. *Value in Health* 2018;21:631-9.

- Connectivity of the sensor to support data transmission to the platform via direct wireless connectivity, Bluetooth connectivity to a mobile device, or web download through connected charging base;
- Sensor's data storage volatility (for example, will the data be lost when the device is switched off) and ability to overwrite with new data when the device's storage is full;
- Changing technological landscape may nullify sensors in use today, thereby posing the challenge of multiple sensors over time measuring the same concept; and
- Updates to sensors may apply revisions of algorithms used to derive outcome measures reported by the sensor which could impact the integrity of the data and the ability to compare data collected before and after the update.

PATIENT-DRIVEN REGISTRIES

A patient registry is an organized system that uses observational study methods to collect uniform data to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure that serves a predetermined scientific, clinical or policy purpose. The registry database is the file(s) derived from the registry.⁶¹ The purpose of the registry may span many areas including:

- recruiting patients for clinical trials;
- inform targeting of future therapeutics;
- understand population behavior patterns and their association with disease development;
- understanding the natural history of a disease or condition;
- developing research hypotheses; and
- improving and monitoring device performance, including conducting active surveillance.

The defining characteristics of patient- or participant-driven (also called "patient-run," "patient-generated," "patient-powered," and "participant-controlled") registries are that they are

- founded, created, owned, and managed by patients or healthcare consumers that participate in the registry;
- primarily focused on the needs and goals of the registry participants;

61 Gliklich RE, Dreyer NA, eds. *Registries for Evaluating Patient Outcomes: A User's Guide*. (Prepared by Outcome DEcide Center [Outcome Sciences, Inc. dba Outcome] under Contract No. HHS29020050035ITO1.) AHRQ Publication No. 07-EHC001-1. Rockville, MD: Agency for Healthcare Research and Quality. April 2007.

- employ social networks and/or community-based engagement to recruit participants; and/or
- collecting participant-reported health data in a structured and/or unstructured format in addition to data generated during the healthcare encounter.⁶²

Analyses of real-world data, using appropriate methods, may in some cases provide similar information with comparable or even superior characteristics to information collected and analyzed through a traditional clinical trial. While clinical trials are often used to generate new knowledge on novel medical devices, they are limited by assessing a fixed number of patients in a highly structured clinical encounter for a predetermined length of time. These studies may also be impractical or excessively challenging to conduct. Many safety events associated with a medical device may not be observed in the clinical trial. Real-world data, such as the information collected in patient registries, may provide valid scientific evidence to guide public health interventions and provide information on emerging signals.

Considerations for developing successful patient-driven registries.

Gliklich et al. identified that the key elements of any registry include the characteristics listed in **Figure 4**.⁶³ According to Terry et al., well-designed technological platforms facilitate the creation of successful patient-driven registries that enable patients to: join the network; report, store, and display information; search for patients with similar experiences or conditions; and link to other resources.⁶⁴ Registries that collaborate with academicians and other scientific advisors tend to be more successful in

⁶² S Terry, D Runkle, P Wicks. Patient- or Participant-generated registries. White Paper, addendum to Registries for evaluating patient outcomes: a user's guide, third edition. (Prepared by L & M Policy Research, LLC, under Contract No 290-2014-00004-C) AHRQ Publication No. 17(18)-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018. www.effectivehealthcare.ahrq.gov.

⁶³ Gliklich RE, Dreyer NA, eds. *Registries for Evaluating Patient Outcomes: A User's Guide*. (Prepared by Outcome DEcIDE Center [Outcome Sciences, Inc. dba Outcome] under Contract No. HHS29020050035ITO1.) AHRQ Publication No. 07-EHC001-1. Rockville, MD: Agency for Healthcare Research and Quality. April 2007.

⁶⁴ S Terry, D Runkle, P Wicks. Patient- or Participant-generated registries. White Paper, addendum to Registries for evaluating patient outcomes: a user's guide, third edition. (Prepared by L & M Policy Research, LLC, under Contract No 290-2014-00004-C) AHRQ Publication No. 17(18)-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018. www.effectivehealthcare.ahrq.gov.

generating data that can be used to conduct research as well as increasing awareness across the relevant researcher community.⁶⁵

Patient-driven registries are often initiated by individuals affected by a disease or condition and/or an advocacy organization dedicated to a condition. Many patient-driven registries not only collect participant-reported data, but also integrate data from the participants' health records, biological samples, and imaging files. Unlike a single sponsor, funder, or investigator team, the advocacy organizations are often committed to collecting the natural history of a disease across many dimensions and over many years.⁶⁶ These natural history studies can help define the disease incidence, understand the variability of the disease, identify causes of morbidity and mortality, define a patient's lifespan, and help develop and validate tests along with outcome measures that can aid in ensuring the clinical trial is adequately designed, particularly for rare diseases.

⁶⁵ S Terry, D Runkle, P Wicks. Patient- or Participant-generated registries. White Paper, addendum to Registries for evaluating patient outcomes: a user's guide, third edition. (Prepared by L & M Policy Research, LLC, under Contract No 290-2014-00004-C) AHRQ Publication No. 17(18)-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018.

⁶⁶ S Terry, D Runkle, P Wicks. Patient- or Participant-generated registries. White Paper, addendum to Registries for evaluating patient outcomes: a user's guide, third edition. (Prepared by L & M Policy Research, LLC, under Contract No 290-2014-00004-C) AHRQ Publication No. 17(18)-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018.
www.effectivehealthcare.ahrq.gov.

Figure 4. Key Components of Registries (derived from Gliklich and Dreyer).⁶⁷

⁶⁷ Gliklich RE, Dreyer NA, eds. *Registries for Evaluating Patient Outcomes: A User's Guide*. (Prepared by Outcome DEcide Center [Outcome Sciences, Inc. dba Outcome] under Contract No. HHS29020050035ITO1.) AHRQ Publication No. 07-EHC001-1. Rockville, MD: Agency for Healthcare Research and Quality. April 2007.

Planning

- What is the purpose?
- Can a registry address the research question?
- Who are the relevant stakeholders?

Registry Design

- Selection of the patients, including determining need for comparison group
- Identification of exposures and outcomes to be measured
- Number of patients needed
- Method for selecting patients (random, systematic sampling, convenience)
- Length of observation (follow up)

Data Elements

- Necessary versus desirable but not essential
- Method for assessing the elements
- Guided by parsimony, validity, and registry's purpose

Data Sources

- Primary data collected for the direct purposes of the registry
- Secondary data collected for other purposes (such as medical records, claims data, birth and death records, other registries)

Ethics, Data Ownership, and Privacy

- Consent processes
- Health Insurance Portability and Accountability Act (HIPAA) protect the privacy of patients
- Oversight and data ownership

Recruitment and Management

- Goals for recruitment, retention, and follow-up explicitly detailed

Data Collection and Quality Assurance

- Method for handling data problems (missing, out of range, logically inconsistent values)

Analysis and Interpretation

- Representativeness of actual studied population to the target population (generalizability)
- Specified statistical analysis plan

Using patient-driven registry data for research and monitoring. Data in the registry also may be used to identify participants who meet the inclusion criteria for a planned trial, to assess the feasibility of enrolling the intended sample size per protocol criteria, to determine desirable geographic locations of enrollment centers and to facilitate trial enrollment by sharing information about the trial with participants who meet the inclusion criteria. It has been asserted that the patient-driven registries may serve as an important channel for participants to learn the results of their own involvement in research, as well as those of other clinical trials for the condition of interest.⁶⁸ In addition to identifying patients for trial participation, patient-driven registries have been successful in establishing the natural history of a disease, which could potentially be used as a comparator in the evaluation of medical products, thereby streamlining the time and cost of clinical trials.^{69,70,71}

Patient-driven registries may be a platform for patients to express what outcomes matter to them and facilitate the creation of new patient-reported outcome (PRO) measures that can be used by other researchers. A PRO measure is an assessment that reflects the status of a patient's health condition that comes directly from the patient without amendment or interpretation by anyone else.⁷² Because these registries may have a wider spectrum of disease severity, they may allow for appropriate calibration of

⁶⁸ S Terry, D Runkle, P Wicks. Patient- or Participant-generated registries. White Paper, addendum to Registries for evaluating patient outcomes: a user's guide, third edition. (Prepared by L & M Policy Research, LLC, under Contract No 290-2014-00004-C,) AHRQ Publication No. 17(18)-EHC017-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018. www.effectivehealthcare.ahrq.gov

⁶⁹ AG Phillips, P Honggaard-Andersen, RA Moscicki, et al. Proceedings of the 2013 CINP summit: innovative partnerships to accelerate CNS drug discovery for improved patient care. *Int J Neuropsychopharmacol* 2014;18: pyu100.

⁷⁰ LB Gordon, J Massaro, RB D'Agostino et al. Impact of farnesylation inhibitors on survival in Hutchinson-Gilford progeria syndrome. *Circulation*. 2014;130:27-34.

⁷¹ EA de Blicke, EF Augustine, FJ Marshall et al. Methodology of clinical research in rare diseases: development of a research program in juvenile neuronal ceroid lipofuscinosis (JNCL) via creation of a patient registry and collaboration with patient advocates. *Contemp Clin Trials* 2013;35:48-54.

⁷² FDA. *Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims.*

<https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf> (accessed October 18, 2018).

PRO measures across the disease continuum.^{73,74} These PRO measures also may be useful tools in postmarket safety efforts.

Challenges with Patient-Generated Health Data

FDA applies a “fit-for-purpose” approach to help ensure any data considered for a specific regulatory decision is of sufficient quality.^{75,76} Patient-generated health data, like any other data should be assessed for quality attributes (completeness, consistency, accuracy) and whether it contains all critical elements needed to evaluate a medical device and its claims.⁷⁷ For example, social media posts alone arguably do not include most of the critical components for adverse event reporting, thereby potentially limiting their utility to contribute directly to an FDA regulatory action. However, they may signal a burgeoning issue that FDA could further explore using other complementary approaches. Alternatively, patient-driven registries, depending on how they are designed, may be able to provide active postmarket monitoring and complement other signal management efforts.

With all the approaches to using patient-generated health data, there are some cross-cutting challenges with using the data in regulatory efforts. Some of the challenges listed below are not limited to patient-generated health data; instead they are data challenges that may apply to many types of studies and data collection approaches. These challenges include:

⁷³ P Wicks. Commentary: Measuring what matters: the case of patient generated PROMS. *BMJ* 2015;350:h54.

⁷⁴ P Wicks. Measuring function in advanced ALS: validation of ALSFRS-EX extension items. *European Journal of Neurology* 2009;16:6.

⁷⁵ FDA. *Use of real-world evidence to support regulatory decision-making for medical devices. Guidance for industry and Food and Drug Administration staff.* Available at <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf> (accessed October 17, 2018).

⁷⁶ G Daniel, M McClellan, C Silcox et al. *Characterizing RWD quality and relevancy for regulatory purposes.* Duke Margolis Center for Health Policy, October 2018. Available at <https://healthpolicy.duke.edu/events/second-annual-duke-margolis-conference-real-world-data-and-evidence> (accessed October 26, 2018).

⁷⁷ FDA. *Use of real-world evidence to support regulatory decision-making for medical devices. Guidance for industry and Food and Drug Administration staff.* Available at <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm513027.pdf> (accessed October 17, 2018).

1. *Consent, security and privacy concerns.* Patient may opt in to using their data for one purpose (i.e., consent) but may not have been given the choice to opt out of using it for other purposes, which may violate their rights.⁷⁸ While social media listening research usually uses publicly available data sets, consent remains an issue for patient-driven registries and sensor-associated applications. Best practices and the use and acceptability of central Institutional Review Boards (IRB) or internet-facilitated shared review systems has been encouraged and explored by the National Patient-Centered Clinical Research Network (PCORnet) for patient-generated health data.^{79,80} When non-public patient-generated health data interfaces with covered entities such as healthcare systems, it is officially considered “protected health information,” (PHI) and thereby subject to HIPAA protections (Privacy Rule) as well.⁸¹

Potential risks to address include the risks of exposing the participant’s identity during data gathering; data dissemination and publication; and quoting directly from online conversations and stories. In addition, the ability to identify patients even from deidentified information using data linkage of several sources and issues with recruiting vulnerable populations such as children to participate in these patient-driven registries further highlight potential risks with contributing to patient-generated health data. These breaches could lead to HIPAA violations, identity theft, and possible discrimination by employers or insurers.⁸²

⁷⁸ DHHS. *Federal policy for the protection of human subjects (“Common Rule”).* Available at <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html> (accessed October 26, 2018).

⁷⁹ RL Florence, AC Beal, SE Sheridan, et al. Patient-powered research networks aim to improve patient care and health research. *Health Affairs* 2014;33:1212-9.

⁸⁰ PCORnet Best Practice sharing session—PCORnet commons tools for electronic consent workgroup webinar series: part 2. <https://pcornet.org/event/pcornet-best-practice-sharing-session-pcornet-commons-tools-electronic-consent-workgroup-webinar-series-part-2-best-practices/> (accessed October 18, 2018).

⁸¹ DHHS. *Summary of the HIPAA security rule.* Available at <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html> (accessed October 26, 2018).

⁸² DHHS. *Summary of the HIPAA security rule.* Available at <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html> (accessed October 26, 2018).

With the many data breaches as well as unauthorized uses of data in recent years, the United Kingdom enacted the General Data Protection Regulation (GDPR) on May 25, 2018 and requires that everyone responsible for using personal data follow strict rules called “data protection principles.”⁸³

Under the GDPR, there are stronger legal protections for biometrics (when used for identification) and health data. For example, patients have the right to be informed about how their data is being used, have their data erased, stop or restrict the processing of their data, and object to how their data is processed in certain circumstances. The GDPR may impact the ability to pool patient-generated health data from outside of the US and could significantly impact the data available to study rare diseases.

2. *Standardization of data elements* facilitates pooling of data from multiple participants and datasets and enables consistent interpretation of the analyses. There are a number of ways to standardize data elements including having a standard definition for given terms, collecting the same data elements for a given condition, and ensuring the tools that are collecting the data elements have been established as valid measures. Having a standard lexicon for adverse events, such as those used in existing health IT standards (such as MedDRA or SNOMED) allows for consistent approaches to tabulating adverse events. This standard lexicon could be used in social media listening and patient-driven registries. Consistent use of Unique Device Identifiers (UDI) can identify and standardize the collection of medical device information in patient-driven registries. These standard elements are often lacking in patient-generated health data which could make it challenging to link it to other data sources or have confidence in the quality of data being collected.

Common data elements (CDE) are data elements that are common to multiple data sets across different studies. The National Institutes of Health encourages intentional use of common data elements to help improve data quality and promote data sharing.⁸⁴ Many registries

⁸³ UK Government. Data Protection. <https://www.gov.uk/data-protection> (accessed October 18, 2018).

⁸⁴ NIH Common Data Element (CDE) Resource Portal. <https://www.nlm.nih.gov/cde/glossary.html#cdedefinition> (accessed October 18, 2018).

including patient-driven registries may not include these CDEs in their data fields, including the core set of symptoms for the given condition, thereby limiting the ability to integrate the data with other data sources.

Validated measures of outcomes are often recommended if the purpose of the data is to inform regulatory efforts. For example, if a patient-driven registry or a sensor mobile app incorporates a patient-reported outcome measure, it should have development work that supports its integration in the research study and shows that it is “fit-for-purpose”.⁸⁵

One challenge to standardizing the terminology is that it is possible to alter the language and meaning used by patients which may be important for regulatory purposes as well as healthcare treatment environments. If the patient-generated health data is being used to support the development of patient-reported outcome measures, then the language used by participants is critical and should not be standardized. By specifying the research question, investigators can ensure that their approach for utilizing and organizing the data is “fit-for-purpose”.

3. *Variable timing of data collection.* The optimal timing for collecting patient-generated health data is unclear. The data could be collected continuously or intermittently, and if it is intermittent the intervals for collection are often not specified or not consistently used across different platforms. For registries and social media, participants often enter data at time points of their own choosing rather than at fixed reporting intervals. There could be non-random drivers of data entry, such as a tendency for participants to enter data more frequently when they experience a change in symptoms, particularly a worsening of their condition, resulting in a potentially inaccurate depiction of their disease severity and a likely biased assessment of the adverse event profile.

⁸⁵ FDA. *Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims.*

<https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf> (accessed October 18, 2018).

4. *Conclusions may vary with the data analytic approach.* Sensors and social media sites generate large volumes of asynchronous or continuous streams of data, which can be challenging to manage and review without intelligent filtering and summarization. Presently, there is no standardized approach to using human or machine processing power to make sense of these large amounts of data, and different approaches can yield different answers.
5. *Missing data* is not unique to patient-generated health data, but given the sheer volume of the data, questions on how to handle the missing data and how much missing data makes the patient-generated health data (especially sensor data and social media) unusable for regulatory purposes remains unanswered. For example, if participants remove their sensor and forget to put it back on for a few days, there may be large gaps in the data for that one participant. However, if the data from the sensor is downloaded every 60 seconds, then it is unclear what the impact of those few days are on the quality of the data. In addition, excessive data checks for quality and to confirm patient identity may frustrate patients and lead to more missing data as opposed to better quality data.
6. *Interoperability.* Many sensors still use proprietary architecture, making it difficult to ensure the interoperability among the patient's multiple devices, the patient-driven registry and electronic health records. Development of new standards to enable "plug-and-play" capability between sensors and a range of electronic health records is being developed but is not yet widely accepted and used by system vendors of electronic health record platforms. In addition, pooling multiple sources of the same data can be an issue when there is variability in data capture and no existing interoperability standards to allow for seamless aggregation of data from multiple sources.
7. *Generalizability.* Low health literacy is a barrier to participating in some of these patient-generated health data sources. Sensors, mobile applications, and internet access are not commonly available across all socioeconomic status and in all parts of the country. Participants in patient-generated health data may not be representative of the patient population as a whole, yet this selective participation may or may not affect the outcome under study. There is a concern that patient-generated health data participants may be those individuals who feel

the strongest (positively or negatively) about a particular treatment, thereby producing a “reporting bias”.⁸⁶

While registries are more likely to be successful when they regularly promote and recruit patients to participate and contribute to the effort, they may foster potential competition for the same patients across multiple patient registries that exist for a single condition. This could create a fractured set of patient data, which could be particularly impactful for rare medical conditions.⁸⁷ Given the aforementioned challenges with interoperability, it may be difficult to pool data from multiple registries.

8. *Data duplication.* It is possible that the same patient could participate in multiple registries, creating challenges for aggregating the data due to the potential for that one patient to have more influence than other patients on the conclusions. Similarly, users of social media can post the same message twice or more on the same forum or on different forums using the same or different pseudonyms with no malevolent intent, but simply to maximize their chances of obtaining an answer or response.⁸⁸

9. *Data Integrity and Verifiability.* Data uploaded to social media sites may not be authentic. Sponsored groups may try to sway public and/or medical opinion through misleading or false information.⁸⁹ Validation of the diagnosis of enrolled patients by health care providers, if absent, may lead to skepticism by the healthcare community. Some patient-driven registries require verification of diagnosis from health care providers, and others link to electronic health records to help support the diagnosis.

Egregious misrepresentation (users pretending to be someone else) is a concern for all patient-generated health data. Social media users

⁸⁶ J Frost., S Okun, T Vaughan, et al. Patient-reported outcomes as a source of evidence in off-label prescribing: analysis of data from PatientsLikeMe. *J Med Internet Res* 2011;13:e6.

⁸⁷ TA Workman. Engaging patients in information sharing and data collection: the role of patient-powered registries and research networks. AHRQ Community Forum White Paper. AHRQ Publication No. 13-EHC124-EF. Rockville, MD: Agency for Healthcare Research and Quality; September 2013.

⁸⁸ J Lardon, R Abdellaoui, F Bellet, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res* 2015; 17:e171.

⁸⁹ J Lardon, R Abdellaoui, F Bellet, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res* 2015; 17:e171.

adopt pseudonyms which could allow malicious persons to spread false rumors with limited risk of being identified as the origin of the rumor.⁹⁰ Participants in patient-driven registries may pretend to have the diagnosis for various reasons. In addition, participants may purposefully misuse sensors, such as placing their sensor on another person to falsely report their physical activity. To combat these issues, diligence should be exercised to examine the metadata and data patterns to help identify misrepresentation and remove it from the analyzable dataset.

10. *Data governance.* "Who owns the data?" is a recurrent theme with the proliferation of ever-expanding data sources and potential financial benefits of marketing that data. Waller and Alcantara warned that "If what is meant by ownership of patient information is the right to exercise complete sovereignty over information, it cannot be said that any one person or entity 'owns' the information."⁹¹ They identified that the true questions are who can do what to which data under what circumstances. Because vendors are increasingly seeing the commercial value of this data, data ownership should be clarified up-front particularly for patient-driven registries and sensor data.

The challenges listed above are not an exhaustive list; however, they highlight opportunities to improve quality and applicability of patient-generated health data for postmarket monitoring.

Conclusion

In protecting the public's health, the FDA uses many approaches to ensure medical devices on the US market are safe and effective. Patient-generated health data may potentially be a complementary tool, augmenting the evidence used to make regulatory decisions. While this area is rapidly developing, sound scientific approaches are needed to

- mitigate the biases,

⁹⁰ J Lardon, R Abdellaoui, F Bellet, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res* 2015; 17:e171.

⁹¹ AA Waller, OL Alcantara. Ownership of health information in the information age. *J AHIMA* 1998;69:28-38.

- analyze the data,
- assess the quality of the data, and
- explore potential applications of patient-generated health data in the regulatory setting.

- **Table 4. Opportunities for Patient-Generated Health Data to Augment the Healthcare System**

Stakeholder Group	Value
Patients and caregivers	greater opportunity to engage in shared decision making, ability to contribute to research with less inconvenience
Clinicians	ability to assess patient concerns, behaviors, adherence to treatment plans across disease states and device use and outside the clinic environment
Researchers & Industry	access to large pre-existing data pools, potential ability to identify future clinical trial participants and/or supplement clinical trial outcome data collection
Regulators & Policymakers	ability to glean post-market data informative for safety monitoring
Policymakers	ability to streamline and standardize information needed to inform policy
Developers & Standards Bodies	determine standards to enable collection and analysis across platforms, data interoperability greater integration into healthcare systems
Payors	Access to additional data sources for coverage and value analysis along with other decision-making purposes

- *Adapted from Office of the National Coordinator for Health Information Technology.*⁹²
-

Appropriate approaches to incorporating patient-generated health data throughout the healthcare ecosystem may provide more holistic and rich data to better inform patient and provider decision-making, enhance postmarket monitoring for regulators and industry, improve our

⁹² Office of the National Coordinator for Health Information Technology. White Paper: Conceptualizing a data infrastructure for the capture, use, and sharing of patient-generated health data in care delivery and research through 2024. January 2018. Contract no. HHSP2332015000931, order no. HHSP23337001T.

understanding of diseases and outcomes important to patients, and support downstream decision-makers such as payors and health systems (**Table 4**).