

PERSPECTIVE OPEN



Modernizing and designing evaluation frameworks for connected sensor technologies in medicine

Andrea Coravos^{1,2,3,4}, Megan Doerr⁵, Jennifer Goldsack³, Christine Manta^{1,3}, Mark Shervey¹, Beau Woods^{4,6,7} and William A. Wood^{3,8}✉

This manuscript is focused on the use of connected sensor technologies, including wearables and other biosensors, for a wide range of health services, such as collecting digital endpoints in clinical trials and remotely monitoring patients in clinical care. The adoption of these technologies poses five risks that currently exceed our abilities to evaluate and secure these products: (1) validation, (2) security practices, (3) data rights and governance, (4) utility and usability; and (5) economic feasibility. In this manuscript we conduct a landscape analysis of emerging evaluation frameworks developed to better manage these risks, broadly in digital health. We then propose a framework specifically for connected sensor technologies. We provide a pragmatic guide for how to put this evaluation framework into practice, taking lessons from concepts in drug and nutrition labels to craft a connected sensor technology label.

npj Digital Medicine (2020)3:37; <https://doi.org/10.1038/s41746-020-0237-3>

INTRODUCTION

Over the past decade, the adoption of digital technologies in medicine—from electronic health records to wearable sensors—has occurred faster than the healthcare community's ability to evaluate and secure these products^{1–3}. The fundamental goal of any biomedical product evaluation is to assure that, in the intended context of use, the benefits of deploying the technology outweigh the potential risks to the participant/patient and the organization. As new technologies enter medicine and biomedical research, manufacturers, regulators, clinicians, and patients are relying upon existing regulatory and evaluation frameworks. For example, traditionally consumer-facing companies, such as Apple, are now approaching the US Food and Drug Administration (FDA) as they develop products for clinical settings⁴. However, the risks of deploying these new technologies are inadequately understood and are not protected by what are quickly becoming legacy evaluation frameworks.

Compared to legacy biomedical products, digital technologies have features that change the benefit-risk calculation; therefore, we must adapt evaluation frameworks. In this manuscript, we focus specifically on connected biometric monitoring technologies, which we will refer to as “connected sensor technologies”. Connected sensor technologies are digital medicine products that perform algorithmic processing of data captured by mobile sensors to generate measures of behavioral and/or physiological function. Examples of connected sensor technologies include smartwatches that measure activity, connected monitors that sit on top of mattresses to measure sleep, wireless arm cuffs that measure blood pressure, and microphones that capture vocal biomarkers signaling changes in brain health^{5–7}. Notably, we intentionally use the phrase “connected sensor technology” or “connected product,” and we avoid the phrase “connected device” because “device” is an FDA term of art that refers specifically to cleared medical devices⁸.

In this paper, we first briefly discuss the benefits of connected sensor technologies and dive deeper into the risks these

technologies present. Next, we outline the frameworks that are emerging across the industry to evaluate digital health, highlighting their strengths and shortcomings with reference to the connected sensor technologies industry. Finally, building on these emerging frameworks as a guide, we outline a practical guide for evaluating fit-for-purpose connected products across biomedical research and clinical care.

FEATURES OF CONNECTED SENSOR TECHNOLOGIES AND THEIR BENEFITS AND RISKS

Connected products are being rapidly adopted, with the number of wearables worldwide estimated to increase from 325 million in 2016 to 929 million by 2021⁹. The proliferation of wearables, ingestibles, and other connected sensors is making it easier than ever to collect high-quality behavioral and physiological data outside of the clinic⁶. Remotely collected data allow clinicians to discover sights that are more reflective of patients' day-to-day experiences.

For drug developers, connected sensor technologies can improve efficacy⁶, increase inclusivity¹⁰, and lower the costs of conducting clinical trials¹¹. For clinicians, these products can capture insights that are more reflective of patients' day-to-day experiences, potentially resulting in major improvements in care delivery.

To capture these potential benefits, risk-benefit analyses are essential to ensure accurate measurement and patient safety in study protocols and clinical care. In the next section we highlight five dimensions that carry risks posed by connected sensor technologies: (1) validation, (2) security practices, (3) data rights and governance, (4) utility and usability; and (5) economic feasibility.

Verification and validation

To determine the appropriateness of using a particular product, a short-cut question that often arises is whether the product is “validated” (e.g., “is this wearable clinically validated?”). Validation carries widely different meanings for different stakeholders. For instance, the pharmaceutical industry may use “validation” as a

¹Elektra Labs, Boston, MA, USA. ²Harvard-MIT Center for Regulatory Science, Boston, MA, USA. ³Digital Medicine Society (DiMe), Boston, MA, USA. ⁴Biohacking Village, Washington, DC, USA. ⁵Sage Bionetworks, Seattle, WA, USA. ⁶Am The Cavalry, Washington, DC, USA. ⁷Atlantic Council, Washington, DC, USA. ⁸University of North Carolina Lineberger Comprehensive Cancer Center, Chapel Hill, NC, USA. ✉email: william_wood@med.unc.edu

substitute for “GxP”, which is a generalized abbreviation for “good practice” quality guidelines and regulations. Examples of GxP are Good Clinical Practice (GCP), Good Manufacturing Practice (GMP), and Good Laboratory Practice (GLP). GxP compliance is a set of quality system of management controls which have been developed over the years with and for stakeholders (e.g., clinical trialists, manufacturers, laboratories), and codified into current regulatory regimes.

For others, validation may be more akin to “analytical or clinical validation,” referring to the quality of the measurement coming from the sensor and algorithms that compose the connected sensor technology. Others may also bundle “validation” with “verification and validation” (V&V), quality management procedures that ensure that the system or product meets specifications and that it fulfills its intended purpose (e.g., “software V&V”). Over time, evaluation frameworks that are developed for connected sensor technologies will likely be codified into revised GxP and related quality management systems; however, to develop good practices, the underlying principles must first be established.

To account for the unique hardware, software, and algorithmic properties of connected biometric monitoring technologies (BioMeTs), we recommend the three-stage process of verification, analytical validation and clinical validation (V3) proposed by Goldsack, Coravos, Bakker et al.¹². In this framework:

- **Verification** evaluates and demonstrates the performance of a sensor technology within a BioMeT, and the sample-level data it generates, against a pre-specified set of criteria.
- **Analytical validation** evaluates the performance of the algorithm, and the ability of this component of the BioMeT to measure, detect, or predict physiological or behavioral metrics.
- **Clinical validation** evaluates whether a BioMeT acceptably identifies, measures, or predicts a meaningful clinical, biological, physical, functional state, or experience in the specified population and context of use.

A strong V3 process serves as the foundational evidence base around the accuracy, reliability, and appropriateness of the data and results from connected sensor technologies. Nonetheless, conducting a successful V3 process is challenging for a number of reasons. First, most sensor-based products are comprised of a modular stack of hardware and software components, from sensors to signal processing to algorithms⁶. Each component may be built by a different company, each of which contribute to the product’s overall V3 results. Second, a change earlier in the data supply chain (e.g., at the signal-processing algorithm in the sensor) may alter the data inputs for an algorithm high-up the chain, which may result in an entirely new V3 valuation¹². Put another way, it is challenging to evaluate a connected sensor technology’s data supply chain, the data flow and data provenance for information generated from hardware, sensors, software, and algorithms. Indeed, it requires modifications to the V&V process for wet-lab tests or clinical outcome assessments (COAs) like electronic patient reported outcomes (ePROs).

Security

By definition, connected sensor technologies transfer data over the internet, which introduces immediate risks when deploying these products, because an actor could attack and access the product remotely and often in near-real time. This second dimension on Security risk evaluates unauthorized uses of data and results; the following section on Data Rights and Governance evaluates authorized uses of data and results. Cybersecurity involves protecting internet-connected systems, data, and networks from unauthorized access and attacks, including human error (e.g., the loss of a company’s unencrypted laptop). Notably, some data and system access may be authorized (or perhaps “not

forbidden”), though unwelcome or undisclosed to the patient or other stakeholders. This type of access will also be covered in the next section. Although the security of a system cannot be guaranteed, quality design and execution can decrease the risk of harm from code flaws, configuration weaknesses, or other issues. A product’s security risk will need to be continuously re-assessed as new technologies and attack methods become available (e.g., advances in quantum technologies and corresponding quantum-resistant encryption standards).

Data rights and governance

When we consider data rights, we prefer to refer to governance rather than privacy, because we believe it’s more important to empower individuals to choose how to share their data—their rights and governance—rather than defaulting to privacy (e.g., a patient with a rare disease may want more freedom rather than barriers to share her data and results with relevant parties).

Over the past year, many popular tech companies have come under greater scrutiny for how they choose to share data with third-parties. For instance, the Cambridge Analytica incident with Facebook was not an unauthorized use or attack on the Facebook network (e.g., it was not a security incident). Aggregation of data in the ways utilized by Cambridge Analytica was part of Facebook’s feature set, though many have argued this feature was not thoroughly disclosed to all parties. Examples of widespread data sharing with inadequate disclosure is also seen in health tech products. Huckvale et al found in a cross-sectional study of 36 top-ranked apps for smoking cessation and depression in public app stores, “29 transmitted data to services provided by Facebook or Google, but only 12 accurately disclosed this in a privacy policy”¹³.

It is important to note that the regulatory environment is far from established when it comes to governing “digital specimens” (e.g., the data generated from connected sensor technologies). With respect to regulation, the FDA has oversight for digital specimen-collecting technologies, like wearables, when they are classified as a medical device. However, due to the narrow definition of device and the revisions with the 21st Century Cures Act, many connected sensor technologies fall outside of the FDA’s purview¹⁴. These narrow frames leave oversight of connected sensor technology functionality and health claims primarily to the Federal Trade Commission, which policies unfair and deceptive trade practices, including enforcing rules against false or misleading advertising¹⁵. In the United States, other agencies like National Institute of Standards and Technology (NIST), Federal Communications Commission (FCC) and Office of the National Coordinator for Health Information Technology (ONC) may each have oversight of components of connected sensor technologies, but no regulator has full responsibility for digital specimens. Given this ambiguous regulatory landscape, end-user license agreements (EULAs) for sensors with downloadable software (e.g., app), terms of service (ToS) for sensors themselves, and privacy policies (PP) have become the de-facto agreements that to retain rights in software and to create rights to monitor, aggregate, and share users’ digital biospecimens (see Box 1)¹⁵.

Box 1. Data rights disclosures: EULAs, ToS and PP intended use cases

Privacy policies (PP) disclose the terms for collection and use of the app/website user’s personal information.

Terms of service (ToS) disclose the rules and requirements of website and/or app use, for example, copyright, allowed uses, and the definition of abusive use.

End-user license agreements (EULAs) are a form of intellectual property licensing that tell people who have purchased software if/how many times they can copy the software and how they can or cannot use those copies.

should be evaluated at the product-level, organization-level or both. Historically evaluations focused on the product, not on the organization or manufacturer (e.g., FDA judges the quality of a specific drug, and not Pfizer overall). With connected sensor technologies, the broader system needs to be taken into context, because a hardware or software change in one component (e.g., an update to the signal processing algorithm) can impact the system overall. Additionally, because software updates can occur frequently—in some instances multiple times per day—regulating these changes can require a new framework to manage risk. In software development the culture and processes at the organizational-level impact multiple products at once. Organizational-level views can also be better than product-level views when considering data rights and governance, as privacy policies and EULAs are often structured at the organizational-level rather than individual product-level. Therefore, some evaluation frameworks have shifted to consider this organizational view such as the FDA Precertification (Pre-Cert) Program, which evaluates the quality of the organization overall and then provides a “streamlined” review pathway for pre-certified organizations²¹.

2. **Research versus clinical care settings.** Second, the same connected sensor technology may be used in either a research setting (e.g., to collect digital endpoint data to support a drug application) or in a clinical setting (e.g., to remotely monitor a patient’s quality of life). An optimal evaluation framework should likely have the same base evaluation for the quality of the connected sensor, and afterwards, adapt the evaluation for different requirements in a clinical and research setting, respectively. This issue is exacerbated when regulatory requirements vary across connected technologies (i.e., some are regulated as medical “devices”, and others are not). The distinction between which digital health technologies are regulated and which are not is still evolving²², and a gray area can be dangerous for public health. Take for instance the controversies surrounding JUUL, a technology that was not well understood when first deployed into the market and now is facing greater regulatory scrutiny²³. In a fast-paced technology

world, it’s not only the responsibility of regulators to develop new evaluation models. With additional forethought, we do not have to wait for public crises to enact thoughtful oversight. It is the responsibility of all the parties involved to work towards understanding safe and effective products.

3. **Evaluation scope: software versus hardware review.** Given the modularity of connected sensors (e.g., a hardware component, sensors, signal processing algorithms, and apps to display the data), some emerging frameworks look at the whole set of components and others conduct a software-only review. Additionally, this modularity split is also showing up in a regulatory context as the FDA has introduced the concept of a “software as a medical device” (SaMD), which is “software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device”²⁴.
4. **Comprehensiveness.** The final theme is that some of the emerging frameworks review all five of the risks posed by connected sensor technologies (validation, security, data rights and governance, utility and usability, and economic feasibility), and others only look at a subset. In the following section, we build off the lessons from the emerging frameworks and propose a pragmatic evaluation criteria to consider when deploying connected sensor technologies in research or clinical care.

BUILDING AN EVALUATION FRAMEWORK FOR CONNECTED SENSOR TECHNOLOGIES

Building on the existing frameworks, we propose a working evaluation framework for connected sensor technologies that reflects the five types of risks identified above (Fig. 2). We constructed this framework using the following principles:

1. **Evaluation criteria should be objective, observable, and verifiable** (see Box 2). Objective criteria are clearly and reliably measurable. Observable criteria can be checked independently, without special or privileged access. Verifiable criteria can be demonstrated or refuted with empirical data.

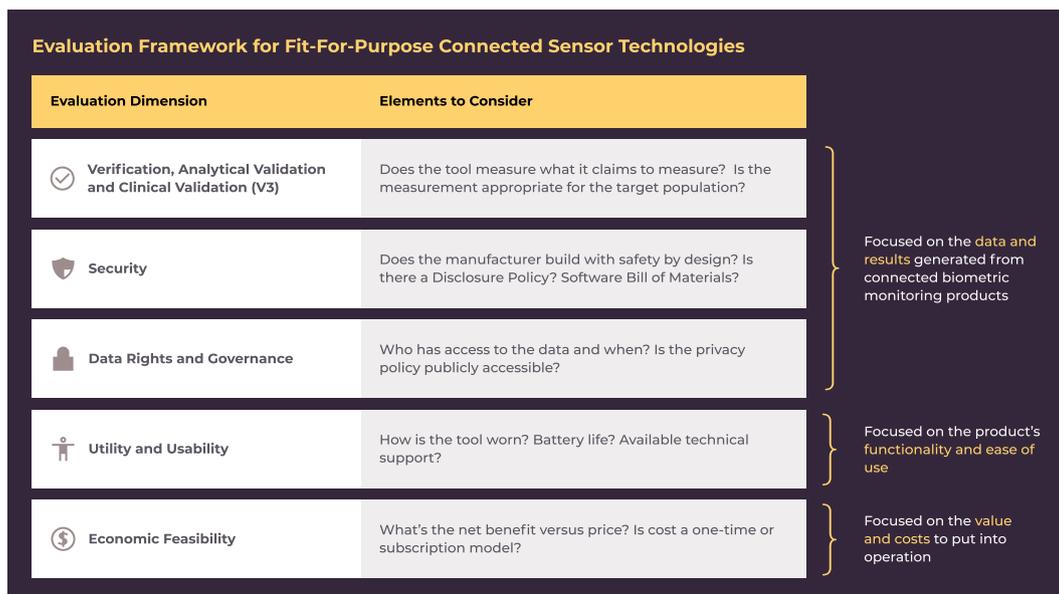


Fig. 2 Proposed Evaluation Framework for Fit-For-Purpose Connected Sensor Technologies. The proposed framework describes the key dimensions to be considered when evaluating connected sensor technologies. Dimensions can be grouped into broader categories of data and results; functionality and ease of use; and value and costs.

Box 2. Objective, observable and verifiable evaluation criteria

Objective That it can be agreed whether the capability is in place - often this leads to binary proofs - it either is or is not, there are no degrees.

Observable That an independent person can know whether the capability exists, without special or privileged access. This characteristic gives public scrutability to the capability.

Verifiable That a capability can be demonstrated or refuted with empirical data. Source: Adapted from comments on NIST considerations⁵¹.

- Evaluations need context.** “What is the best drug?” or “What is the best food?” are meaningless questions without additional context (e.g., does the person need less sugar in her diet? Or more protein?). Similarly, “what is the best heart-rate monitor?” is an empty question without a clearly articulated context.
- Evaluations should be multidimensional** (e.g., avoid a single metric “score”). While scoring a food by a single metric such as total calories can be helpful, calorie count in itself is not a way to construct an overall healthy diet. Similarly, we argue that compressing an evaluation of a Fitbit versus an Apple Watch into a single overall score lacks meaningful nuance.
- Evaluation components can have required minimum thresholds and optional features that enhance the desirability of the product.** Required thresholds of each component may depend on the risk level and context of use.

We propose a systematic and standardized approach to evaluate whether connected biometric monitoring technologies are “fit-for-purpose” across five dimensions:

- Verification, analytical validation, and clinical validation (V3)¹²;
- Security practices
- Data rights and governance
- Utility and usability; and
- Economic feasibility

The first three dimensions evaluate the data and subsequent results generated by connected biometric monitoring products. The fourth dimension, Utility and Usability, evaluates the ease of implementation and adoption of the product, and the last dimension, Economic Feasibility, evaluates the economic feasibility of adoption.

Notably, excellence in one dimension does not necessarily imply excellence in another. Indeed, significant deficiencies in any one dimension may lead to problems when using a connected sensor technology in research or practice. Thus, we propose a framework to simplify the evaluation process of connected sensor technologies for their intended uses.

KEY EVALUATION CRITERIA AND METRICS

Verification, Analytical Validation, and Clinical Validation (V3)

The following documentation is necessary to determine net benefit using V3 principles:

- Verification** can look like performance specifications for integrated hardware, overviews of software system tests, or output data specifications¹². Often this information is on the manufacturers’ website and not as a peer-reviewed journal article.
- Analytical validation** may look like studies that follow Good Clinical Practice (GCP) requirements, and could show up as a regulatory submission (e.g., 510k), white paper, or peer-reviewed journal article.

- Clinical validation** may look like a clinical study report (CSR), regulatory submission, journal paper or published conference proceedings¹².

Of all forms of documentation, those that make available complete data sets for external review should be weighted most heavily, particularly when machine learning algorithms are used²⁵. Emerging standards exist for the assembly and documentation of datasets²⁶ and the visualization of underlying data quality²⁷. There is also a need for a living database of published studies with these types of data, as efficiently gathering data across the medical literature in each of these areas can be challenging for end-users. The Clinical Trials Transformation Initiative’s (CTTI’s) Interactive Database of Feasibility Studies for Mobile Clinical Trials is a useful and welcome start²⁸. The Digital Medicine Society (DiMe)’s crowdsourced library is another useful resource that lists connected sensor technologies that collected data used to derive digital endpoints for industry-sponsored studies of new medical products or new applications of existing medical products^{29,30}. More work needs to be done to ensure reproducibility of machine learning algorithms in healthcare³¹, and whether or not publishing either the algorithms or datasets for an independent external review³² would be a constructive way to increase reproducibility and decrease bias¹.

When evaluating V3 for a connected sensor technology, the product should have documentation for each of the three stages. The documentation should align with the intended patient population and context of use. Measures that have been clinically validated in one group of patients cannot be assumed to be valid in another group in which patient or environment characteristics may affect measure performance (e.g., gait in Parkinson’s population may be evaluated differently than in a population with Multiple Sclerosis). As a desired threshold, the measure should be evaluated and published in multiple populations and data sets. Regulatory decisions (e.g., from FDA or EMA) can be distracting or misleading when evaluating a product. For instance, a connected sensor technology’s 510(k) FDA clearance as a device for clinical practice has no impact on whether it would be a suitable product for a drug clinical trial. All the decision indicates is that an external body has reviewed the manufacturer’s marketing claims.

We recognize that many connected sensor technologies may not meet the minimum threshold for verification and validation. For such technologies, we recommend identifying where along the data supply chain the product is missing documentation, and then conducting and/or soliciting research to complete the V3 process.

Sample threshold criteria for the V3 process

- Minimum threshold or pass/fail:** a proposal or initial white paper for how the product plans to run its V3 process (e.g., what’s been completed and where are the gaps).
- Desired threshold:** white paper or equivalent data describing elements of the V3 process; depending on the intended use (unless an analytical or clinical validation study is planned), published V3 data within the context of use.
- High quality:** well-documented V3 specifications; analytic and clinical validation data published in multiple populations and data sets.

Security

The next two dimensions of the evaluation framework, Cybersecurity and Data Rights and Governance, address unauthorized and authorized access to data and systems. Although all connected sensor technologies will eventually fail, not all failures have to cause harm. In 2016, a group of security researchers published a

Box 3. Cybersecurity considerations

Evaluation criteria considerations

Does the company have a Coordinated Vulnerability Disclosure (CVD) Policy and what does it contain? A Coordinated Disclosure (CVD) Policy outlines the process, rules of engagement, and expectations for security researchers (or others) who find and report potential security issues, in good faith⁵². When financial incentives are offered to individuals for reporting issues, this may be called a bug bounty program⁵³.

Does the organization publish its security support lifetime and issue secure, prompt, and agile software updates once security issues are discovered? Software updates can correct and mitigate potential security and safety issues much more effectively and efficiently than hardware replacement. Transparency and clear communication increases update participation rates, protecting more patients. Sharing security documentation, such as minimum supported lifetimes and known Common Vulnerabilities and Exposures (CVEs) allows users of connected sensor technologies to make more informed decisions about purchasing and use.

Does the organization track and share a Software Bill of Materials (SBOM)? An SBOM list of components in a piece of software is analogous to a list of ingredients on food packaging⁵⁴. Given that nearly all software products today are comprised of third-party components, an SBOM makes it easier to identify and address potential security risks from software component vulnerabilities.

set of principles for creating more resilient and secure systems. These principles, encapsulated as the Hippocratic Oath For Connected Medical Devices, include designing the product with cyber safety in mind, ensuring third-party collaboration, capturing evidence to improve safety investigations, resilience and containment, and an emphasis on cyber safety updates and patching³³. All of these five principles were incorporated into the FDA's pre- and post-market guidances for cybersecurity of connected medical devices^{34,35}.

When evaluating a product's cybersecurity considerations, it is critical to ensure the assessment is risk-based and the context of use is considered. There are many organizations that will conduct customized risk-based security assessments. Or, an organization may earn a certification, like HITRUST³⁶.

When considering the cybersecurity metrics, we propose a mix of passed assessments and certifications as well as process-related objective metrics. For technologies connected to the internet, the question isn't "if" the product will be compromised, but "when." Because of this, we focus on process metrics around identifying vulnerabilities and successfully patching them³³:

Sample threshold criteria for the cybersecurity considerations (see Box 3)

- **Minimum threshold or pass/fail:** the connected sensor technology manufacturer has a publicly accessible Coordinated Vulnerability Disclosure (CVD) policy and publishes details about updates, including security vulnerabilities addressed.
- **Desired threshold:** the connected sensor technology manufacturer maintains a software bill of materials (SBOM).
- **High quality:** the organization keeps track of patch uptake rate and makes its SBOM available to customers and regulators.

Data rights and governance

The third dimension addresses issues related to authorized (or "not unauthorized") access to data and systems. Given the ambiguity in the regulatory landscape and the lack of clear regulatory oversight of this risk area, we focus on the individual product policies. As a result, the metrics we recommend evaluating for connected sensor technologies are related to EULAs/ToS and PPs. Sage Bionetworks, a nonprofit organization that promotes open science and patient engagement in the research process, has pioneered a number of open toolkits for informed consent and privacy best practices that we have drawn from in this example^{37,38}.

Box 4. Data rights and governance considerations

Evaluation criteria considerations

Does the device have EULA/ToS and PP? EULAs are traditionally associated with software that is downloaded by the user; ToS are associated with devices that do not require additional download of software. All sensor/software/devices should additionally have privacy policies.

Are these policy documents comprehensive? EULA and ToS: are the limits of use and liability* disclosed? Are termination of access, transfer of holdings, and the policy for changes/updates to terms described?

PP: Is what digital specimens will be collected, how they will be used (including by third parties), and what recourse the user has regarding their digital specimen described?

Are these documents physically accessible? Documents should be publicly accessible online.

Is the information contained in them comprehensible by broad audiences? Is documentation written at or below the 5th grade reading level or are plain language summaries provided? Is documentation 508 compliant⁵⁵?

*It is important that researchers understand the limits of liability disclosed by the device/sensor manufacturer and ensure that these do not constitute exculpatory language per FDA/OHRP guidance⁵⁶.

Sample threshold criteria for data rights and governance considerations (see Box 4)

- **Minimum threshold or pass/fail:** EULA/ToS and PP exist, are comprehensive, and are publically accessible online.
- **Desired threshold:** Documents are comprehensible by broad audiences, as well as accessibility compliance (e.g., 508 Compliant) for people with disabilities³⁹.
- **High quality:**
 - EULA/ToS does not contain exculpatory language.
 - Users can opt-in or opt-out of third party transfer/use of their data.
 - The rights of users are not changed even in the case of a change in ownership of the connected technology/sensor manufacturer.

Utility and usability considerations at the participant, site, and software integration levels

In this dimension, we focus on utility and usability from three perspectives: the patient, the site, and the software-integration level.

For the person wearing the connected sensor technology there are many practical considerations. This includes the form factor (e.g., how is a wearable sensor worn, how does it feel, what does it look like), battery life, water resistance, and materials used in its construction (e.g., does it cause a rash or irritation?) Availability of technical support, features of associated software that display sensor-generated data, and how engaging the product is can affect the overall participant experience. If the product is too difficult for the patient to use (i.e., low patient usability), data may not be suitable for analysis because of high rates of data missingness or participant dropout.

For the clinical site or practice that must integrate the connected sensor technology into a trial or routine care, workflow considerations are important. For example, sites should consider statistical analysis and how the data from the connected sensor technologies are aggregated and made available for reporting (e.g., administrative dashboards, or EHR displays if applicable). Workflow issues might include how participants are trained to use the connected sensor technology; what level of site monitoring or support is needed as the data are obtained; what clinical actions (if any) are needed in response to the data and how thresholds or alerts to guide these actions are designed; and what type of site-level technical support is needed to ensure that the technology functions appropriately. If not adequately considered, the site may be unable to obtain health measurements or deliver interventions

as desired, reducing the effectiveness and impact of the technology⁴⁰.

By “software-integration perspective,” we refer to considerations related to the integration of a connected sensor technology with software necessary for the transfer, visualization, and analysis of the connected data. This may include the use of standard data models (e.g., FHIR or OMOP Common Data Model) and open application programming interfaces (APIs) to facilitate ingestion and analysis of digital specimens.

Finally, usefulness under ideal circumstances is different for particular sub-populations. For example, a connected sensor technology may perform well from a participant and site perspective when the population is a group of high-functioning, healthy volunteers and the environment is a highly equipped and controlled digital laboratory within a sponsor organization. The same technology may perform less well when the population is a group of older, multimorbid patients with life-limiting illnesses and low physical function, and when the environment is a busy medical center participating in a clinical study, or a “real world” direct to participant approach as part of a decentralized trial¹⁰.

When evaluating a product, we recommend that the connected sensor technology have data on its utility and usability in healthy volunteers, published in a related white paper or equivalent. As a desired threshold, we look for peer-reviewed evidence of the technology in healthy volunteers and patients in trials or practice.

Additional value is added if published data match the intended use of the technology, in the intended target population, for the intended duration of data collection, with low rates of data missingness, and with ease of data aggregation and analysis. Further value is added based on the level of published or unpublished detail available for each of the above that can guide those applying the connected sensor technology to the intended population. Lastly, an indirect criteria for usability of a connected sensor technology could be the extent to which it is currently used in research and practice. Thus, further value is added based on the extent to which it is being used in clinical trials, grant applications, and real-world settings.

Sample threshold criteria for utility and usability considerations

- **Minimum threshold or pass/fail:** white paper or equivalent data about the connected sensor technology in healthy volunteers.
- **Desired threshold:** peer-reviewed publication about the use of the connected sensor technology in patients.
- **High quality:** matching of feasibility data to intended context of use; high adherence/low data missingness; ease of data aggregation; detail to provide guidance; past and current use in research and practice.

Economic feasibility

When considering economic feasibility, it is important to calculate both the advantages and costs when using the product. The advantages often will come from improved quality and outcomes from using the product. The costs will depend on the product’s pricing model and expected changes over time (e.g., to cover long-term data storage, analysis, and security updates). When conducting an economic feasibility analysis, the feasibility perspectives from an insurance company, clinician and patient would likely differ, and therefore it is also important to consider all of the potential stakeholders and the values of those participating in the system. One stakeholder’s advantage may be another stakeholder’s cost. Economic feasibility calculations for a system are often complex as many of the advantages do not always have financial metrics associated and the lifetime costs can be difficult to estimate.

OPERATIONALIZING AND DEPLOYING A CONNECTED SENSOR TECHNOLOGY EVALUATION FRAMEWORK

An evaluation framework is an intellectual exercise until it is put into practice. In this section, we outline several pathways for operationalizing the framework we describe in this manuscript. We anticipate roles for regulators such as FDA, standards bodies like Xcertia and IEEE, and those who develop GxP processes for connected sensor technologies. We also need to have better communication infrastructure to make information more accessible to the stakeholders who evaluate and deploy these technologies, the patients, doctors, software engineers and so on.

The technique of labeling has been deployed as a communication tool for complex products in healthcare. Prescription drug labels communicate essential, science-based information to healthcare professionals. Nutrition labels communicate essential, easy-to-read information for consumers managing their diets. Labels are critical pieces of infrastructure that provide transparency and clarity for life-critical products. Efforts to improve Internet of Things (IoT) security transparency has the attention of the multiple government entities. In the United States Congress, the Cyber Shield Act has been proposed, which includes labeling elements to help consumers review the security features of devices they’re buying⁴¹; the National Telecommunications and Information Administration (NTIA) is exploring transparency tools such as product packaging labels and consumer-facing websites⁴²; the United Kingdom government has developed labeling proposals and researched consumer perceptions⁴³. Outside of governments, several more security and data-rights focused labels have been proposed such as Privacy and Security Nutrition Label for Smart Devices⁴⁴, Symantec’s IoT label⁴⁵, the Data Transparency Label⁴⁶, and the Patient Privacy Rights Information Governance Label⁴⁷, as well as efforts to make a cybersecurity “Energy Star”⁴⁸. These types of labels create a visual representation for metrics like the type of data collected, retention period, update capabilities (e.g., automatic), and authentication methods, making it easier for less technical consumers to understand important features and gaps in a product.

Mock visualizations

We propose that a connected sensor technology label could be a useful piece of infrastructure for an evaluation framework, making it easier for decision-makers to understand critical aspects of the technology. The label format could demonstrate which connected sensor technologies are available that contain the proposed measurements, and whether each of five dimensions (V3, security practices, data rights and governance, utility and usability, and economic feasibility) meet minimal (pass/fail) or desired quality thresholds. It also displays additional quality, if any, for each dimension. A prototype is provided below, one that we hope will be the starting point for a broader industry discussion and ultimately, widespread implementation (Figs 3 and 4).

CONCLUSIONS

The adoption of connected sensor technologies is not slowing down, and so it is of utmost importance to provide our community with improved and practical evaluation toolkits. Our hope is that the use of these toolkits will slow the adoption of substandard connected sensor technologies that lack evidence and/or harm individuals while accelerating the adoption of safe and effective ones. To achieve this vision, we need to define shared ontologies that better reflect the unique considerations of products in the digital era. Evaluation frameworks will need to be easy to understand and implement in order to shift the adoption curve at scale. Frameworks will also need to provide risk-benefit evaluations in a timely manner, as software organizations have the ability to make changes to connected products faster compared to

Connected Sensor Facts	
Make (e.g., Fitbit, Apple, Actigraph)	
Model (e.g., Charge 3, Apple Watch, wGT3X-BT)	
Verification, Analytical Validation, and Clinical Validation	
Peer Reviewed Papers	87
ClinicalTrials.gov Studies	36
...	...
Security Practices	
Coordinated Vulnerability Disclosure Policy	Yes
Software Bill of Materials (SBOM)	Yes
...	...
Data Rights and Governance	
Privacy Policy	Yes
Publicly Accessible Online	Yes
Last Update	2017
...	...
Utility and Usability	
Battery Life (hours)	72
Waterproof Rating	IP67
...	...
Price	
Physical Product (e.g., hardware)	\$99
Data Analysis and Storage (per month)	\$35

Last Updated: January 12, 2020
Suggest an update?
<https://github.com/acoravos/connected-sensor-facts>

Fig. 3 Connected sensor facts. This is an example of a nutrition label-type visualization that includes the key dimensions of the evaluative framework for connected sensor technologies, and metrics related to facts within each dimension. This approach permits a user to have a concise but comprehensive picture of a sensor's appropriateness and fitness for use.

other medical products (e.g., ability to ship software updates multiple times per day versus releasing a new drug or medical device every few years). Additional work should be done to ensure that our proposed framework appropriately meets the needs of stakeholders throughout the digital medicine ecosystem, and our proposed framework should be continually reevaluated to best capture benefits and risks when adopting evolving technologies. Communities, like the Digital Medicine Society (DiMe), will be essential to develop a professional home to serve software engineers, data scientists, clinicians, and other experts working in this field as new frameworks are developed⁴⁹.

There are many ways to operationalize an improved evaluation framework for connected sensor technologies, and the optimal result will likely have a mix of actions ranging from regulatory, standards bodies, and communication tools like accessible labeling. The underlying principles we outlined for connected sensor technologies can be adapted for related digital medicine technologies⁵⁰ like ePROs or interventional products like digital therapeutics (DTx)¹⁴, although neither are the immediate focus of this work. Our hope is that the community will build on the five

Comparison Table		 Armband by Company A	 Smartwatch by Company B
Verification, Analytical Validation and Clinical Validation (V3)			
Regulatory Decisions		3	1
Peer-reviewed papers		15	3
Trials registered in ClinicalTrials.gov		15	3
Security			
Responsible Disclosure Policy		Available	Available
Bug Bounty Program		Bugcrowd	Hackerone
Software Bill of Materials		Not found	Not found
List of Known Vulnerabilities		Not found	Available
Data Rights and Governance			
Publicly Available		End User License Agreement Privacy Policy Terms of Service	End User License Agreement Privacy Policy Terms of Service
Comprehensibility of Privacy Policy		5th grade or below	College and up
Utility and Usability			
MEASUREMENTS			
Heart Rate		Available with evidence	Available
Activity		Available with evidence	Available with evidence
Sleep		Available with evidence	Available
Temperature		Available with evidence	Not available
FORM FACTOR			
		Strap/Band	Watch
WEAR LOCATION			
		 Upper Arm	 Wrist
SPECIFICATIONS			
		Accelerometer Altimeter Gyroscope Photoplethysmography (PPG) Thermometer	Accelerometer Ambient Light Meter Electrocardiogram (with chest strap)
Battery Life		1-2 days	5-7 days
Waterproof		Water resistant to IP22 standards	Water resistant to IP57 standards
Connectivity		Bluetooth Wired	Bluetooth Cellular WiFi
Compatible OS		iOS	iOS Android
API		Available	Not Found
Product Manual		Download	Download
Economic Feasibility			
Cost		\$2800.00	\$129.99

Fig. 4 Connected sensor comparison table. This is an example of how sensors might be compared in a side by side visualization that incorporates dimensions from the proposed evaluative framework. In this illustrative visualization, additional detail and metrics are provided for the utility and usability dimension, in order to facilitate a rapid evaluation for appropriateness within an intended population.

dimensions described here developing more robust thresholds so more technologies are worthy of the trust society places in them.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

No datasets were generated or analysed during the current study.

Received: 21 October 2019; Accepted: 11 February 2020;

Published online: 13 March 2020

REFERENCES

- Hailu, R. Fitbits and other wearables may not accurately track heart rates in people of color. <https://www.statnews.com/2019/07/24/fitbit-accuracy-dark-skin/> (2019).
- Yazdi, D. The Apple Watch 4 is an iffy atrial fibrillation detector in those under age 55. <https://www.statnews.com/2019/01/08/apple-watch-iffy-atrial-fibrillation-detector/> (2019).
- Brandom, R. UK hospitals hit with massive ransomware attack. <https://www.theverge.com/2017/5/12/15630354/nhs-hospitals-ransomware-hack-wannacry-bitcoin> (2017).
- Apple. ECG app and irregular heart rhythm notification available today on Apple Watch. <https://www.apple.com/newsroom/2018/12/ecg-app-and-irregular-heart-rhythm-notification-available-today-on-apple-watch/> (2018).
- Sim, I. Mobile devices and health. *N. Engl. J. Med.* **381**, 956–968. <https://doi.org/10.1056/NEJMra1806949> (2019).
- Coravos, A., Khozin, S. & Mandl, K. D. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digital Med.* **2**, 14 (2019).
- Wang, T., Azad, T. & Rajan, R. The emerging influence of digital biomarkers on healthcare. <https://rockhealth.com/reports/the-emerging-influence-of-digital-biomarkers-on-healthcare/> (n.d.).
- US Department of Health and Human Services, Food and Drug Administration. Device—not a device. <https://www.fda.gov/medical-devices/classify-your-medical-device/device-not-device> (2018).
- Liu, S. Number of connected wearable devices worldwide from 2016 to 2022 (in millions). <https://www.statista.com/statistics/487291/global-connected-wearable-devices/> (2019).
- Khozin, S. & Coravos, A. Decentralized trials in the age of real-world evidence and inclusivity in clinical investigations. *Clin. Pharmacol. Therapeutics* **106**, 25–27 (2019).
- Dodge, H. H. et al. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. <https://doi.org/10.1371/journal.pone.0138095> (2015).
- Goldsack, J. C. et al. Verification, Analytical Validation, and Clinical Validation (V3): The Foundation of Determining Fit-for-Purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digital Medicine, In Press. “IN SAME SPECIAL ISSUE.” (2020).
- Huckvale, K., Torous, J. & Larsen, M. E. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Netw. Open* **2**, e192542. <https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2730782> (2019).
- Coravos, A. et al. Digital medicine: a primer on measurement. *Digital Biomark.* **3**, 31–71. <https://www.karger.com/Article/FullText/500413> (2019).
- Matwyshyn, A. M. The ‘internet of bodies’ is here. *Are courts and regulators ready?* <https://www.wsj.com/articles/the-internet-of-bodies-is-here-are-courts-and-regulators-ready-1542039566> (2018).
- FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. *Food and Drug Administration (US)*. <https://www.ncbi.nlm.nih.gov/books/NBK326791> (2016).
- Nielsen, J. *Usability 101: Introduction to usability*. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/> (2012).
- Elezea. Utility is more important than usability. <https://elezea.com/2016/01/utility-is-more-important-than-usability/> (2016).
- Cambridge. Cambridge Business English Dictionary. (Cambridge University Press, Cambridge, England. <https://www.cambridge.org/us/cambridgeenglish/catalog/business-professional-and-vocational/cambridge-business-english-dictionary/> (2011).
- US Department of Health and Human Services, Food and Drug Administration. Statement on new steps to advance digital health policies that encourage innovation and enable efficient and modern regulatory oversight. <https://www.fda.gov/news-events/press-announcements/statement-new-steps-advance-digital-health-policies-encourage-innovation-and-enable-efficient-and> (2019).
- US Department of Health and Human Services, Food and Drug Administration. Digital health software precertification (Pre-Cert) program. <https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program> (2019).
- Li, V. Senior Medical Device Regulation Expert. Changes to existing medical software policies resulting from Section 3060 of the 21st Century Cures Act. <http://www.fdalawblog.net/2019/10/changes-to-existing-medical-software-policies-resulting-from-section-3060-of-the-21st-century-cures-act/> (2019).
- US Department of Health and Human Services, Food and Drug Administration. FDA warns JUUL Labs for marketing unauthorized modified risk tobacco products, including in outreach to youth. <https://www.fda.gov/news-events/press-announcements/fda-warns-juul-labs-marketing-unauthorized-modified-risk-tobacco-products-including-outreach-youth> (2019).
- US Department of Health and Human Services, Food and Drug Administration. Software as a Medical Device (SaMD). <https://www.fda.gov/medical-devices/digital-health/software-medical-device-samd> (2018).
- Anonymous. Reproducibility in machine learning for health. *OpenReview.net*. <https://openreview.net/forum?id=HyLG52lpLN> (2019).
- Gebru, T. et al. Datasheets for datasets. <https://arxiv.org/pdf/1803.09010.pdf> (2019).
- Ross, H. & Bassoff, N. W. The “Dataset nutrition label project” Tackles dataset health and standards. <https://cyber.harvard.edu/story/2019-01/dataset-nutrition-label-project-tackles-dataset-health-and-standards> (2019).
- Clinical Trials Transformation Initiative. CTTI unveils new database of feasibility studies on mobile technologies in clinical research. <https://www.ctti-clinicaltrials.org/news/ctti-unveils-new-database-feasibility-studies-mobile-technologies-clinical-research> (2019).
- Goldsack, J., Chasse, R. A. & Wood, W. A. Digital endpoints library can aid clinical trials for new medicines. <https://www.statnews.com/2019/11/06/digital-endpoints-library-clinical-trials-drug-development/> (2019).
- Digital Medicine Society (DiMe). DiMe’s Library of Digital Endpoints. <https://www.dimesociety.org/index.php/knowledge-center/library-of-digital-endpoints> (2019).
- Beam, A. L., Manrai, A. K. & Ghassemi, M. Challenges to the reproducibility of machine learning models in health care. *JAMA*. <https://doi.org/10.1001/jama.2019.20866>. <https://jamanetwork.com/journals/jama/fullarticle/2758612> (2020).
- Coravos, A., Chen, I., Gordhanas, A. & Stern, A. D. We should treat algorithms like prescription drugs. <https://qz.com/1540594/treating-algorithms-like-prescription-drugs-could-reduce-ai-bias/> (2019).
- Woods, B., Coravos, A. & Corman, J. D. The case for a hippocratic oath for connected medical devices: Viewpoint. *J. Med. Internet Res.* **21**, e12568. <https://www.jmir.org/2019/3/e12568/> (2019).
- US Department of Health and Human Services, Food and Drug Administration. Content of premarket submissions for management of cybersecurity in medical devices. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/content-premarket-submissions-management-cybersecurity-medical-devices> (2018).
- US Department of Health and Human Services, Food and Drug Administration. Postmarket management of cybersecurity in medical devices. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/postmarket-management-cybersecurity-medical-devices> (2016).
- Wikipedia. HITRUST. <https://en.wikipedia.org/wiki/HITRUST> (2019).
- Sage Bionetworks. Elements of Informed Consent. https://sagebionetworks.org/tools_resources/elements-of-informed-consent/ (2019).
- PrivacyToolkit. A toolkit for biomedical researchers of best practices for privacy in digital studies. http://designmanual.sagebionetworks.org/privacy_toolkit.html (n.d.).
- HHS.gov. Section 508. <https://www.hhs.gov/web/section-508/additional-resources/index.html> (2019).
- Izmailova, E. S. et al. Evaluation of wearable digital devices in a phase I clinical trial. *Clin. Transl. Sci.* **12**, 247–259. <https://doi.org/10.1111/cts.12602> (2019).
- Eldeas, A. The Internet of Things and cybersecurity: A Q&A with Sen. Edward J. Markey (D-MA) and Rep. Ted Lieu (D-CA). <https://www.aei.org/technology-and-innovation/the-internet-of-things-and-cybersecurity-a-qa-with-sen-edward-j-markey-d-ma-and-rep-ted-lieu-d-ca/> (2018).
- National Telecommunications and Information Administration (NTIA). Communicating IoT Device Security Update Capability to Improve Transparency for Consumers. https://www.ntia.doc.gov/files/ntia/publications/draft_communicating_iiot_security_update_capability_-_jul_14_2017_-_ntia_multistakeholder_process.pdf (2017).
- Harris Interactive. Consumer Internet of Things Security Labelling Survey Research Findings. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798543/Harris_Interactive_Consumer_IoT_Security_Labelling_Survey_Report.pdf (n.d.).
- Emami-Naeini, P., Agarwal, Y. & Cranor, L.F. Privacy and security nutrition label for smart devices. *Privacy Forecast*. <https://privacy.shorensteincenter.org/nutrition-labels> (2019).
- Shen, Y. & Vervier, P. A. Why we need a security and privacy “Nutrition Label” for IoT devices. <https://www.symantec.com/blogs/expert-perspectives/why-we-need-security-and-privacy-nutrition-label-iiot-devices> (2019).
- DataLabel.org <https://www.datalabel.org/> (2019).
- Gropper, A. Patient Privacy Rights Information Governance Label. *SSRN*. <https://ssrn.com/abstract=3439701> (2019).
- Peterson, M., Creating A. Cybersecurity “Energy Star”. <https://www.publicknowledge.org/blog/creating-a-cybersecurity-energy-star/> (2018).

49. Digital Medicine Society (DiMe). Digital medicine society now accepting members. <https://www.prnewswire.com/news-releases/digital-medicine-society-now-accepting-members-300848402.html> (2019).
50. Goldsack, J. Laying the foundation: Defining digital medicine. <https://medium.com/digital-medicine-society-dime/laying-the-foundation-defining-digital-medicine-49ab7b6ab6ef> (2019).
51. NIST. Considerations for a core IoT cybersecurity capabilities baseline. https://www.nist.gov/sites/default/files/documents/2019/02/01/final_core_iot_cybersecurity_capabilities_baseline_considerations.pdf (2019).
52. National Telecommunications and Information Administration. Multistakeholder process: cybersecurity vulnerabilities. <https://www.ntia.doc.gov/other-publication/2016/multistakeholder-process-cybersecurity-vulnerabilities> (2016).
53. Wikipedia. Bug bounty program. https://en.wikipedia.org/wiki/Bug_bounty_program (2019).
54. Fahs, G. et al. Building resilient medical technology supply chains with a software bill of materials. NPJ Digital Med., "IN SAME SPECIAL ISSUE." (2020).
55. Fields, H. 508 compliance: Making your website more accessible. <https://www.webdevelopmentgroup.com/2017/09/508-compliance-making-websites-accessible-for-people-with-disabilities/> (2017).
56. US Department of Health and Human Services, Food and Drug Administration. Guidance on Exculpatory Language in Informed Consent. <https://www.fda.gov/media/81521/download> (2011).

ACKNOWLEDGEMENTS

For refinement terms relating to digital medicine and illustrative examples we gratefully acknowledge Adam Conner- Simons and Pavan Choksi. For formatting work, we gratefully acknowledge Lisa Maturo. For design collaboration for the connector sensor label we gratefully acknowledge Indicius, a digital design agency.

AUTHOR CONTRIBUTIONS

All authors (A.C., M.D., J.G., C.M., M.S., B.W., and W.W.) equally contributed to the following: (1) substantial contributions to the conception or design of the work; (2) drafting the work and revising it critically for important intellectual content; (3) final approval of the completed version; and (4) accountability 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. A.C., C.M., M.S., and design agency, Indicius, created Figs. 2 and 4. A.C. created Fig. 3.

COMPETING INTERESTS

J.G. is an employee of HealthMode. M.D. is a co-inventor of the Cleveland Clinic's MyLegacy intellectual property portfolio, now licensed to Family Care Path, Inc.; as part of this license, she is entitled to a share in both royalties and returns on equity. W.W., B.W. are advisors and equity-holders, and A.C., C.M., and M.S. are employees of Elektra Labs, a healthcare-security company working in digital medicine. W.W. is also an advisor for Koneksa Health, a consultant for Best Doctors, and has received research funding within the last 12 months from Pfizer and Genentech. The other authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41746-020-0237-3>.

Correspondence and requests for materials should be addressed to W.A.W.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020