U.S. FOOD & DRUG ADMINISTRATION

Key Considerations for the Development and Use of Digitally Derived Measures for Clinical Investigations

August 2026

Table of Contents

Executive Summary

Digital health technologies (DHTs), including those enabled by artificial intelligence (AI), have the potential to capture information about a person's health, continuously in real time, outside of health care settings, for a variety of purposes. For the purposes of this paper, we refer to measures derived from data collected using DHTs as digitally derived measures (DDMs).

This paper highlights key considerations drawing from existing U.S. Food and Drug Administration (FDA) guidances to support the development and use of DDMs as outcomes in clinical investigations. DDMs may be used as clinical outcome assessments (COAs), biomarkers, or as part of multicomponent endpoints derived from multimodal data. FDA encourages the assessment of clinical outcomes that are both clinically relevant and capture what is meaningful to patients.[1] When considering incorporating a DDM as an outcome in a clinical investigation, it can be helpful to start with a justification, which may be updated over time, to identify what evidence will be needed to support the DDM selection, relevance, and validity.

In the context of clinical investigations, a DHT used to generate a DDM should be verified and validated to be considered fit for purpose.[2] Evidence for validation should demonstrate that users understand and can follow the instructions for use. When using a DDM as an outcome in a clinical investigation, it is important to identify potential sources of error and factors that might negatively impact the validity of the DDM. It is also important to involve patients, caregivers, and clinicians in the development process to determine meaningfulness and clinical relevance of the DDM. This paper identifies key considerations from FDA guidances that may facilitate the development of DDMs that quantify meaningful aspects of health, consistent with best practices for patient engagement, measurement science, and software and AI verification and validation. This paper also highlights FDA's continued focus on a risk-based approach to facilitate innovations in clinical investigations and medical product development by tailoring evidence generation commensurate with the intended use of the DDM. FDA believes a collaborative, patient-centered approach may help unlock the full potential of DHTs to support medical product development and innovation.

[1] See FDA guidances Patient Engagement in the Design and Conduct of Medical Device Clinical Studies and Principles for Selecting, Developing, Modifying, and Adapting Patient-Reported Outcome Instruments for Use in Medical Device Evaluation.

[2] See FDA guidance Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.

Introduction

A digital health technology (DHT) is a system that uses computing platforms, connectivity, software, and/or sensors for health care and related uses. These technologies span a wide range of uses, from applications including non-device general wellness functions to medical device functions.[3] DHTs, including those enabled by artificial intelligence (AI), have the potential to capture information about a person's health, continuously in real time, outside of health care settings, for a variety of purposes, including clinical research and clinical care.[4]

As a result, DHTs have the potential to bring research and clinical care to participants and patients where they are located and at various stages along their health journey. The ability to collect continuous data outside of the health care setting may help detect important changes in health over time that may not be evident from data collected only during visits to research or clinical care locations.

For the purposes of this paper, digitally derived measures (DDMs) refer to measures derived from data collected using DHTs. DDMs leverage technology to advance health assessments and may span a wide range of categories. DDMs have the potential to facilitate early disease detection, support treatment, and enhance patient care outside of the clinic. DDMs may be more reflective of a patient's health than traditional measures used in clinical research. DDMs also have the potential to streamline clinical investigations through earlier detection of treatment effects or safety signals. DDMs may potentially be used as clinical outcome assessments, biomarkers, or as part of multicomponent endpoints derived from multimodal data. Additionally, DDMs may support the implementation of decentralized trials.

Aligned with FDA's mission of protecting, promoting, and advancing public health, and building on the Agency's longstanding commitment to support innovations in the development and regulation of medical products, FDA's Center for Biologics Evaluation and Research (CBER), Center for Drug Evaluation and Research (CDER), Center for Devices and Radiological Health (CDRH), and Oncology Center of Excellence (OCE) are jointly publishing this paper to highlight key considerations drawn from existing FDA guidances to support the development and use of DDMs as outcomes in clinical investigations. This paper reflects FDA's continued focus on a risk-based approach to facilitate responsible innovations in clinical investigations and medical product development by tailoring evidence generation commensurate with the intended use of the DDM. This paper is not intended to cover the entire process for developing a DDM nor is it intended to address all use cases for DHTs in clinical care and research. DDM developers should refer to all relevant FDA guidances. Sponsors should engage with the relevant FDA review division early and prior to incorporating a DDM in a clinical investigation for medical product development.

[3] To the extent that a DHT meets the definition of a device under section 201(h) of the Federal Food, Drug and Cosmetic Act, the device will be subject to applicable legal requirements for devices. FDA's guidance documents include policies regarding DHTs used in clinical investigations. See e.g. FDA guidance Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.

[4] See FDA guidance Digital Health Technologies for Remote Data Acquisition in Clinical Investigations. Also see FDA Digital Health and Artificial Intelligence Glossary.

Measuring What Is Meaningful to Patients

FDA encourages the use of clinical outcomes that are both clinically relevant and capture what is meaningful to patients.[5] FDA has described a meaningful aspect of health as any specific aspect of feeling or functioning in daily life that is important to patients.[6] Meaningful aspects of health that are captured by DDMs derived from DHTs are important to be identified early and to continue to be reevaluated as the DDM is refined. At the early stage, it is critical to define the context of use, especially the target population, i.e., those intended to use or benefit from the medical product. Attributes of the target population, such as age, sex, and relevant co-morbidities, are important to be considered during DDM development. These factors may influence how a DHT performs and its usability as well as what constitutes a meaningful change as defined by the DDM when used as an outcome. To support the characterization of the target population, it may be useful to develop a disease model that represents how a disease or condition affects bodily structures and/or processes and all the resulting effects on patients in terms of meaningful aspects of health.[7] Developers are encouraged to consult relevant FDA guidance when characterizing the target population.[8]

When considering whether to use a DHT to capture a DDM, it is important to evaluate the potential value and burden to patients, and caregivers as applicable, in the target population. Determining the meaningful aspect of health is not exclusive to clinicians and regulators; patients and caregivers can be helpful in this determination. In addition to input from clinical experts, qualitative data from patients and/or caregivers play an essential role in ensuring that the meaningful aspects of health associated with the disease or condition are identified and clearly described.[9]

[5] See FDA guidances Patient Engagement in the Design and Conduct of Medical Device Clinical Studies and Principles for Selecting, Developing, Modifying, and Adapting Patient-Reported Outcome Instruments for Use in Medical Device Evaluation.

[6] See FDA guidances Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments and Principles for Selecting, Developing, Modifying, and Adapting Patient-Reported Outcome Instruments for Use in Medical Device Evaluation.

[7] See FDA guidance Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.

[8] See FDA guidance Principles for Selecting, Developing, Modifying, and Adapting Patient-Reported Outcome Instruments for Use in Medical Device Evaluation.

[9] See FDA guidance Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments.

DHT and DDM Selection and Evidence

The concept of interest is the aspect of an individual's clinical, biological, physical, or functional state, or experience that the assessment is intended to capture (or reflect).[10] The concept of interest may be any aspect of health that might be affected, positively or negatively, by the medical product.[11] Note that the meaningful aspect of health is the focus for the measurement and may or may not be synonymous with the concept of interest. Drawing from the Patient-Focused Drug Development (PFDD) methodological guidance series, the meaningful aspect of health and concept of interest should be clearly defined.[12] Once the meaningful aspect of health, concept of interest, and context of use have been considered, it is important to consider the technical evidence for the verification and validation of the DHT and the validity evidence for the DDM when selecting these to integrate into the clinical investigation. When incorporating a DDM as an outcome, it can be helpful to start with a justification for the relationship between the DDM and the meaningful aspects of health to identify what evidence will be needed to support the DDM selection, relevance, and validity. This justification may be updated over time.

Some have suggested that using a conceptual model to help identify and convey the concept of interest may help ensure that specific health experiences are captured.[13] For example, it may be important to identify and define which aspects of walking (e.g., bout frequency, bout definition, bout length, and gait speed) are relevant in the context of the disease or condition being studied. Conceptual clarity may help to ensure that the DDM aligns with the goals of the measurements and helps identify why that concept is clinically relevant to the specific target patient population. Conceptual clarity also helps articulate key components of the DDM, such as those aspects of walking described above and their relevance to the target patient population.

A clear justification for the DDM selected may include delineation of any possible advantages of the selected DDM over other types of COAs or biomarkers. When building the justification, it may be helpful to consider incorporating evidence from multiple sources to support the relevance of the values generated by the DDM.

[10] See BEST (Biomarkers, EndpointS, and other Tools) Resource.

[11] See also FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making.

[12] See FDA guidance Principles for Selecting, Developing, Modifying, and Adapting Patient-Reported Outcome Instruments for Use in Medical Device Evaluation. See also FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making.

[13] Manta, C, B Patrick-Lake, and JC Goldsack, 2020, Digital Measures That Matter to Patients: A Framework To Guide the Selection and Development of Digital Measures of Health, Digit Biomark, 4(3): 69–77.

If comparing the DDM to a reference measure, it is important to understand the context of the comparison and, for example, whether the reference measure and the DDM attempt to capture the same constructs.[14] However, there may be no appropriate reference measure for direct comparison to novel concepts. In this case, it may be helpful to discuss how to support the rationale and validity of the DDM with the relevant FDA review division. In the walking bout example, the rationale would articulate why bout length versus total step count alone is a better measure to capture the ambulatory function relevant to the patients with cardiovascular disease.

When selecting a DHT that generates a DDM, both may be considered fit-for-purpose when they have sufficient evidence of validity and are found to be appropriate for the context of use.[15] Specifically, the target population, the clinical study design in which the DDM is to be used, the measurement schedule (e.g., duration of wearing a sensor, timeframe of data collection from ecological momentary assessment), and the clinical event or characteristic of the disease or condition of interest to be measured can be considered. For DDMs that use AI algorithms, developers might consider whether to include a credibility assessment commensurate with model risk specific to the proposed context of use in the fit-for-purpose determination. This is consistent with the FDA's proposed risk-based credibility assessment framework for AI/ML models used to produce information or data to support regulatory decision-making for drugs and biologics described in the draft guidance Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products.[16]

FDA's guidance for industry, FDA staff, and other stakeholders titled Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments, outlines eight evidence-based considerations that should be considered for inclusion in the rationale and supporting evidence or justification that a clinical outcome measure is appropriate in a particular context of use.[17] This includes evidence of the interpretability of the DDM in the clinical investigation.[18] While these principles for clinical outcome assessments and those related to biomarker development may generally be applicable to DDMs, not all evidence-based considerations may be relevant, and their applicability may also depend on the type of DHT being used to generate the DDM. Furthermore, there may be specific considerations unique to DHTs that may be worth discussing with the Agency.

[14] See FDA guidance Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments.

[15] See FDA guidance Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.

[16] See FDA's Artificial Intelligence for Drug Development.

[17] See FDA guidance Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments.

[18] See FDA guidances Principles for Selecting, Developing, Modifying, and Adapting Patient-Reported Outcome Instruments for Use in Medical Device Evaluation; Patient-Focused Drug Development: Collecting Comprehensive and Representative Input; and Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments.

Verification and Validation Considerations

In the context of clinical investigations, a DHT used to generate a DDM should be verified and validated to be considered fit-for-purpose.[19] From a DHT perspective, verification is confirmation by an examination and provision of objective evidence that the parameter that defines the DDM (e.g., acceleration, temperature, pressure) is measured accurately and precisely. Validation of a DDM consists of two major parts – analytical validation and clinical validation.[20] Analytical validation is confirmation by examination and provision of objective evidence that the selected DHT appropriately assesses the clinical event or characteristic in the proposed target population (e.g., physical functioning, atrial fibrillation). Clinical validation confirms that the DDM measures what it purports to measure at the conceptual level, and it reflects a meaningful aspect of health and tracks changes in the patient's disease or clinical status.

To illustrate these steps, this paper uses a recurring example of a DDM based on walking bout length in patients with cardiovascular disease. This example is carried through the verification and validation sections that follow. In this example, bout length is derived from a DHT that measures steps. The DHT measures accelerations that are then converted into step count. The acceleration measurements and any associated AI algorithm are verified. The step count derived from those accelerations is analytically validated against a reference measure in the target patient population. Walking bouts are then derived from step count, using a defined algorithm to identify and characterize bouts of continuous walking. The resulting DDM of walking bout length is then clinically validated in the context of use, including determining the clinical relevance and meaningfulness to patients with cardiovascular disease. The following sections describe each of these steps in more detail.

Verification

The verification step builds evidence that the DHT accurately and precisely measures the physical or chemical parameters that it purports to measure. In the walking bout example, this means confirming that the accelerometer accurately captures acceleration values that will later be converted to step count. The verification process includes evaluating the performance and accuracy of the DHT to ensure it functions properly and provides reliable data. During the design and development of the DHT, verification is a key quality assurance technique that involves tests and analyses. It is helpful for sponsors to select and apply appropriate verification techniques based on the generally accepted practices for the technologies. Leveraging verification data made available by DHT manufacturers or other third parties either publicly or by right of reference, including where applicable in device labeling, may be appropriate. Verification can include demonstrating conformance to consensus performance standards where they exist. When using DHTs in clinical investigations, engaging with DHT manufacturers, patients, caregivers, and other technical and clinical experts as part of the DHT verification process may also be appropriate. FDA guidances on digital health such as those specific to software and AI may be helpful to DDM developers to support verification and validation.[21]

[19] See FDA guidance Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.

[20] Resources like this paper can be helpful to provide additional information on analytical and clinical validation and how it relates to digitally derived measures. Bakker, J, Barge, R, Centra, A, et al. V3+ extends the V3 framework to ensure user-centricity and scalability of sensor-based digital health technologies. NJP Digital Med, 8 (51).

[21] See FDA's Guidances with Digital Health Content.

Validation

Validation of a DDM consists of two parts – analytical validation and clinical validation, as defined above.[22]

Analytical Validation

For many DHTs, multiple parameters may require analytical validation. For example, a DDM for 50th percentile of stride length derived from steps taken over a specified duration (e.g. one week as an illustrative example) requires evidence that the DHT measures both strides and stride length accurately and precisely in the target population.

In this example, the first step in analytically validating the stride length, as a component of bout length, may be to confirm the ability of the selected DHT to detect strides and measure stride length in the target population. To determine what evidence is needed to show that the DHT captures strides and stride parameters, the developer might consider what is known about how strides are detected and measured in a healthy population, how the disease may impact how a patient takes a stride compared to a comparable healthy population, and how the DHT detects strides and measures a stride length from the raw sensor signal. Where validation studies are warranted, the DHT developer might consider gathering evidence that strides and stride parameters correspond to a set of reference measures, if available.[23]

The need for and scope of validation studies in the target population is determined based on the current state of evidence and the intended use of the DDM, e.g. context of use, role in the clinical investigation. For example, a DDM serving as a primary endpoint in a pivotal study is expected to have prospective validation with pre-specified performance thresholds in the intended target population, whereas a DDM used as an exploratory endpoint may rely on retrospective or bridging evidence.[24]

[22] See BEST (Biomarkers, EndpointS, and other Tools) Resource.

[23] Bakker, JP, SJ McClenahan, P Fromy, et al, 2025, A Hierarchical Framework for Selecting Reference Measures for the Analytical Validation of Sensor-Based Digital Health Technologies, J Med Internet Res, 27:e58956.

[24] See FDA guidance Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments.

When a reference measure measures the same concept, the new DDM may be compared to the reference measure using tools of correspondence analysis appropriate for the scale of measurement (e.g., continuous, categorical). These tools provide evidence whether the two measurement modalities capture the same construct and help to evaluate the accuracy of the new DDM relative to the reference measure. In other cases, analytical validation may involve generating evidence that a DDM captures a similar aspect of health that is best represented by a set of reference measures. As noted previously, there may be no reference measure for direct comparison to novel concepts. In this case, it is helpful to discuss how to best demonstrate analytical validity with the relevant FDA review division.

The impact of missing data and variable-quality data may be considered during analytical validation. It is helpful for sponsors to assess how much data is needed to provide reliable estimates of the DDM. For example, how many minutes of data or time spent walking are needed to estimate a percentile of the stride length distribution, which in turn informs the precision of the walking bout length DDM with a known level of uncertainty. The determination of the length of recording may depend on multiple factors, e.g., which percentile of the distribution is of interest, and the precision of the measurement. Also, it is helpful to assess the impact of missing time periods and study dropout on the DDM estimations.

Clinical Validation

Evidence of clinical validity reflects that the DDM acceptably identifies, measures, or predicts the concept of interest.[25] A DDM intended to capture a clinical event of interest can demonstrate that it:

- reflects a patient's, caregiver's, or clinician's impression of the disease severity or status,
- reliably measures disease status when the patient, caregiver or clinician does not report any changes,
- distinguishes between groups of known disease status, or
- captures changes in disease status when reported by the patient, caregiver or clinician.

In the walking bout example, walking bout length is distinguished between patients with varying degrees of cardiovascular disease severity and capture changes in ambulatory function when reported by the patient or clinician. The values generated from the DDM then correspond to the specific health experiences of the patient and relate to the meaningful aspects of health as defined in the context of use.

[25] See BEST (Biomarkers, EndpointS, and other Tools) Resource.

Role of Usability Studies

Evidence for validation should demonstrate that users understand and can follow the instructions for use, typically assessed through human factors or usability studies.[26] The rigor and scope of usability studies are commensurate with the regulatory role of the DDM and the characteristics of the intended population. Usability evaluations can be used to determine if the design of the DHT allows participants in the clinical investigation to use the DHT as directed in the trial protocol, and efforts are made to ensure the DHT is used accordingly to avoid inaccurate or imprecise measurements. In addition, the tolerability of the DHT for the target population is helpful to be taken into account. User interfaces accommodate anticipated sensory, cognitive, language, and motor capabilities of the target population. This is particularly important for populations such as children, older adults, and patients with impairments.[27] For example, DHTs should be user-friendly for their intended purposes, require minimal effort to complete tasks, have a low burden of data entry, and provide options for user control preferences when appropriate.[28] For DHTs intended for prolonged or repeated use, sponsors should evaluate adherence, user burden, sustained usability, and potential user fatigue throughout the study duration.

When applicable, calibration processes should be evaluated as well. Certain DHTs may require calibration by the user or caregiver, with or without assistance by clinical trial personnel (e.g., calibrating a wearable accelerometer for individual stride length to support accurate computation of walking bout). The calibration process should be validated to ensure accurate and precise measurements of the clinical characteristic or event of interest, and the appropriate frequency of calibration should be determined.

Cognitive interviews with patients regarding device instructions combined with pilot testing can confirm whether patients or caregivers understand the instructions and have appropriate training and support to use the DHT as intended. If the DHT may be used in a remote setting, a comparative evaluation between measurements obtained remotely and those obtained in a clinical setting using the same DHT should be conducted, and any differences should be justified.

[26] See FDA guidance Applying Human Factors and Usability Engineering to Medical Devices.

[27] Kruizinga, MD, F Stuurman, V Exadaktylos, et al, 2020, Development of Novel, Value-Based, Digital Endpoints for Clinical Trials: A Structured Approach Toward Fit-for-Purpose Validation, Pharmacol Rev, 72(4):899–909.

[28] Mathews, SC, MJ McShea, CL Hanley, et al, 2019, Digital Health: A Path to Validation, NPJ Digital Med, 2:38.

Potential Sources of Error

When using a DDM as an outcome in a clinical investigation, it is important to identify potential sources of error and factors that might negatively impact the validity of the DDM. This process might begin with an understanding of the sensors and algorithms and how they work to determine what types of experimental conditions are needed to characterize the errors.

When using a DHT to capture a DDM, the measurement may be influenced by multiple factors including inherent noise in the DHT sensor, use errors, interactions between the user and the DHT (e.g., feedback provided by the DHT to the user) and environmental factors that may change between the home, clinic, and other locations. Other examples include incorrect placement of a wearable DHT (e.g., wrist versus hip), interference with the measurement that may be misinterpreted as the clinical event or characteristic of interest (e.g., riding in a vehicle misclassified as a walking bout), changes to the DHT and software versions (e.g., alterations in relevant algorithms), and environmental characteristics (e.g., temperature or humidity). Physical factors could impact measurement performance, such as user posture, motion, or sensor-to-heart height difference (e.g., for blood pressure).

Changes in the values generated by the DDM over time, even in the absence of an intervention or change in disease activity, may also be a source of error. Considerations of this variability and identification of any associated trends may help determine whether an observed change in the values generated by a DDM is attributable to a change in disease state.

If a DHT or associated technology, such as a general computing platform or AI algorithm, is updated during a clinical investigation (e.g., operating system updates), sponsors and other relevant parties should ensure that the DHT remains fit-for-purpose and that updates do not affect the DDMs. In situations where the DDMs may be affected, it may be helpful to validate the measurements (e.g., using previously collected data or a new prospective study) after introduction of the update to ensure that no changes to the measurements occurred.

Conclusion

FDA's medical product Centers developed this paper to reaffirm our commitment to support the development and use of DDMs as outcomes in clinical investigations. This paper highlights FDA's continued focus on a risk-based approach to facilitate innovations in clinical investigations and medical product development by tailoring evidence expectations to the intended use of the DDM. This commitment aligns with our mission to ensure patient access to medical products that are safe and effective for their intended uses and underscores our dedication to facilitating innovation.

This paper identifies key considerations from FDA guidances that may facilitate the development of DDMs that quantify meaningful aspects of health, consistent with best practices for patient engagement, measurement science, and software and AI verification and validation. It is important to involve patients, caregivers, and clinicians in the development process to determine meaningfulness and clinical relevance of the DDM. FDA believes a collaborative, patient-centered approach may help unlock the full potential of DHTs to support medical product development and innovation.