

<https://doi.org/10.1038/s41746-024-01322-2>

# V3+ extends the V3 framework to ensure user-centricity and scalability of sensor-based digital health technologies



Jessie P. Bakker<sup>1,2,3</sup>✉, Roland Barge<sup>4</sup>, Jacob Centra<sup>1</sup>, Bryan Cobb<sup>5</sup>, Chas Cota<sup>6</sup>, Christine C. Guo<sup>7</sup>, Bert Hartog<sup>8</sup>, Nathalie Horowicz-Mehler<sup>9</sup>, Elena S. Izmailova<sup>10</sup>, Nikolay V. Manyakov<sup>11</sup>, Samantha McClenahan<sup>1</sup>, Stéphane Motola<sup>12</sup>, Smit Patel<sup>1</sup>, Oana Paun<sup>13</sup>, Marian Schoone<sup>14</sup>, Emre Sezgin<sup>15</sup>, Thomas Switzer<sup>5</sup>, Animesh Tandon<sup>16,17,18,19</sup>, Willem van den Brink<sup>14</sup>, Srinivasan Vairavan<sup>120</sup>, Benjamin Vandendriessche<sup>1,21</sup>, Bernard Vrijens<sup>13,22</sup> & Jennifer C. Goldsack<sup>1</sup>

We propose the addition of *usability validation* to the extended V3 framework, now “V3+”, and describe a pragmatic approach to ensuring that sensor-based digital health technologies can be used optimally at scale by diverse users. Alongside the original V3 components (verification; analytical validation; clinical validation), usability validation will ensure user-centricity of digital measurement tools, paving the way for more inclusive, reliable, and trustworthy digital measures within clinical research and clinical care.

Sensor-based digital health technologies (sDHTs) represent a paradigm shift in the way clinical data are captured, by providing the means to collect high-resolution data in real-world remote settings over long time periods, and reflecting meaningful functional changes that are less prone to observer bias<sup>1</sup>. Within the period of 2019–2024, we have observed a 10-fold increase in the number of sDHT-derived measures adopted in industry-sponsored interventional trials, over 100 of which are positioned as primary endpoints<sup>2</sup>. In 2023, the first pivotal trial adopting a digital measure as a United States Food and Drug Administration (FDA) endorsed primary endpoint was reported<sup>3</sup>, and the European Medicines Agency (EMA) qualified a digital measure as a primary efficacy endpoint<sup>4</sup>. More recently, the FDA qualified atrial fibrillation burden as the first medical device development tool captured by an sDHT<sup>5</sup>. Although the integration of sDHTs for remote patient monitoring and its use in clinical practice has been more gradual, implementation has been fueled by the development of a more robust infrastructure designed to tackle the complexities of large healthcare systems<sup>6,7</sup>.

The increasing trust and continued investment in sDHTs by healthcare providers, study sponsors, regulators, payers, and patients has been supported by V3, a modular framework for evaluating the quality of sDHTs according to<sup>8</sup>:

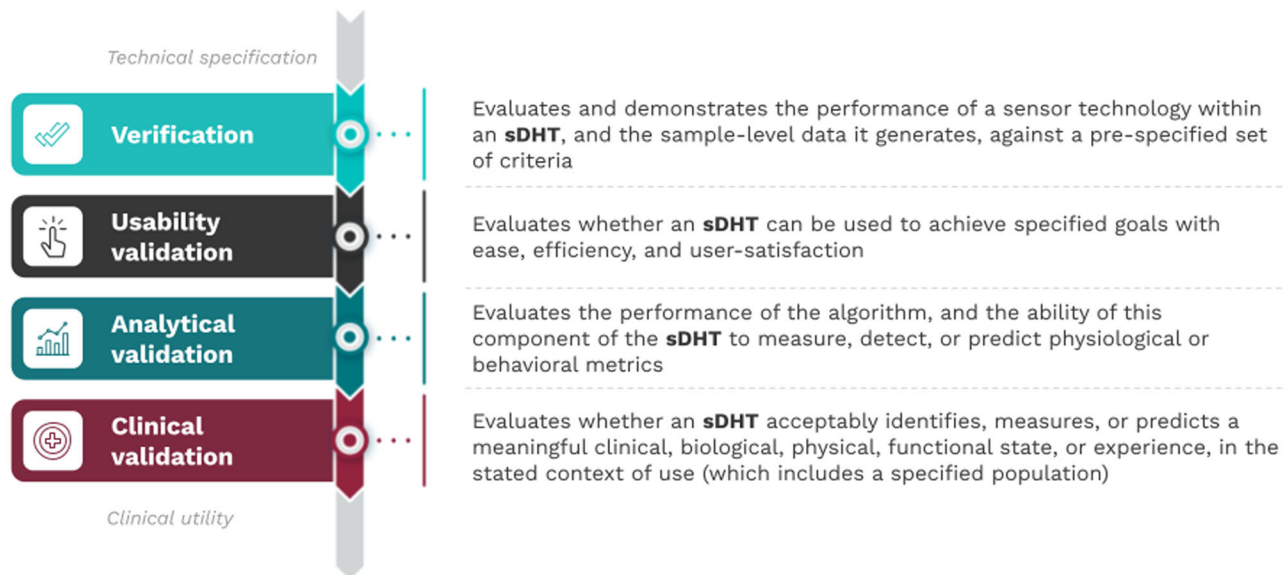
- *Verification*, which evaluates the performance of the sensor(s) against a pre-specified set of criteria;
- *Analytical validation*, which evaluates the performance of the algorithm(s) in terms of its ability to measure, detect, or predict physiological or behavioral metrics; and

- *Clinical validation*, which evaluates the extent to which the digital measure acceptably identifies, measures, or predicts a meaningful clinical, biological, physical, functional state, or experience in the specified context of use including the patient population.

Since its publication in 2020, the V3 framework has become foundational to the evaluation of sDHTs for technical, scientific, and clinical performance, having been adopted and/or referred to by individuals at the EMA<sup>9,10</sup>, FDA<sup>11</sup>, and over 250 industry and academic researchers<sup>12</sup>.

## Development and implementation of sDHTs at scale requires an extension of V3

In the 4 years since the V3 framework was published, regulations focused on the use of sDHTs for remote data capture have matured<sup>13</sup> and reimbursement pathways have been developed, with >10 common procedural terminology codes now available for services related to digital measurement<sup>14</sup>. Furthermore, large-scale adoption of general wellness products has empowered individuals to monitor their own health, creating expectations for implementing sDHTs into both research and healthcare settings<sup>15</sup>. As clinical research sponsors and healthcare organizations take digital clinical measures to scale, challenges related to the implementation of sDHTs across diverse populations, different settings, and multifarious methodological approaches have become pressing<sup>16,17</sup>. For example, tremor classification data were missing for 50% of participants in the *Wearable Assessment in the Clinic and at Home in Parkinson's Disease* study due to the inadvertent deactivation of device permissions which prevented transfer of passive data



**Fig. 1 | From V3 to V3+.** The V3+ framework includes four components, with usability validation as the latest addition. Although depicted sequentially, V3+ is a modular framework as shown in Fig. 5. sDHT sensor-based digital health technologies. See Table 1 for definitions.

from the smartwatch to the study database<sup>18</sup>. It is possible that this data missingness could have been prevented or minimized had the user interface been designed to minimize or avoid the accidental change of permissions and/or more extensive preliminary testing amongst target users. Moreover, consider the FDA recall of a specific blood glucose monitor because the product may inadvertently switch the unit of measure from mg/dL to mmol/L or vice versa during battery insertion during normal use, highlighting a potential impact on usability-related safety<sup>19</sup>. Although the underlying causes of these examples are multifactorial and complex, both highlight the lack of a unified set of methodological best practices to support optimal sDHT usability, defined as the extent to which an sDHT can be used to achieve specified goals with ease, efficiency, and user satisfaction (see Table 1 for definitions of terms).

sDHTs adopt a variety of form factors, measure a range of different health concepts, and may be regulated medical devices or non-regulated products<sup>20</sup> (see Box 1 for a description of the technologies in-scope for V3+). This heterogeneity has almost certainly contributed to the disparate methodological approaches and evaluation criteria adopted in sDHT usability studies published over the last decade, as described in our recent systematic scoping review<sup>21</sup>. To embrace this heterogeneity and support the development of fit-for-purpose sDHTs, our goal is to build on the foundation of V3 by adding an evidence-based component to the framework addressing sDHT usability validation, as shown in Fig. 1 (see Box 2 for a terminology rationale). Through dissemination of the extended V3 framework, referred to henceforth as V3+, we will ensure that sDHTs can be developed and evaluated according to state-of-the-art approaches to human factors engineering, and support the further advancement of high quality, well-validated, and easy-to-use digital measures for optimizing scientific, clinical, regulatory, and payer decision-making.

### Existing guidance and standards relevant to sDHT usability and related concepts

Recently finalized FDA guidance clarifies that digital tools used for remote data acquisition in a clinical investigation may be a regulated medical device per the definition under Section 201(h) of the FD&C Act, or a non-regulated product such as a general wellness product<sup>20</sup>. Many, but not all, of the latter are developed for and marketed primarily towards consumers, and are generally considered to be low-risk products<sup>22</sup>. In addition, sDHTs that are intended for use as medical

devices may be considered high risk (such as some implantables<sup>23</sup>) or low/moderate risk (which may be the case for some ingestibles<sup>24</sup>, wearables<sup>25</sup>, and ambient<sup>26</sup> tools, depending on the nature of the product itself and the developer's claims).

Given this substantial variability in the ways sDHTs are regulated and marketed, best practices relevant to sDHT usability and related concepts may be found in numerous global regulatory documents<sup>20,27–34</sup> as well as industry standards developed for medical devices (IEC 62366-1:2015<sup>35</sup> and IEC/TR 62366-2:2016<sup>36</sup>), interactive systems (ISO 9241-210:2019)<sup>37</sup>, everyday products (ISO 20282-1:2006<sup>38</sup>), consumer products (ISO/TS 20282-2:2013<sup>39</sup>), and others<sup>40</sup>. Although the recommendations in these documents are not completely aligned, many elements of the usability validation component of V3+ described below have been drawn from these existing descriptions of best practices and we have endeavored to remain consistent with the concepts and terminology wherever possible.

### The usability validation component of V3+ is comprised of four key activities

Following the development of a proposed intended use statement, which is applicable to all four components of V3+ and which we recommend developing for both regulated and non-regulated sDHTs (see Box 3 for further information), the usability validation process includes four key activities as shown in Fig. 2.

#### Key activity 1: develop the use specification

The intended use statement has a direct impact on both the technical specification (a comprehensive description of the sDHT, including but not limited to the dimensions and materials; the units of the sampled data, sampling frequency, and the sampling range of each sensor; battery life; data storage; data transmission protocols; environmental limits; and other technical information) and the use specification (a living document containing a comprehensive description of who the intended sDHT user groups are; where, when, and how each user group will interact with the sDHT including the data generated; and their motivations for doing so, which may also be defined using 'use cases'). The use specification is intentionally depicted as the counterpart to the technical specification, of equal importance and with a bidirectional relationship. The process for creating the use specification, the first activity of usability validation, is summarized in Table 2.

Table 1 | Glossary of terms

|                                                                                                                                                                                                                                                                                                                                                                                                                            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Abnormal use:</b> Intentional reckless use or sabotage <sup>27,31,32</sup> , beyond the scope of the <i>use-related risk analysis</i> .                                                                                                                                                                                                                                                                                 |
| <b>Actionability:</b> The extent to which <i>users</i> of diverse backgrounds, languages, and varying levels of health literacy understand the actions or <i>user tasks</i> they should complete in response to clinical data or other information presented to them <sup>59</sup> , typically assessed in a <i>knowledge task study</i> .                                                                                 |
| <b>Clinical utility:</b> The extent to which implementing an sDHT leads to improved health outcomes or provides useful information about diagnosis, treatment, management, or prevention of a disease <sup>27</sup> .                                                                                                                                                                                                      |
| <b>Cognitive walkthrough:</b> A <i>formative evaluation</i> in which A) <i>usability</i> experts break down <i>user tasks</i> and identify possible <i>use-errors</i> <sup>31</sup> and/or B) <i>usability</i> experts guide <i>users</i> through <i>user tasks</i> while encouraging <i>users</i> to think aloud <sup>27,32</sup> .                                                                                       |
| <b>Context of use:</b> A statement that fully and clearly describes the way the sDHT is to be used and the purpose of the use <sup>58</sup> .                                                                                                                                                                                                                                                                              |
| <b>Contextual inquiry:</b> Observation of <i>users</i> interacting with a functional sDHT in the intended <i>use environment</i> , with staff asking questions during or after use <sup>27,31</sup> .                                                                                                                                                                                                                      |
| <b>Critical task:</b> A <i>user task</i> that, if not performed or performed incorrectly, would or could lead to serious harm <sup>27,31,32,34</sup> .                                                                                                                                                                                                                                                                     |
| <b>Ease of use:</b> The ease with which a <i>user</i> is able to perform <i>user tasks</i> <sup>44</sup> , captured through self-report (such as the mental demand or effort required to complete a task) or objective measures (such as the number of actions, number of attempts, or time required to complete a task).                                                                                                  |
| <b>Efficiency:</b> The ease with which a <i>user</i> is able to perform <i>user tasks</i> after having learned how to use the sDHT <sup>44</sup> .                                                                                                                                                                                                                                                                         |
| <b>End-user:</b> A user from whom sDHT-derived clinical data are captured; that is, the patient or participant.                                                                                                                                                                                                                                                                                                            |
| <b>Error recovery:</b> The ability of a <i>user</i> to make a correction following a <i>use-error</i> in order to complete a <i>user task</i> <sup>59</sup> .                                                                                                                                                                                                                                                              |
| <b>Fit-for-purpose:</b> A conclusion that the level of validation associated with an sDHT is sufficient to support its <i>context of use</i> <sup>58</sup> .                                                                                                                                                                                                                                                               |
| <b>Formative evaluation:</b> A research study or activity undertaken to evaluate <i>usability</i> of a prototype sDHT, with the goals of understanding <i>user</i> interactions with the sDHT and gathering information to inform design modifications <sup>27,31,32</sup> .                                                                                                                                               |
| <b>Gap analysis:</b> A systematic approach to comparing two or more statements or scenarios. For the <i>usability validation</i> component of V3+, a gap analysis will identify differences between the <i>intended use</i> and <i>context of use</i> statements.                                                                                                                                                          |
| <b>Human factors:</b> The application of knowledge about human behavior, abilities, limitations, and other characteristics of <i>users</i> to the design and development of an sDHT to optimize <i>usability</i> within a defined <i>intended use</i> or <i>context of use</i> . This definition incorporates terminology and concepts from FDA <sup>27</sup> , MHRA <sup>31</sup> , and NMPA (translated) <sup>32</sup> . |
| <b>Human-centered design:</b> An approach to interactive systems that aims to make systems usable and useful by focusing on the <i>users</i> , their needs, and requirements, and by applying <i>human factors</i> and <i>usability</i> knowledge and techniques <sup>37</sup> .                                                                                                                                           |
| <b>Indications for use:</b> A statement that describes the disease or condition the sDHT is designed to diagnose, treat, prevent, cure, or mitigate, including a description of the patient or participant population for which the sDHT is intended <sup>60</sup> .                                                                                                                                                       |
| <b>Intended use:</b> A statement that describes the specific clinical circumstance or purpose for which an sDHT is being developed, including the <i>indications for use</i> <sup>61</sup> .                                                                                                                                                                                                                               |
| <b>Knowledge task study:</b> A study undertaken to assess <i>understandability</i> and <i>actionability</i> <sup>27</sup> .                                                                                                                                                                                                                                                                                                |
| <b>Learnability:</b> The ease with which a <i>user</i> is able to perform <i>user tasks</i> during their first encounter with the sDHT <sup>59</sup> .                                                                                                                                                                                                                                                                     |
| <b>Memorability:</b> The ease with which a <i>user</i> is able to perform <i>user tasks</i> after a period of non-use, assessed in a test-retest paradigm <sup>59</sup> .                                                                                                                                                                                                                                                  |
| <b>Production-equivalent:</b> A sample sDHT of final design assembled in a way that differs from, but is equivalent to, the manufacturing processes used for the marketed sDHT <sup>62</sup> .                                                                                                                                                                                                                             |
| <b>Sensor-based digital health technologies:</b> Connected digital medicine products that process data captured by mobile sensors using algorithms to generate measures of behavioral and/or physiological function, also referred to as biometric monitoring technology <sup>8</sup> .                                                                                                                                    |
| <b>Summative evaluation:</b> A research study undertaken on a <i>production-equivalent</i> or marketed sDHT including all components of the <i>user interface</i> , with the goal of demonstrating <i>usability</i> amongst a representative sample under conditions reflecting the <i>intended use</i> or <i>context of use</i> <sup>31,32</sup> , referred to by the FDA as “human factors validation” <sup>27</sup> .   |
| <b>Technical specification:</b> A comprehensive description of the sDHT dimensions and materials; the units of the sampled data, sampling frequency, and the sampling range of each sensor; battery life; data storage; data transmission protocols; environmental limits; and other technical information.                                                                                                                |
| <b>Understandability:</b> The extent to which <i>users</i> of diverse backgrounds, languages, and varying levels of health literacy understand the clinical data or other information (such as instructions, cautions, warnings, and contraindications) presented to them <sup>59</sup> , typically assessed in a <i>knowledge task study</i> .                                                                            |
| <b>Usability:</b> The extent to which an sDHT can be used to achieve specified goals with <i>ease</i> , <i>efficiency</i> , and <i>user satisfaction</i> within a defined <i>intended use</i> or <i>context of use</i> . This definition incorporates terminology and concepts from FDA <sup>27</sup> , MHRA <sup>31</sup> , NMPA (translated) <sup>32</sup> , and ISO 9241-210:2019 <sup>35</sup> .                       |
| <b>Usability validation:</b> Evaluation and demonstration of <i>usability</i> .                                                                                                                                                                                                                                                                                                                                            |
| <b>Use environment:</b> The setting(s) in which the sDHT is intended to be used <sup>27,31,32</sup> .                                                                                                                                                                                                                                                                                                                      |
| <b>Use-error:</b> An action or lack of action which may result in a <i>use-related hazard</i> <sup>27</sup> . “Use-error” is preferable to “user-error” as it avoids the implication that the user is at fault.                                                                                                                                                                                                            |
| <b>Use-related hazard:</b> A source of potential harm resulting from a <i>use-error</i> . <sup>27</sup> Use-related hazards are those associated with <i>user</i> interactions, rather than issues associated with sDHT technical performance or hazards such as sharp edges, unsafe operating temperatures, or non-biocompatible materials <sup>27</sup> .                                                                |
| <b>Use-related risk analysis:</b> A living document describing reasonably foreseeable risks associated with use of an sDHT, and a detailed plan to mitigate those risks <sup>28,31</sup> .                                                                                                                                                                                                                                 |
| <b>Use scenario:</b> Rich descriptions of several likely <i>use environments</i> and how interactions with the sDHT might differ between them <sup>27,31,32</sup> .                                                                                                                                                                                                                                                        |
| <b>Use specification:</b> A living document containing detailed descriptions of all <i>user</i> groups, all <i>use environments</i> , and all aspects of the sDHT <i>user interface</i> <sup>31</sup> .                                                                                                                                                                                                                    |
| <b>Usefulness:</b> The extent to which a <i>user</i> finds the sDHT, or its specific features/functions, to be valuable, productive, and/or helpful <sup>63</sup> .                                                                                                                                                                                                                                                        |
| <b>User:</b> Any individual who may interact with an sDHT as part of normal use, including the <i>end-user</i> and their care partner(s) as well as individuals acting in a professional capacity such as those in clinical, research, and/or administrative roles <sup>27,31,32</sup> .                                                                                                                                   |
| <b>User interface:</b> All points of interaction a <i>user</i> may have with the sDHT as a holistic system, including hardware, software, all components and accessories, packaging, instructions for use and other documentation, and user training <sup>27,31,32</sup> .                                                                                                                                                 |
| <b>User satisfaction:</b> The extent to which a <i>user</i> finds the sDHT to be pleasant to use <sup>59</sup> , which may reflect trust, comfort, esthetics, engagement, desirability, emotional response/s, and other considerations.                                                                                                                                                                                    |
| <b>User task:</b> An action or set of actions performed by a <i>user</i> to achieve a specific goal, often referred to simply as “task” <sup>27</sup> .                                                                                                                                                                                                                                                                    |

Definitions presented in this glossary describe each term as used in V3+, and are not always verbatim from the referenced source.

## Box 1 | What technologies does V3+ apply to?

The V3+ framework applies to **sensor-based digital health technologies (sDHTs)**, defined as *connected digital medicine products that process data captured by mobile sensors using algorithms to generate measures of behavioral and/or physiological function*<sup>8</sup>. In the original V3 paper, these technologies were referred to as biometric monitoring technologies (BioMeTs). sDHTs are a subset of the broader category of technologies referred to as DHTs, defined as *systems that uses computing platforms, connectivity, software, and/or sensors, for health care and related uses*, which includes non-sensor based tools such as electronic patient-reported outcome (ePRO) platforms.

sDHTs may be:

- **Wearable:** Worn on the body, such as a chest strap or adhesive patch;
- **Implantable:** Implanted into the body, such as a loop recorder;

- **Ingestible:** Swallowed and excreted, such as a core body temperature sensor; or
- **Ambient:** Placed in the environment and including both passive and active data capture, such as a microphone, a connected pharmaceutical package or drug delivery device, a camera, a mattress pad, or connected weight scales.

sDHTs include **hardware**, **firmware**, and **software** components. These components may be combined into a single product, or split across multiple products; for example, by using the hardware/firmware of a hardware-based sDHT to collect signals for analysis in a separate sDHT software application.

**V3+ applies to all sDHTs used for generating digital clinical measures**, regardless of a tool's regulatory status or the way(s) in which it is marketed, purchased, or accessed.

## Box 2 | Language matters: the rationale for usability validation

Many evaluation frameworks that informed the development of V3<sup>66,67–69</sup> include **clinical utility** as the framework component directly following clinical validation: *The extent to which implementing a medical product leads to improved health outcomes or provides useful information about diagnosis, treatment, management, or prevention of disease*<sup>67</sup>. Notably, however, all of these frameworks position clinical utility in terms of real-world health outcomes—that is, after the point at which the product is introduced to the market—whereas V3+ applies to both pre-market and post-market sDHTs. As such, only some of the questions and processes described within clinical utility in the aforementioned frameworks, such as the importance of user-centricity and the reliance on trustworthy

measurements for appropriate decision-making, are applicable to sDHTs within V3+.

We are therefore not proposing the term clinical utility to describe the fourth component of V3+; instead, we have identified **usability** as being the necessary process that allows implementation and sustainable adoption of an sDHT to be user-centric, scalable, and informative for clinical and scientific decision-making.

As is the case for analytical and clinical validation, we have adopted the term **usability validation** to refer to this entire component of the V3+ framework (see Fig. 1), recognizing that several preliminary activities and studies may precede the final study/ies intended to generate evidence supporting usability.

### Key activity 2: conduct a use-related risk analysis

The use specification contains the information necessary to complete a preliminary version of the use-related risk analysis, which is undertaken iteratively and collaboratively to identify foreseeable risks associated with the use of the sDHT and develop a plan for minimizing or eliminating known risks. The analysis should address reasonably foreseeable misuse, including use of the sDHT by unintended users, distinct from abnormal use which involves intentional reckless use or sabotage<sup>27,31,32,34</sup>.

The analysis begins by compiling a list of all user tasks and potential use-errors, which are actions (or lack thereof) which may result in a use-related hazard; see Table 1 for definitions, and note that “use-error” is preferable to “user-error” as it avoids the implication that the user is at fault. Once identified, use-related hazards should be categorized according to the seriousness of the potential harm in order to identify critical tasks<sup>27</sup> (see Fig. 3). Established risk management approaches also account for the severity of potential harm and the likelihood of occurrence<sup>41</sup>. As described in FDA guidance<sup>27,34</sup>, the ideal approach is for use-errors to be “designed out”, known as inherent safety by design; however, if this is not feasible, protective measures such as alarms or error messages may be adopted. The least favorable measures include the provision of instructions to avoid use-errors which can be included in the user manual or user training.

Importantly, harm may arise not only as a result of use-errors, but as a result of poor usability leading to sub-optimal sDHT adherence and consequently, excessive missing data. Harm resulting from missing or unreliable sDHT data may include false-negative diagnostic tests, missed signs of

clinical deterioration requiring intervention, or inappropriate treatment-related decisions<sup>42</sup>. For example, a 2021 study found that use of a wearable temperature sensor allowed for the detection of fever amongst individuals with cytokine release syndrome ~5 h earlier than standard temperature checks, considered a clinically significant improvement for administering antibiotics<sup>43</sup>, which would not be possible in a scenario with extensive missing data.

### Key activity 3: conduct iterative formative evaluation of sDHT prototypes

Formative evaluations include any activity or research study undertaken with the goals of describing user tasks, identifying use-errors, and gathering the information necessary to inform design improvements<sup>27,31,32</sup>. The circular arrow in Fig. 2 indicates that formative evaluation and sDHT hardware, software, and workflow design proceed iteratively as prototypes become increasingly mature; however, the design process itself is out of scope for V3+. Similarly, formative evaluations will allow the use-related risk analysis to be updated as new use-errors are discovered, which can then be addressed in an updated design and assessed during a subsequent formative evaluation. This iterative process continues until the sDHT demonstrates sufficient usability to progress to summative evaluation, as described below. Common data capture methods for formative evaluations are described in Table 3.

Formative evaluation is typically conducted as a series of incremental steps, beginning with small samples (for example,  $n \leq 5$ ) before recruiting



### Box 3 | The sDHT intended use and context of use

The V3+ process for *pre-market* sDHTs begins with defining the **intended use statement**, which describes the specific clinical circumstances or purpose for which the sDHT is being developed including the indications for use. For *post-market* sDHTs, the V3+ process begins with defining the **context of use statement**, which fully and clearly describes the way the sDHT is to be used and the purpose of the use.

The intended use and context of use statements address similar questions; the difference is that the former summarizes the technology developer’s claim/s or statement/s regarding what the sDHT does, whereas the latter describes the manner in which the sDHT will be implemented.

A stakeholder responsible for implementing an sDHT that is a regulated medical device can be assured that some degree of usability validation evidence was generated to support the original regulatory submission. This evidence, however, aligns with the original **intended use** which may differ from the proposed **context of use**, prompting the need for further evaluation. For example, using a home sleep test device in a clinical trial for seven consecutive nights when the device was developed for single-night use is likely to require additional usability validation data even though the sDHT is a regulated medical device. Thus, in some cases existing medical device usability data may be necessary but not sufficient for supporting the assessment of fit-for-purpose.

Note that “intended use” and “context of use” are regulatory terms which we recognize are not used with respect to non-regulated products; however, the questions addressed in each statement are critical for all four components of V3+. We, therefore, recommend that equivalent statements be developed for general wellness or other products for which regulations are either not applicable or not enforced.

| Intended use:                                                                                                                                                                                                                                                             | Context of use:                                                                                                                                                                                                                                                                         |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>• <i>What</i> does the sDHT do?</li><li>• <i>Who</i> are the intended users?</li><li>• <i>Where</i> should the sDHT be used?</li><li>• <i>When</i> should the sDHT be used?</li><li>• <i>How</i> should the sDHT be used?</li></ul> | <ul style="list-style-type: none"><li>• <i>What</i> will the sDHT be used for?</li><li>• <i>Who</i> are the population(s) of interest?</li><li>• <i>Where</i> will the sDHT be used?</li><li>• <i>When</i> will the sDHT be used?</li><li>• <i>How</i> will the sDHT be used?</li></ul> |

larger samples of participants that increasingly represent the diversity of the intended end-user population, as well as all other user groups described in the use specification (see Table 2). We recommend that sDHT remote data capture be undertaken as soon as is feasible during the formative evaluation stage, in order to address any use-errors or technical difficulties associated with data storage and transfer while the sDHT design is flexible enough to accommodate required modifications (see Box 4 for an overview of human-centered design principles).

**Key activity 4: complete summative evaluation of the final sDHT**  
The purpose of summative evaluations (referred to by the FDA as human factors validation studies<sup>27,34</sup>) is to demonstrate that the final (or production-equivalent<sup>44</sup>) version of the sDHT—including all components, accessories, packaging, instructions for use, additional documentation, and user training—is sufficiently usable within the proposed intended use including the intended user population(s)<sup>31,32</sup>.

Summative studies should be designed to evaluate all essential user tasks including critical tasks, and all components of the user interface.

Recommended methods include simulated-use and actual-use, during which the study participants interact with the sDHT independently and naturally without prompts or feedback from study staff (see Table 3)<sup>27,31,32</sup>. The setting and duration of use should ideally reflect the conditions of actual use, and it is critical that the study design includes remote data capture without expert supervision beyond the training and support that is intended to be offered to intended users, such as helpdesk troubleshooting.

As is the case with analytical validation, the summative study sample(s) should represent the intended use population; however, a unique characteristic of usability validation studies is that all user groups identified in the use specification—such as carepartners, clinicians, investigators, research staff, and administrative staff—must be studied (see Table 2)<sup>27,31,32</sup>. Within each user group, efforts should be made to ensure socio-demographic diversity, as well as diversity across the user characteristics that the investigators believe will have the greatest impact on sDHT use. For example, body habitus might be important for ensuring appropriate sizing of a wearable sDHT, while dexterity, which can be impacted by age or certain health conditions, might be important for an sDHT that requires a user to input data on a small touchscreen.

Summative evaluation is the final activity in the usability validation component of V3+, and as is the case with verification and analytical validation, quantitative pass/fail criteria should be specified a priori. Table 4 provides real-world examples where each of the four usability validation activities has been applied during development or evaluation of sDHTs.

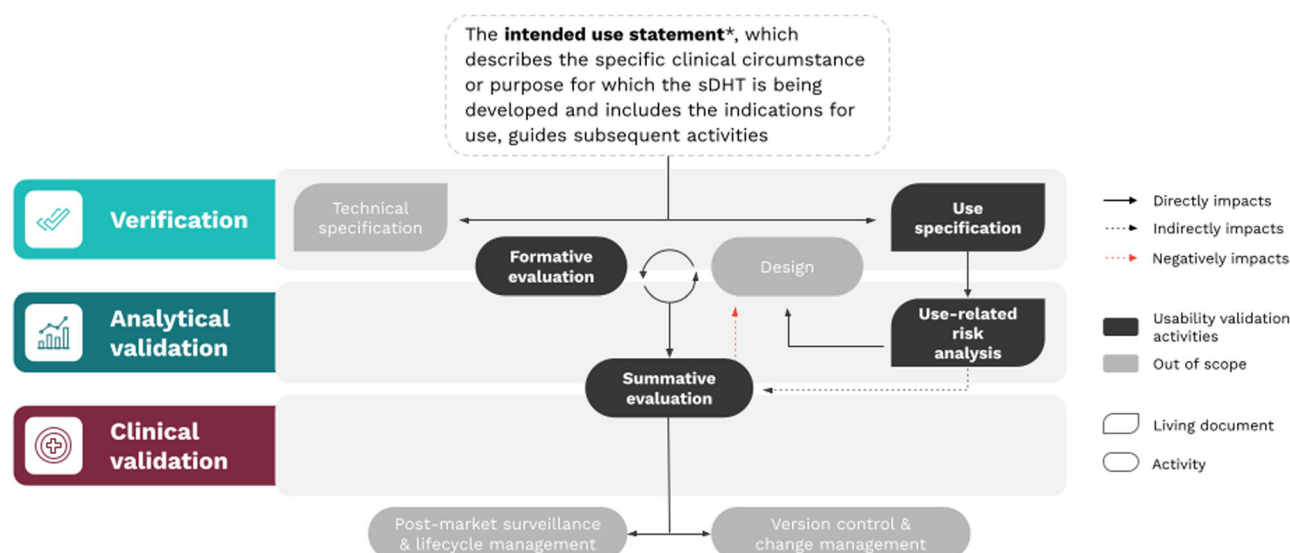
### Usability validation is applicable to pre- and post-market sDHTs

As is the case for the original V3 framework, there are two primary use cases to which V3+ usability validation is applicable. The first relates to technologies that are under development (that is, pre-market technologies), which may be completely novel products or next-generation versions of existing products. In this scenario, the technology developer is the party responsible for undertaking the processes described in V3+. Pre-market processes of regulated medical devices fall within a controlled regulatory space with applicable regulations and guidance; however, this is not the case for sDHTs that are not marketed as medical devices<sup>45</sup>. The second use case is for technologies that are commercially available (that is, post-market technologies). In this scenario, the party interested in deploying the sDHT—either the clinical research sponsor or the healthcare provider—is responsible for ensuring that the product has been assessed satisfactorily according to V3+. In addition, it is recommended that where possible, sponsors and healthcare providers establish a collaborative relationship with the sDHT developer to improve usability of next-generation products and ensure that previously unidentified use-errors are accounted for.

Figures 2 and 4 depict the elements of the usability validation component of V3+ that are applicable to pre-market and post-market sDHTs, respectively. The activities themselves are identical in each case; the difference is that only a subset are applicable to post-market sDHTs, as there is no immediate opportunity to modify the product design or undertake formative evaluation. All post-market activities are therefore relevant to the pre-market use case, but the reverse is not true. Thus, throughout this paper, we describe the pre-market use case in depth before describing how the post-market use case differs, even though there is arguably a greater need for a usability evaluation framework in the post-market scenario where little guidance and standards exist.

In addition, there are scenarios which may be considered a hybrid between pre- and post-market application of usability validation. For example, the developer of a pre-market stand-alone sDHT software application that relies on analyzing signals captured from a post-market hardware-based sDHT may proceed through the steps shown in Fig. 2 including formative evaluation of the software, but they will have no opportunity to modify the technical specification of the hardware.

The activities undertaken during usability validation of a post-market sDHT are driven by a gap analysis comparing the original intended use statement and the current context of use statement. The gaps



**Fig. 2 | The four key activities of the usability validation component of V3+.** This figure depicts the four usability validation activities of V3+ in the order in which they are typically undertaken, as well as the approximate alignment with verification, analytical validation, and clinical validation. Importantly, usability validation is

rarely linear, and therefore the timing and order of activities will vary case-by-case. See Table 1 for definitions. \*The intended use statement is a key component of the labeling of regulated medical devices. We recommend development of an equivalent statement for non-regulated sDHTs.

**Table 2 | Creating the use specification**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Part One: Identify all user groups and compile a series of representative user descriptions                                                                                                                                                                                                                                                                                                                                                                                                                            |
| End-users are the individuals from whom sDHT-derived clinical data will be captured. In a research study, the end-users are the participants including healthy controls, while in clinical settings the end-users are the patients being cared for. Additional sDHT users might include carepartners, clinicians, investigators, research staff, and administrative staff.                                                                                                                                             |
| For each user group, describe aspects including but not limited to socio-demographics and cultural customs; health and technology literacy; physical, sensory, and cognitive capabilities; anthropometry; disease characteristics including comorbidities; and vulnerable populations. Not all characteristics will apply to all user groups.                                                                                                                                                                          |
| Part Two: Identify all likely use environments and compile a series of use scenarios                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| sDHTs may be used in many environments such as the home, clinic/lab settings, assisted living facilities, workplaces, educational institutions, community and leisure spaces, and during transit.                                                                                                                                                                                                                                                                                                                      |
| For each use environment, describe aspects including but not limited to the temperature, humidity, lighting, noise levels, space availability, cleanliness and sterility, privacy and the presence of other individuals, security, and risk of theft, power and network availability, and the presence of (and interoperability with) other sDHTs.                                                                                                                                                                     |
| Part Three: Compile an in-depth description of all aspects of the sDHT user interface                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| List the various ways users might interact with the sDHT (including hardware, software, and all accessories and components such as packaging, chargers and cables, batteries, replacement/consumable parts) and their motivations for doing so. Consider visual, auditory, and tactile cues. Compile all written materials and instructions, and describe what training, troubleshooting, and support will be available to each user group (typically identified and documented as part of a training needs analysis). |
| Finally, use the information from all three parts of the use specification to describe the various foreseeable interactions that users from within each user group will have with the sDHT.                                                                                                                                                                                                                                                                                                                            |

between how the sDHT was developed and the way that it is planned to be used will guide the responsible stakeholders as they determine whether existing usability data are sufficient to demonstrate that the sDHT is fit-for-purpose, or whether further evidence is required. If the latter, the use specification and use-related risk analysis should be undertaken as described above (see Table 2 and Fig. 3), before proceeding directly to summative evaluation. In many cases, summative evaluation may take the form of a bridging study designed to gather validation evidence addressing only the gaps between the intended use and context of use, thereby building on the existing evidence base.

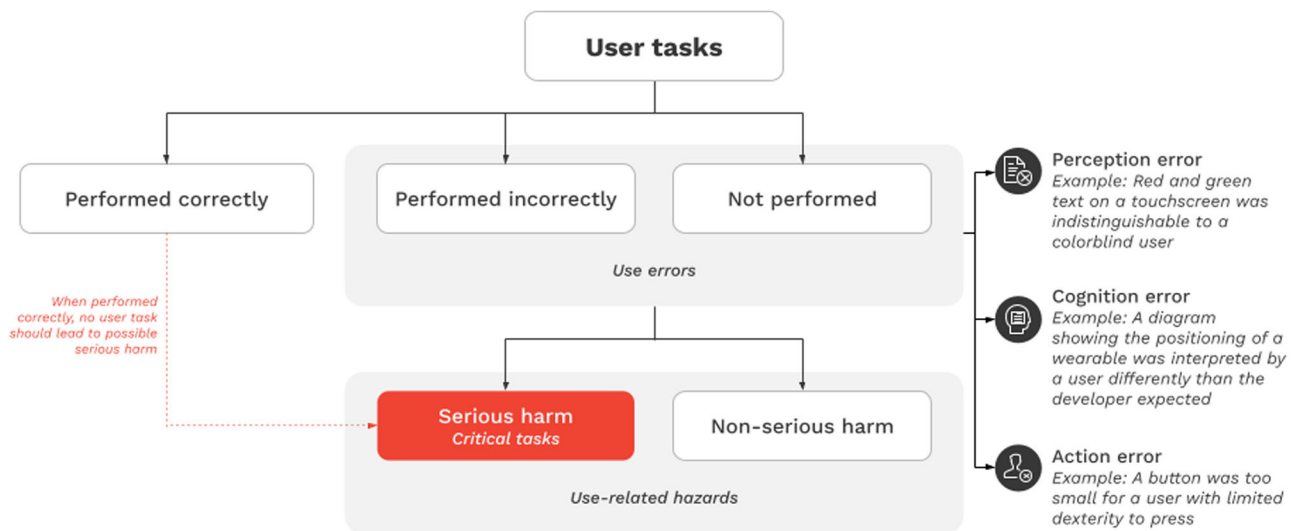
The context of use is also important in terms of implementation in research versus healthcare settings. For example, end-users have expressed that the ability to access, understand, and potentially act on their own sDHT-derived clinical data is a major driver of user satisfaction and engagement<sup>46</sup>; however, real-time data sharing may not be suitable during a research study due to the risk of introducing behavior change<sup>20</sup>. As a result, perceived usefulness and user satisfaction may differ substantially between settings.

Table 5 summarizes the process of usability validation, while Table 6 describes the application of usability validation in practice.

## Evaluation criteria, sample size, and statistical considerations

Several models for evaluating usability and related concepts are available, such as the Technology Acceptance Model<sup>47</sup> and the Unified Theory of Acceptance and Use of Technology<sup>48</sup>. While such well-established models are informative, we encourage the field to develop and implement sDHT-specific usability evaluation methods that address the challenges related to these technologies. In addition to ease of use, efficiency, and user satisfaction, common usability evaluation criteria include learnability, memorability, usefulness, use-errors, and ease of error recovery (defined in Table 1). In some cases, it may be possible to leverage the sDHT backend infrastructure to capture additional metrics, such as the time taken to complete specific tasks.

While the sample size of formative evaluations is typically small, and primarily driven by the extent to which new use-errors are uncovered, the sample size of summative studies depends on the study objectives and the nature of the analyses. Regardless, all user groups described in the use specification should be described during protocol development, and in many cases, it will be appropriate to recruit subgroups within each user group; for example, based on diversity and inclusion, language,



**Fig. 3 | Relationships between the elements of a use-related risk analysis.** A use-error is an action or lack of action which may result in a use-related hazard. An error that does not lead to foreseeable harm does not need to be described in the use-related risk analysis; however, such errors should be accounted for to optimize

usability. See Table 1 for definitions. The three examples illustrated in this figure relate to the end-users of sDHTs; however, use-errors apply to all sDHT users including carepartners, clinicians, researchers, and administrators.

**Table 3 | A comparison of formative and summative evaluations**

|                                   | Formative evaluations                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Summative evaluations                                                                                                |
|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|
| Purpose                           | Describe user tasks and identify use-errors<br>Inform iterative design modifications                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | Demonstrate usability amongst a representative sample under conditions reflecting the intended use or context of use |
| Activities                        | Activities such as consumer preference testing or market research<br>Institutional Review Board or ethics committee approved or exempt research studies                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                      |
| sDHT maturity                     | Prototypes with no, partial, or full functionality                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | Final or production-equivalent version<br>Marketed version                                                           |
| Example Procedures                | Cognitive walkthrough: <sup>a</sup> Study staff guide each participant through user tasks while encouraging users to think aloud or usability experts break down user tasks and identify possible use-errors<br>Contextual inquiry: Study staff observe users interacting with the sDHT in the intended use environment, asking questions during and/or after use<br>Knowledge task studies: Assessment of users' understandability and actionability<br>Simulated-use, closely mimicking conditions reflecting the intended use or context of use, with no involvement of study staff<br>Actual-use with no involvement of study staff |                                                                                                                      |
| Example usability testing methods | Interviews<br>Focus groups<br>Qualitative and quantitative surveys<br>Real-time verbalization (think-aloud)<br>Observation and/or timing of user tasks, use-errors, and recovery from use-errors<br>Captured by the sDHT itself; for example, logging use-errors or timing user tasks. Product-related considerations such as connectivity loss, app crashes, and page load times may also be informative.                                                                                                                                                                                                                              |                                                                                                                      |
| Sample                            | Typically begins with small samples, becoming larger and increasingly similar to intended users as iterative evaluations continue.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | Representative of all user groups described in the use specification.                                                |
| Setting                           | Typically begin in-lab before progressing to remote data capture.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Representative of, or generalizable to, all use environments described in the use specification.                     |
| Duration                          | Typically begin with single visits before progressing to longer evaluations.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Ideally reflecting the period of intended use, including extended durations.                                         |
| Evaluation criteria               | Typically focused on assessing use-errors rather than formal evaluation criteria.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Quantitative pass/fail criteria must be pre-specified.                                                               |

<sup>a</sup>Global regulatory guidance documents define cognitive walkthrough differently; however, both approaches are suitable procedures for formative evaluations. See Table 1 for additional definitions.

Box 4 | Principles of human-centered design

Optimizing sDHT usability involves more than just minimizing use-errors and use-related hazards; the goal is to create tools that are tailored to the user’s needs, resulting in products that are intuitive, accessible, and enjoyable to use for a sustained period within the user population of interest.

Adopting human-centered design, which prioritizes the needs, capabilities, and behaviors of users during the design process<sup>37</sup>, can result in

products with superior usability, typically associated with greater accessibility, engagement, and adherence<sup>52,70</sup>. This approach involves understanding the user perspective, designing for their needs, and collaborating with representative users as partners throughout the entire end-to-end design process. This approach is particularly important for sDHTs, where tension may exist between the technical specification and user requirements.

Table 4 | Real-world case studies in which usability validation activities have been applied

| Usability validation activities           | Case studies                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Key activity 1: Use specification         | A 2018 publication by Pillalamarri et al. <sup>53</sup> provides an overview of the approach taken during development of the Omnipod DASH Insulet Management System (Insulet Corp, Billerica MA, USA). Table 1 provides an abridged summary of representative end-users, including their characteristics, needs, and motivations for using the sDHT.                                                                                                                                                                                |
| Key activity 2: Use-related risk analysis | A 2017 publication by Preusse et al. <sup>64</sup> describes a heuristic analysis, in which one expert evaluated a commercially-available sDHT (Fitbit One; Fitbit Inc, San Francisco CA, USA) against Nielsen’s usability heuristics <sup>65</sup> , followed by a secondary review undertaken by two additional experts. Violations were presented in Table 1 along with examples and possible use-errors that may result.                                                                                                        |
| Key activity 3: Formative evaluation      | A 2020 publication by Stubberud et al. <sup>54</sup> describes the development of an sDHT smartphone app conducted over three iterative formative evaluation cycles. Usability data were captured using contextual inquiry during the first two phases, followed by a 2-week period of actual use in the intended use environment. Improvements were made to the app between each phase by incorporating user feedback.                                                                                                             |
| Key activity 4: Summative evaluation      | A 2022 publication by Domingos et al. <sup>55</sup> describes a summative evaluation of the Mi Band 2 (Xiaomi Inc, Beijing, China), in which 110 older adults used the sDHT in the intended use environment for 15 days before completing three validated usability surveys (Technology Acceptance Model; System Usability Scale; User Satisfaction Evaluation Questionnaire). Study hypotheses were clearly described, a power calculation was performed a priori, and statistical data were presented addressing each hypothesis. |

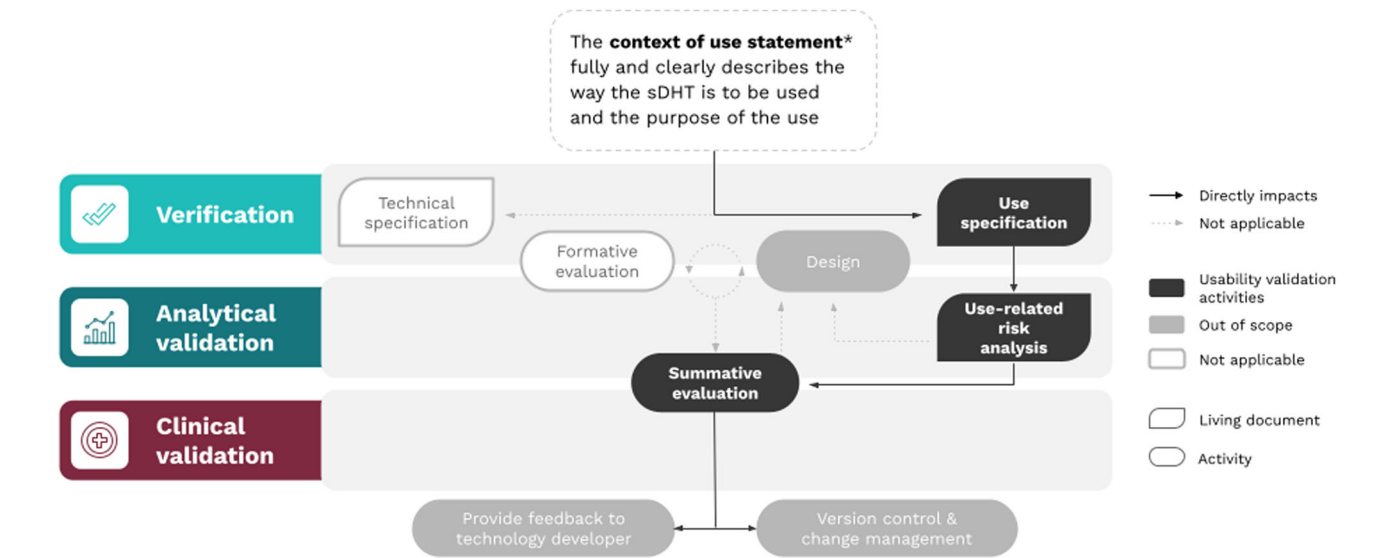


Fig. 4 | The three V3+ usability validation activities applicable to post-market sDHTs. This figure depicts the subset of V3+ usability validation activities that are applicable to post-market sDHTs, during which there is no immediate opportunity to modify the technical specification, product design, or undertake formative evaluation. See Table 1 for definitions. \*The context of use statement is a key component of the labeling of regulated medical devices. We recommend development of an equivalent statement for non-regulated sDHTs. This statement should be compared against the original intended use (or equivalent statement) of the sDHT; this gap analysis will guide subsequent activities.

educational background, disease severity, and other important determinants of usability.

Two influential papers published in the early 1990s<sup>49,50</sup> led to the “rule of thumb” that 80% of use-errors can be uncovered by assessing 5–10 participants per user group, depending on the likelihood of problem detection. While minimization of use-errors is important, sample sizes of

this magnitude do not allow for sufficient diversity to understand generalizability of study results. We, therefore, encourage investigators to take a hypothesis-driven approach to summative usability validation studies where appropriate, requiring a power calculation based on preliminary data demonstrating a meaningful effect size along with a robust statistical analysis plan. The sample size and recruitment plan developed for a quantitative



Table 5 | Summary of usability validation

|        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Who?   | Human factors engineers, user experience experts, designers, hardware/software developers, data scientists/analysts/statisticians, clinical researchers, and user representatives.                                                                                                                                                                                                                                                                                                                                            |
| What?  | Study protocol for data capture from human participants including the pre-specified pass/fail criteria, or procedural description of an activity other than a formal research study such as a consumer preference test or market research survey.<br>Study report or other documentation containing sufficient information to determine the descriptive characteristics of user groups, circumstances, and setting(s) of data capture, evaluation criteria, and descriptive statistics or statistical testing as appropriate. |
| When?  | Iteratively throughout sDHT prototype development, and following finalization of the sDHT design.                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Where? | Research or clinical laboratories and remote data capture environments.                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Why?   | To evaluate the extent to which an sDHT can be used to achieve specified goals with ease, efficiency, and user satisfaction.                                                                                                                                                                                                                                                                                                                                                                                                  |

Table 6 | Usability validation in practice

|                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |
|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Documentation you can expect               | Documentation of activities or studies should include: <ul style="list-style-type: none"><li>• Use specification; see Table 2.</li><li>• Use-related risk analysis; see Fig. 3.</li><li>• Summative study protocol(s)</li><li>• Summative study IRB/ethics committee documentation</li><li>• Study report</li><li>• White paper</li><li>• Peer-reviewed manuscript</li><li>• Regulatory submission, if applicable.</li></ul> |
| Questions answered by usability validation | Can the sDHT be used to achieve specified goals with ease, efficiency, and user satisfaction, within the stated intended use or context of use which includes a description of all intended users?                                                                                                                                                                                                                           |

Integration and alignment of V3+ with the original V3 framework

An important characteristic of the V3 framework—which also applies to V3+—is that it is modular, meaning that changes limited to one component do not necessarily require collection of new evidence within the other components. As shown in Fig. 5, which is an expanded version of Fig. 3 in ref. 8, changing the use specification is a prompt to either repeat usability validation or compile documentation demonstrating that existing usability data is generalizable to the latest use specification. Expansion to a new population, however, may prompt the need to repeat usability, analytical, and clinical validation, and in many cases, it may be possible to incorporate multiple objectives into a single study.

Future directions

The implementation of sDHTs into both clinical research and healthcare has expanded rapidly in recent years, creating the need for a common lexicon and standardized approach to usability validation that is aligned to global regulatory guidance and industry standards (see Box 5 for key takeaways). Although such guidance and standards have existed for some time, V3+ represents the first evaluation framework that is specific to sDHTs, which is of particular importance given the multidisciplinary nature of the growing field of digital medicine and the need to ensure a common understanding amongst engineering and clinical experts in addition to those with specific training in user experience. In addition to ensuring accessibility and user-centricity of sDHTs, and that appropriate scientific and clinical decisions are made on the basis of sDHT-derived data, following a pragmatic approach to usability validation offers many advantages to research study sponsors and clinicians; for example, improved workflow efficiency, reduced risk of missing data, improved clinical decision-making, and cost reduction resulting from the avoidance of unnecessary product re-design or inappropriate product selection.

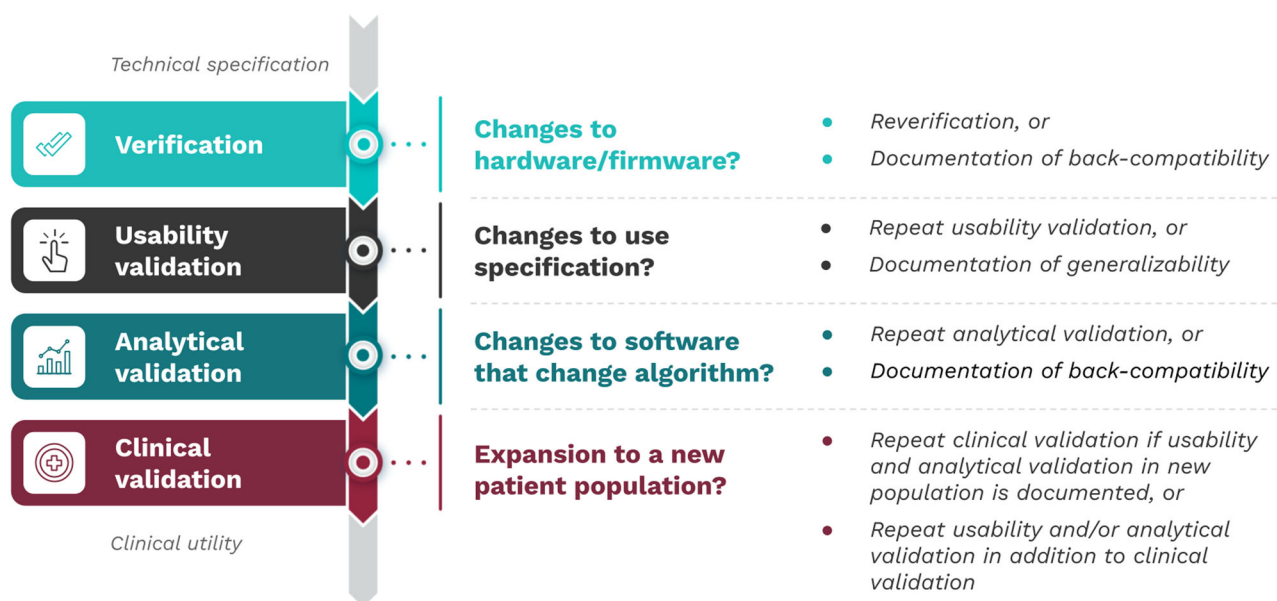
Box 5 | Key takeaways

1. The rapid proliferation of sensor-based digital health technologies (sDHTs) has led to some troubling examples of what may result from inadequate attention to usability evaluation, including compromised patient safety and extensive loss of clinical research data.
  2. The latest addition to V3—now the extended “V3+” framework—describes usability validation, a streamlined approach to evaluating the extent to which sDHTs can be used to achieve specified goals with ease, efficiency, and user satisfaction within a defined intended use or context of use.
  3. We propose and describe four key activities within usability validation: development of a use specification, implementation of a use-related risk analysis, formative evaluation of prototype products, and summative evaluation of final products demonstrating sufficient usability according to pre-specified success criteria.
  4. The approach described here is multidisciplinary, iterative, and cyclical in nature, and applicable to sDHTs both under development (pre-market) and commercially available (post-market).
  5. We urge investigators to take a hypothesis-driven approach to usability study design, and encourage the development of standardized
- usability study outcome measures tailored to the unique considerations of sDHTs.

  6. As was the case with the original V3 framework, V3+ is modular, meaning that changes to an sDHT limited to one component of the framework do not necessarily require collection of new evidence within the other components, optimizing efficient evidence generation supporting sDHTs as fit-for-purpose.
  7. The original components of V3—verification, analytical validation, and clinical validation—remain unchanged.
  8. Implementation of the usability validation component of V3+ offers benefits to end-users and their care partners, as well as clinicians, investigators, the technology sector, ethics committees or institutional review boards, data and safety monitoring boards, regulators, and payers, all of whom may rely on sDHT-generated data for decision-making purposes.

research study should account for representation and diversity in addition to confidence levels and error margins. A similarly robust approach should be taken when determining the sample size of qualitative studies, by considering additional factors such as the scope, data collection method, and expertise of the researchers involved.

The usability validation component of V3+ emphasizes the importance of identifying quantitative pre-specified pass/fail criteria, particularly during the design of summative evaluations, which can be challenging to define given the relative lack of well-established and widely adopted sDHT-specific usability study outcomes<sup>21</sup>. We, therefore, encourage the



**Fig. 5 | V3+ in action.** This figure is an expanded version of Fig. 3 in ref. 8, emphasizing the modular nature of V3+; that is, changes made within one component of the framework will require repeat testing or documentation of generalizability of that component, but may not always require repetition of the full evidence stack.

development and psychometric evaluation of standardized usability measures that address the unique considerations of sDHTs, including prolonged use in unsupervised environments and industry benchmarking, which will allow direct comparison both across tools and over time. Such measures can only become widely adopted if usability studies are made available in the public domain, ideally in the peer-reviewed literature, and we therefore encourage sDHT developers and evaluators to publish their usability research including negative studies. Similarly, DATAcc by DiMe is committed to expanding our database of case studies as an open-access resource to allow the digital medicine community to share experiences, best practices, and lessons learned<sup>51</sup>.

Although V3+ has been developed specifically for sDHTs, many of the principles and activities described in the framework are applicable to other products and workflows. As the field of digital medicine continues to mature, we anticipate the need to apply similarly streamlined approaches in emerging areas such as AI/ML applications in healthcare, and encourage other professional bodies to build on our work. For example, there is scope to explore the design process in digital health in more depth<sup>52</sup>. Our hope is that by developing recommendations for usability validation best practices and supporting their implementation, sDHT user experience will become a key product differentiator as it has for consumer electronics. When all stakeholders demand optimal sDHT usability as a critical requirement rather than a nice-to-have, investment in robust usability validation will follow, ensuring optimal care for the diverse patient populations we serve<sup>53–55</sup>.

Received: 23 February 2024; Accepted: 30 October 2024;  
Published online: 24 January 2025

## References

1. Goldsack, J. C., Dowling, A. V., Samuelson, D., Patrick-Lake, B. & Clay, I. Evaluation, acceptance, and qualification of digital measures: from proof of concept to endpoint. *Digit. Biomark.* **5**, 53–64 (2021).
2. Taguibao, C. Celebrating the fourth anniversary of the Library of Digital Endpoints! *Digital Medicine Society (DiMe)* <https://dimesociety.org/celebrating-the-fourth-anniversary-of-the-library-of-digital-endpoints/> (2023).
3. Bellerophon. A study to assess pulsed inhaled nitric oxide in subjects with pulmonary fibrosis at risk for pulmonary hypertension (REBUILD). *Clinical Trials* <https://clinicaltrials.gov/study/NCT03267108> (2023).
4. Servais, L. et al. First regulatory qualification of a digital primary endpoint to measure treatment efficacy in DMD. *Nat. Med.* **29**, 2391–2392 (2023).
5. U.S. Food and Drug Administration. Medical Device Development Tool (MDDT) summary of evidence and basis of qualification Apple atrial fibrillation history feature. U.S. Food and Drug Administration <https://www.fda.gov/media/178230/download> (2024).
6. Li, R. C., Asch, S. M. & Shah, N. H. Developing a delivery science for artificial intelligence in healthcare. *NPJ Digit. Med.* **3**, 107 (2020).
7. Marwaha, J. S., Landman, A. B., Brat, G. A., Dunn, T. & Gordon, W. J. Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation. *NPJ Digit. Med.* **5**, 13 (2022).
8. Goldsack, J. C. et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *NPJ Digit. Med.* **3**, 55 (2020).
9. Mantua, V., Arango, C., Balabanov, P. & Butlen-Ducuing, F. Digital health technologies in clinical trials for central nervous system drugs: an EU regulatory perspective. *Nat. Rev. Drug Discov.* **20**, 83–84 (2021).
10. Colloud, S. et al. Evolving regulatory perspectives on digital health technologies for medicinal product development. *NPJ Digit. Med.* **6**, 56 (2023).
11. Sacks, L. & Kunkoski, E. Digital health technology to measure drug efficacy in clinical trials for Parkinson's disease: a regulatory perspective. *J. Parkinsons Dis.* **11**, S111–S115 (2021).
12. Digital Medicine Society (DiMe). Resources in action. *Digital Medicine Society (DiMe)* <https://dimesociety.org/resources-in-action/> (2022).
13. Center for Drug Evaluation & Research. Digital Health Technologies (DHTs) for drug development. U.S. Food and Drug Administration <https://www.fda.gov/science-research/science-and-research-special-topics/digital-health-technologies-dhts-drug-development> (2023).
14. American Medical Association. Future of health- commercial payer coverage for digital medicine codes. American Medical Association <https://www.ama-assn.org/system/files/issue-brief-commercial-payer-coverage-digital-care.pdf> (2023).
15. Kang, H. S. & Exworthy, M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. *JMIR mHealth uHealth* **10**, e35684 (2022).

16. Adedinsawo, D. et al. Health disparities, clinical trials, and the digital divide. *Mayo Clin. Proc.* **98**, 1875–1887 (2023).
17. Landers, M., Dorsey, R. & Saria, S. Digital endpoints: definition, benefits, and current barriers in accelerating development and adoption. *Digit. Biomark.* **5**, 216–223 (2021).
18. Adams, J. L. et al. Using a smartwatch and smartphone to assess early Parkinson's disease in the WATCH-PD study. *NPJ Parkinson's Dis.* **9**, 64 (2023).
19. U.S. Food and Drug Administration. Class 2 device recall. U.S. Food and Drug Administration <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfres/res.cfm?id=43539> (2006).
20. Center for Drug Evaluation & Research. Digital health technologies for remote data acquisition in clinical investigations. U.S. Food and Drug Administration <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/digital-health-technologies-remote-data-acquisition-clinical-investigations> (2023).
21. Tandon, A. et al. Human Factors, Human-Centered Design, and Usability of Sensor-Based Digital Health Technologies: Scoping Review. *J. Med. Internet Res.* **26**, e57628 (2024).
22. Center for Devices & Radiological Health. General wellness: Policy for low risk devices—guidance for industry and FDA staff. U.S. Food and Drug Administration <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices> (2019).
23. U.S. Food and Drug Administration. Premarket approval (PMA). U.S. Food and Drug Administration <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P160048S026> (2023).
24. Center for Devices & Radiological Health. Ingestible telemetric gastrointestinal capsule imaging system - final class II special controls guidance document for industry and FDA. U.S. Food and Drug Administration <https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/ingestible-telemetric-gastrointestinal-capsule-imaging-system-final-class-ii-special-controls> (2021).
25. U.S. Food and Drug Administration. Device classification under section 513(f)(2)(De Novo). U.S. Food and Drug Administration <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/denovo.cfm?id=DEN200011> (2023).
26. U.S. Food and Drug Administration. 510(k) premarket notification. U.S. Food and Drug Administration <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?id=K150557> (2015).
27. Center for Devices & Radiological Health. Applying human factors and usability engineering to medical devices. U.S. Food and Drug Administration <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices> (2016).
28. U.S. Food and Drug Administration. Human factors studies and related clinical study considerations in combination product design and development- draft guidance for industry and FDA staff. U.S. Food and Drug Administration <https://www.fda.gov/media/96018/download> (2016).
29. Center for Drug Evaluation & Research. Comparative analyses and related comparative use human factors studies for a drug-device combination product submitted in an ANDA: Draft guidance for industry. U.S. Food and Drug Administration <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/comparative-analyses-and-related-comparative-use-human-factors-studies-drug-device-combination> (2017).
30. European Medicines Agency. Quality documentation for medicinal products when used with a medical device—scientific guideline. European Medicines Agency <https://www.ema.europa.eu/en/quality-documentation-medicinal-products-when-used-medical-device-scientific-guideline> (2022).
31. Department of Health and Social Care. Guidance on applying human factors and usability engineering to medical devices including drug-device combination products in Great Britain. Medicines, Healthcare products Regulatory Agency [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/970563/Human-Factors\\_Medical-Devices\\_v2.0.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/970563/Human-Factors_Medical-Devices_v2.0.pdf) (2021).
32. National Medical Products Association (NMPA). Guiding principles for technical review of human factors design of medical devices [DiMe, Trans.]. Digital Medicine Society (DiMe) <https://datacc.dimesociety.org/wp-content/uploads/2023/09/NMPA-Human-Factors-Guidance-English-Translation-FINAL.pdf> (2023).
33. The European Parliament and the Council of the European Union. Regulation—2017/745—EN—medical device regulation. EUR-Lex <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745> (2017).
34. Center for Devices & Radiological Health. Purpose and content of use-related risk analyses for drugs, biological products, and combination products. U.S. Food and Drug Administration <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/purpose-and-content-use-related-risk-analyses-drugs-biological-products-and-combination-products> (2024).
35. International Organization for Standardization. IEC 62366-1:2015. ISO <https://www.iso.org/standard/63179.html> (2021).
36. International Organization for Standardization. IEC/TR 62366-2:2016. ISO <https://www.iso.org/standard/69126.html> (2024).
37. International Organization for Standardization. ISO 9241-210:2019. ISO <https://www.iso.org/standard/77520.html> (2019).
38. International Organization for Standardization. ISO 20282-1:2006. ISO <https://www.iso.org/standard/34122.html> (2020).
39. International Organization for Standardization. ISO/TS 20282-2:2013. ISO <https://www.iso.org/standard/62733.html> (2024).
40. Digital Medicine Society (DiMe). Library of Human Factors Resources for Digital Health Technologies. DATAcc by DiMe <https://datacc.dimesociety.org/resources/interactive-index-of-human-factors-resources-for-digital-health-technologies-dhts-2/> (2023).
41. International Organization for Standardization. ISO 14971:2019. ISO <https://www.iso.org/standard/72704.html> (2019).
42. Zhou, Y. et al. Missing data matter: an empirical evaluation of the impacts of missing EHR data in comparative effectiveness research. *J. Am. Med. Inform. Assoc.* **30**, 1246–1256 (2023).
43. Flora, C. et al. High-frequency temperature monitoring for early detection of febrile adverse events in patients with cancer. *Cancer Cell* **39**, 1167–1168 (2021).
44. Office of Regulatory Affairs. Design controls. U.S. Food and Drug Administration <https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/inspection-guides/design-controls> (2023).
45. Bakker, J. P. et al. Regulatory pathways for qualification and acceptance of digital health technology-derived clinical trial endpoints: considerations for sponsors. *Clin. Pharmacol. Ther.* <https://doi.org/10.1002/cpt.3398> (2024).
46. Schaefer, S. E., Van Loan, M. & German, J. B. A feasibility study of wearable activity monitors for pre-adolescent school-age children. *Prev. Chronic Dis.* **11**, E85 (2014).
47. Davis, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* <https://doi.org/10.2307/249008> (1989).
48. Shachak, A., Kuziemsky, C. & Petersen, C. Beyond TAM and UTAUT: future directions for HIT implementation research. *J. Biomed. Inform.* **100**, 103315 (2019).
49. Virzi, R. A. Refining the test phase of usability evaluation: how many subjects is enough? *Hum. Factors* <https://doi.org/10.1177/001872089203400407> (1992).
50. Lewis, J. R. Sample sizes for usability studies: additional considerations. *Hum. Factors* **36**, 368–378 (1994).
51. Digital Medicine Society (DiMe). DATAcc by DiMe Resources. DATAcc by DiMe <https://datacc.dimesociety.org/resources/> (2024).
52. Duffy, A., Christie, G. J. & Moreno, S. The challenges toward real-world implementation of digital health design approaches: narrative review. *JMIR Hum. Factors* **9**, e35693 (2022).



53. Pillalamarri, S. S., Huyett, L. M. & Abdel-Malek, A. Novel bluetooth-enabled tubeless insulin pump: a user experience design approach for a connected digital diabetes management platform. *J. Diabetes Sci. Technol.* **12**, 1132–1142 (2018).
54. Stubberud, A., Tronvik, E., Olsen, A., Gravdahl, G. & Linde, M. Biofeedback treatment app for pediatric migraine: development and usability study. *Headache* **60**, 889–901 (2020).
55. Domingos, C., Costa, P., Santos, N. C. & Pêgo, J. M. Usability, acceptability, and satisfaction of a wearable activity tracker in older adults: observational study in a real-life context in northern Portugal. *J. Med. Internet Res.* **24**, e26652 (2022).
56. Shoemaker, S. J., Wolf, M. S. & Brach, C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ. Couns.* **96**, 395–403 (2014).
57. FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource* <https://www.ncbi.nlm.nih.gov/books/NBK326791/> (National Library of Medicine, 2016).
58. FDA-NIH Biomarker Working Group. Glossary. in *BEST (Biomarkers, EndpointS, and other Tools) Resource* <https://www.ncbi.nlm.nih.gov/books/NBK338448/#IX-C> (U.S. Food and Drug Administration and National Institutes of Health, 2021).
59. Nielsen, J. Usability 101: Introduction to Usability. Nielsen Norman Group <https://www.nngroup.com/articles/usability-101-introduction-to-usability/> (2012).
60. U.S. Food and Drug Administration. Glossary of terms on clinical trials for patient engagement advisory committee meeting. U.S. Food and Drug Administration <https://fda.report/media/108378/Clinical-Trial-Glossary-Terms.pdf> (2017).
61. Center for Devices & Radiological Health. General/specific intended use—guidance for industry. U.S. Food and Drug Administration <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/generalspecific-intended-use-guidance-industry> (2020).
62. International Organization for Standardization. ISO 13485:2016. *ISO* <https://www.iso.org/standard/59752.html> (2020).
63. Schnall, R., Higgins, T., Brown, W., Carballo-Dieguez, A. & Bakken, S. Trust, perceived risk, perceived ease of use and perceived usefulness as factors related to mHealth technology use. *Stud. Health Technol. Inform.* **216**, 467 (2015).
64. Preusse, K. C., Mitzner, T. L., Fausset, C. B. & Rogers, W. A. Older adults' acceptance of activity trackers. *J. Appl. Gerontol.* **36**, 127–155 (2017).
65. Nielsen, J. *Usability Inspection Methods* Nielsen Norman Group <https://www.nngroup.com/books/usability-inspection-methods/> (John Wiley & Sons, 1994).
66. IVD Clinical Evidence Working Group. Framework for developing clinical evidence for regulatory and coverage assessments in in vitro diagnostics (IVDs). *MDIC* <https://mdic.org/resource/framework-for-developing-clinical-evidence-for-regulatory-and-coverage-assessments-in-in-vitro-diagnostics-ivds/> (2019).
67. Fryback, D. G. & Thornbury, J. R. The efficacy of diagnostic imaging. *Med. Decis. Mak.* **11**, 88–94 (1991).
68. Haddow, J. E., Palomaki, G. E., Haddow, J. & Palomaki, G. E. ACCE: a model process for evaluating data on emerging genetic tests. <https://www.scienceopen.com/document?vid=397ab49d-0b83-45cc-9ad0-c45b20c833e6> (2003).
69. National Academies of Sciences, Engineering, and Medicine. *An Evidence Framework for Genetic Testing* <https://doi.org/10.17226/24632> (The National Academies Press, 2017).
70. Henni, S. H., Maurud, S., Fuglerud, K. S. & Moen, A. The experiences, needs and barriers of people with impairments related to usability and accessibility of digital health solutions, levels of involvement in the design process and strategies for participatory and universal design: a scoping review. *BMC Public Health* **22**, 35 (2022).

## Acknowledgements

The authors wish to acknowledge Bethanie McCrary and Danielle Stefko for assistance with project management. On November 9th, 2023, DATAcc hosted a multistakeholder workshop in order to gather and incorporate expert feedback on the content presented in this manuscript; we are grateful to the 29 individuals who participated in refining and improving our work.

## Author contributions

J.P.B., R.B., J.C., B.C., C.C., C.C.G., B.H., N.H.M., E.S.I., N.V.M., S.Mc., S.Mo., S.P., O.P., M.S., E.S., T.S., A.T., W.vd.B., S.V., B.Va., B.Vr., and J.C.G. contributed to the conceptual development of the V3+ framework and the drafting of this manuscript; all authors have read and approved the final submitted version, and agree to be accountable for all aspects of the work presented here.

## Competing interests

J.P.B. is a former employee of Philips and Signifier Medical Technologies and holds company stock or stock options, and reports consulting income from Apnimed and Koneksa Health. R.B. is an employee of Regeneron Pharmaceuticals Inc. and holds company stock or stock options. B.C. is an employee of Genentech and holds company stock. C.C. is an employee of Stel Life Inc. C.C.G. is an employee of ActiGraph LLC. B.H. is a former employee of Johnson & Johnson Innovative Medicine and holds company stock or stock options. N.H.M. is an employee of Exponent Inc. and holds company stock or stock options. E.S.I. is an employee of Koneksa health and may own company stock. N.V.M. is an employee of Johnson & Johnson Innovative Medicine and holds company stocks or stock options. S.Mo. is an employee of Synnav Healthcare. O.P. is an employee of AARDEX Group. E.S. is a member of the NPJ Digital Medicine editorial board. T.S. is an employee of Genentech and owner of shares of Roche stock. A.T. is a paid consultant for Synergen Health Technologies; Siemens Healthineers; Gabi Smartcare; and Medtronic (relationship is concluded). S.V. is an employee of Johnson & Johnson Innovative Medicine and holds company stock or stock options. B.Va. is a former employee of Byteflies and holds company stock or stock options. B.Vr. is an employee of AARDEX Group and holds company stock or stock options. The remaining authors declare no competing interests: J.C., S.Mc., S.P., M.S., W.vd.B., J.C.G.

## Additional information

**Correspondence** and requests for materials should be addressed to Jessie P. Bakker.

**Reprints and permissions 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025



<sup>1</sup>Digital Medicine Society, Boston, MA, USA. <sup>2</sup>Division of Sleep and Circadian Disorders, Mass General Brigham, Boston, MA, USA. <sup>3</sup>Division of Sleep Medicine, Harvard Medical School, Boston, MA, USA. <sup>4</sup>Regeneron Pharmaceuticals Inc, New York, NY, USA. <sup>5</sup>Genentech, South San Francisco, CA, USA. <sup>6</sup>Stel Life Inc, Philadelphia, PA, USA. <sup>7</sup>ActiGraph LLC, Pensacola, FL, USA. <sup>8</sup>Johnson & Johnson Innovative Medicine, Breda, The Netherlands. <sup>9</sup>Exponent Inc, New York, NY, USA. <sup>10</sup>Koneksa Health, New York, NY, USA. <sup>11</sup>Johnson & Johnson Innovative Medicine, Beerse, Belgium. <sup>12</sup>Sysnav Healthcare, Vernon, France. <sup>13</sup>AARDEX Group, Liège, Belgium. <sup>14</sup>Netherlands Organisation for Applied Scientific Research (TNO), Leiden, The Netherlands. <sup>15</sup>Nationwide Children's Hospital, Columbus, OH, USA. <sup>16</sup>Department of Heart, Vascular, and Thoracic, Division of Cardiology and Cardiovascular Medicine, Children's Institute, Cleveland Clinic Children's, Cleveland, OH, USA. <sup>17</sup>Cleveland Clinic Children's Center for Artificial Intelligence (C4AI), Department of Heart, Vascular, and Thoracic, Children's Institute, Cleveland Clinic Children's, Cleveland, OH, USA. <sup>18</sup>Department of Pediatrics, Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, OH, USA. <sup>19</sup>Department of Biomedical Engineering, Cleveland Clinic Lerner Research Institute, Cleveland, OH, USA. <sup>20</sup>Johnson & Johnson Innovative Medicine, New Brunswick, NJ, USA. <sup>21</sup>Department of Electrical, Computer, and Systems Engineering, Case Western Reserve University, Cleveland, OH, USA. <sup>22</sup>Department of Public Health, Liège University, Liège, Belgium. ✉ e-mail: [jessie.bakker@dimedociety.org](mailto:jessie.bakker@dimedociety.org)