

Setting the Stage for a Fit-For-Purpose DTx Evidentiary Standard

IN PARTNERSHIP WITH



www.curebase.com

www.dtxalliance.org

What Is a Digital Therapeutic?

Digital therapeutics (DTx) deliver to patients evidence-based therapeutic interventions that are driven by high-quality software programs to treat, manage, or prevent a disease or disorder. They are used independently or in concert with medications, devices, or other therapies to optimize patient care and health outcomes.

DTx products incorporate advanced technology best practices relating to design, clinical evaluation, usability, and data security. They are certified or cleared by regulatory bodies as required to support product claims regarding risk, efficacy, and intended use.

DTx empower patients, clinicians, and payors with intelligent and accessible tools for addressing a wide range of conditions through high-quality, safe, and effective data-driven interventions.

PER INDUSTRY STANDARDS, DIGITAL THERAPEUTIC PRODUCTS SHOULD ADHERE TO THESE FOUNDATIONAL PRINCIPLES:

- 1 Treat, manage, or prevent a disease or disorder.
- 2 Produce a medical intervention that is driven by software.
- 3 Incorporate design, manufacturing, and quality best practices.
- 4 Engage end users in product development and usability processes.
- 5 Incorporate patient privacy and security protections.
- 6 Apply product deployment, management, and maintenance best practices.
- 7 Publish trial results inclusive of clinically meaningful outcomes in peer-reviewed journals.
- Be reviewed and cleared or certified by regulatory bodies as required to support product claims of risk, efficacy, and intended use.
- Make claims appropriate to clinical evaluation and regulatory status.
- 10 Collect, analyze, and apply real-world evidence and/or product performance data.

For more information please visit https://dtxalliance.org/understanding-dtx/what-is-a-dtx/

Contents

Inti	roduction
Par	t l: Foundational Concepts
	DTx Intervention Classifications
	Clinical Outcome Domains
	Clinical Study Participant Selection
	DTx Clinical Evidence Endpoints
	Selection and Assessment of Endpoints9
	DTx Endpoint Types
	Clinical Outcome Categories11
	Process Variables
	Biomarkers
	Reported Outcomes
	Control Arm Considerations
	Control Arm Design
	Blinding Clinical Studies

Part II: Clinical Evidence Types
Part III: Demonstrating Clinical Evidence Quality17
Part IV: Clinical Evidence Timing Recommendations
DTx Evaluation Considerations
Discussion
Next Steps
Conclusion
Acknowledgements
Annex
Endnotes

Introduction

The past decade has seen a profound acceleration of digital health innovation, particularly with the rise of evidence-based digital therapeutics (DTx) and publication of an increasing number of randomized controlled trial (RCT) and real-world evidence (RWE) studies demonstrating DTx clinical benefits beyond those provided by traditional medical treatments. This has resulted in DTx regulatory approvals, favorable health technology assessments, incorporation into clinical practice guidelines, establishment of permanent reimbursement pathways, and expanded DTx access for target patient populations.

Globally, many decision makers are recognizing the need to develop guidelines that are tailored to the unique characteristics of DTx. Organizations such as the United States Food and Drug Administration (FDA), National Institute for Health and Care Excellence (NICE) in the United Kingdom, World Health Organization (WHO), and others are developing recommendations focused on documenting best practices in clinical evidence generation for digital health and therapeutic products.

Digital therapeutics require a fit-for-purpose evidentiary standard due to their agile development processes, mechanisms of action, ability to generate real-time outcomes, ongoing iterative nature, lower potential risk profiles, and place in clinical therapy. Although a DTx-specific standard will incorporate aspects of existing drug and medical device frameworks, it must include tailor-made components that enable appropriate and efficient DTx product design, capability, and performance assessments.

This publication is intended for healthcare decision makers (HCDMs) who are responsible for DTx product evaluation, patient access determinations, and ongoing product performance assessments. It sets the stage for this new DTx-specific evidentiary standard by providing foundational principles that apply to the DTx category of medicine, in addition to baseline expectations for HCDMs related to the types, quality, and timing of clinical trials necessary to evaluate and implement DTx therapies in real-world settings.

HCDMs—including government bodies, payors, health system leaders, and clinicians—require reliable, harmonized evaluation frameworks to appropriately determine the safety, efficacy, and impact of DTx treatments. Developing a common understanding of what types and quality of DTx clinical evidence qualifies as a necessary and sufficient evidence package will prevent unnecessary delays in providing patients around the world with appropriate access to DTx products.

"Every good product starts with a clear answer to the questions 'what problem are you solving? For whom? How?'. It is imperative that digital health products and services have a clear purpose. Suppliers and potential suppliers should understand how their innovation or technology will result in better provision and/or outcomes for people and the health and care system. This could be through:

- » improvements in patient outcomes or experience
- » generation of new knowledge and capabilities
- » generation of a firmer evidence base, and reduction in uncertainty
- » efficiency improvements"

Source: UK Department of Health & Social Care. *Guidance: A guide to good practice for digital and data-driven health technologies.* Updated January 19, 2021.¹

Part I: Foundational Concepts

DTx products generate multiple evidence types during their life cycle, including clinical evidence, RWE, health economic and outcomes research (HEOR), and context-specific implementation pilots. Although this publication focuses primarily on clinical evidence requirements, comprehensive DTx evidence dossiers should provide HCDMs with sufficient insights necessary to assess the value and effect of DTx interventions at the individual patient and population health levels of care.

"In general, it is not possible to set a blanket threshold for all types of statistical assessments of clinical validation, as these will differ depending on the clinical measurement, patient population, and context of use."

Source: Goldsack, J.C., Coravos, A., Bakker, J.P. *et al.* (2020). Verification, analytical validation, and clinical validation (V3): The foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *npj Digital Medicine 3*, 55.²

A fit-for-purpose DTx evidentiary standard—applying both to products provided via prescription and non-prescription routes—must include the types, quality, timing, and levels of clinical evidence considered to be sufficient for DTx product regulatory, reimbursement, and clinical use purposes.

As a first step, the following considerations should factor into developing a fit-for-purpose evidentiary standard framework:



DTx Intervention Classifications

The International Classification of Health Interventions (ICHI) defines a health intervention as "an act performed for, with or on behalf of a person or population whose purpose is to assess, improve, maintain, promote, or modify health, functioning, or health conditions" (WHO, 2019).³ Flowing from that definition, the classification encompasses interventions across all sectors of the health system and is built around three axes:

- » Target. The entity on which the action is carried out
- » Action. The deed done by an actor to the target
- » Means. The processes and methods by which the action is carried out

This publication provides guidelines for clinical evidence generation to demonstrate the value and effect of DTx interventions to treat, manage, or prevent medical conditions. It does not provide guidance for products primarily responsible for diagnosis, monitoring, screening, or general education of medical conditions.

Clinical Outcome Domains

DTx treatment developers and manufacturers primarily focus on six domains when undertaking clinical evidence evaluations. These domains form the foundation of DTx clinical evidence dossiers and a guide for HCDMs as they assess the sufficiency of available DTx evidence.

Primary clinical outcome domains include:

1. Safety. The Code of Federal Regulations (CFR) Title 21 defines safety outcomes as the valid scientific evidence to adequately demonstrate the absence of unreasonable risk of illness or injury associated with the intended use of the device and condition of its use.⁴

This publication does not focus on product cybersecurity as a subset of safety. Cybersecurity is most often evaluated via other mechanisms (i.e., alignment with regional or national requirements, certifications) instead of clinical trials.

2. Benefit. Efficacy trials determine whether an intervention produces the expected result under controlled circumstances.⁵

Effectiveness studies measure the degree of beneficial effect under "real-world" settings.⁶

3. Durability and duration of response. Durability of an intervention may be defined as its ability to postpone or delay progression of disease in a safe and well-tolerated manner.⁷

Duration of response is the period of time the treatment effects persist after treatment is completed or discontinued.⁸

4. Usability and accessibility. Usability is the characteristic of the product that establishes effectiveness, efficiency, ease of user learning, and user satisfaction.⁹

Accessibility refers to how a technological product can be used by people from a population with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use.¹⁰

5. User engagement. User engagement with technological aspects refers to how, how frequently, and for what duration a user makes use of the DTx system, including software, hardware, and any other components necessary for the DTx to function.¹¹

User engagement, however, reflects a complex set of factors that go beyond standard adoption and utilization measures (i.e., attrition), and extends into the quality of patients' interaction with the intervention (i.e., task success, user satisfaction, net promoter score).

6. Behavior change. Any alteration or adjustment of behavior that affects a patient's functioning brought about by psychotherapeutic, other interventions, or occurring spontaneously.¹²

Clinical Study Participant Selection

Equitable participant selection is fundamental for enhancing the social value of DTx products and promoting their clinical effectiveness. Individuals selected for a study should reflect the population for whom the product is intended. The target population is represented by enrolling participants who meet specific enrollment criteria that match the key characteristics of the intended target population.

Insufficient participation from some groups within a population can result in inadequate information pertaining to the safety and effectiveness of the DTx product in important subpopulations.

Although outside the scope of this publication, additional insights about product appropriateness and usability by target populations and subpopulations may be generated through real-world data (RWD) and RWE evaluations. These real-world insights may be used to provide greater context to clinical evidence outcomes and influence the development of future product iterations.

DTx Clinical Evidence Endpoints

HCDMs assessing the quality of DTx evidence outcomes may rely on the following considerations related to the use of endpoints in data generation, assessment, and analysis.

Selection and Assessment of Endpoints¹³

A clinical endpoint is a measure of the benefit or reduced harm of an intervention and should therefore be a key symptom or sign of a disease, a valid measure of clinical benefit due to treatment, clinically relevant, responsive to change, and accepted by clinicians. Clinical endpoints should be reproducible, facilitate comparisons across studies, valid, and quantify what was intended to be measured. Protocols should clearly define endpoints and pre-specify clinically relevant effects based on published standards or consensus among independent clinicians and/or patients. In some cases, authoritative organizations have developed standards or guidance documents specifying how to define and implement specific types of digital health clinical endpoints in clinical trials (i.e., FDA, Consumer Technology Association, Digital Medicine Society).

Clinical endpoints are intended to measure the impact of an intervention on how a patient feels, functions, or survives. Because most interventions affect more than one disease characteristic, most trials measure multiple endpoints to document the benefits of the intervention. To illustrate the issue of multiple endpoints in clinical trials and studies, consider the example of efficacy measures. As described above, efficacy endpoints are measures intended to reflect the effects of an intervention. Efficacy spans several domains including clinical events, patient symptoms, measures, or functions. Most diseases have impacts on multiple domains and often on multiple organ systems. Therefore, it is often necessary to evaluate the effect of an intervention on several different endpoints.

Testing multiple endpoints increases the risk that any statistical significance observed in the analysis may be the result of chance and drive a false conclusion about the intervention's effects. To mitigate this risk, study protocols must pre-specify an outcome hierarchy and make a prospective determination of outcome classification according to their type—primary, secondary, and exploratory. For an efficacy study, the set of primary endpoints are the outcomes or outcomes that establish the effectiveness of the intervention. Power calculations should be part of study protocols and the sample size must be based on the desired power level and significance level to detect an expected effect size for the primary endpoint. Statistical analysis plans must specify approaches to adjust for multiple comparisons.

Designating primary endpoints is based on clinical importance and the likelihood of demonstrating an effect. Secondary endpoints are outcomes selected to demonstrate additional effects of the intervention after success on the primary endpoint or outcomes that provide evidence of the mechanism of action underlying the demonstrated clinical effect. The study may be powered to explore secondary endpoints; however, positive results on the secondary endpoints can be interpreted only if a treatment effect on the primary endpoint family is demonstrated first. Exploratory endpoints may include clinically important events that are expected to occur too infrequently to show a treatment effect or endpoints that for other reasons are thought to be less likely to show an effect but are included to explore new hypotheses.

DTx Endpoint Types

Types of endpoints that DTx products use include:

- » Objective endpoints. Objective endpoints are those that can be measured without being influenced by the beliefs or expectations of the user. They tend to be quantitative, well-defined, and reliable regardless of the observer. Examples include wearables for measuring heart rate variability, sleep, or gait.
- » Subjective endpoints. Subjective endpoints can be influenced by the perceptions and beliefs of the users. These could be tested and validated tools such as questionnaires but nonetheless be impacted or influenced by individual interpretation or perception. Examples include assessments of depression or anxiety questionnaires or assessments of pain, and patient reported outcome surveys.
- » Composite endpoints. Composite endpoints combine numerous endpoints into a single outcome.¹⁴ Such endpoints may be useful for providing a more multifaceted view of effects and can potentially help reduce bias. An example is the European Medicines Agency qualified PROActive composite measure, which combines patient-reported questionnaire data with activity monitor data in chronic obstructive pulmonary disorder (COPD) patients.
- » **Intermediate clinical endpoints.** An intermediate clinical endpoint is defined as "a measure of a therapeutic effect that is considered reasonably likely to predict the clinical benefit of an intervention, such as an effect on irreversible morbidity and mortality."¹⁵ An example is tracking heart rate and body movement during sleep as a proxy for sleep disturbance or nightmares.
- » Surrogate endpoints. Surrogate endpoints are measures that can be used as a substitute for a clinically meaningful endpoint. Although surrogate endpoints may predict a clinical benefit, they are not themselves clinical benefits.¹⁶ Surrogate endpoints may take a shorter time to observe or may be easier to observe when the clinical endpoint may be too difficult, expensive, or unethical to measure. These must be validated. Examples include HbA1c for diabetic complications.

Clinical Outcome Categories

Clinical outcomes are measures of mortality and morbidity, complications, symptom reduction, and functional status improvements for the disease of interest. The domains of health that a digital therapeutic study may assess include:¹⁷

- » Survival outcomes. Survival outcomes span all-cause survival and cause-specific survival.
- » Physiological and clinical outcomes. Physiological and clinical outcomes that measure signs and symptoms of disease, laboratory and other clinical test parameters, and anthropometric measurements.
- » Life impact or functioning outcomes. Life impact or functioning outcomes that measure the impact of the intervention on the following subdomains: physical functioning (i.e., the ability to perform physical activities of daily living); social functioning (i.e., the ability to participate in social activities and to operate within society); role functioning (i.e., the ability to fulfill the requirements of an individual's role such as caregiver or employee); emotional functioning (i.e., the impact of the disease on well-being and psychological status); cognitive functioning (i.e., the impact of the disease on cognitive domains such as memory and attention); and quality of life (i.e., perception of an individual's position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards, and concerns).
- » Adverse events. Adverse events are defined by the FDA as any untoward medical occurrence associated with the use of a drug/intervention in humans, whether or not considered drug/ intervention related. Tolerability is defined as "the degree to which overt adverse effects can be tolerated by the subject" by the International Conference on Harmonization (ICH) and measured as withdrawal from intervention.

Process Variables

Apart from the health domains listed above, studies of digital interventions may also evaluate the delivery of care and how the participant interacts with the intervention. Such aspects of process may include the following:

» Adherence and compliance. Although the two terms are often used interchangeably, an implicit difference exists between them. *Compliance* is defined as "the extent to which the patient's behavior matches the prescriber's recommendations," whereas *adherence* is defined as the "active, voluntary, and collaborative involvement of the patient in a mutually acceptable course of behavior to produce a therapeutic result."¹⁸ The main distinction is that the adherence metric emphasizes patient participation in the choice of treatment goals and regimen.

Adherence consists of two components: compliance and persistence. Compliance, measured as a proportion, refers to using the digital therapeutic per the recommended time period (i.e., twice daily sessions for four weeks), whereas persistence refers to the accumulation of time from initiation to discontinuation of therapy (i.e., four weeks).¹⁹ Although conceptually similar, compliance refers to the intensity of exposure to the intervention during the duration of therapy, whereas persistence, often measured as patient retention, refers to the overall duration of therapy.

- » User engagement. User engagement is a concept that goes beyond whether the individual adheres to and persists with the intervention, but instead evaluates the quality of their interaction with the intervention. Although the definition of engagement is still evolving—especially as it pertains to clinically informed, data-driven measurement processes—current metrics include capturing users' level of attention, task success, and interactivity with the digital intervention, as well as the particular sections and elements of the intervention the individual engages with.²⁰ Beyond "generic" measures of engagement (i.e., number of sessions, weekly active usage, or program completion), *measures of meaningful engagement* may refer to "a clinically informed and data-driven approach to identify specific engagement metrics that uniquely predict the long-term value for a digital therapeutic."²¹
- » User satisfaction and preference. Treatment satisfaction is defined as a patient's perception of the extent to which all aspects of the intervention (i.e., mode of delivery, duration, and benefit) meets their health needs.²²

User preference endpoints aim to capture the value that patients place on aspects of the intervention. The FDA defines patient preference endpoints as "qualitative or quantitative assessments of the relative desirability or acceptability to patients of specified alternatives or choices among outcomes or other attributes that differ among alternative health interventions."²³

Note that while process variables are valuable endpoints and measure metrics that may contribute to the effectiveness of the intervention, they must be accompanied by and correlated with a meaningful clinical outcome demonstrating treatment benefit.

Biomarkers

The Biomarkers, EndpointS, and other Tools (BEST) glossary defines a *biomarker* as "a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions."²⁴ Biomarkers may have molecular, histologic, radiographic, physiologic, or behavioral characteristics.

Highly relevant to the evaluation of digital interventions are digital biomarkers defined as a characteristic or set of characteristics, collected from digital health technologies, that are measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.²⁵ Digital biomarkers may represent digitization of measurement of previously established traditional biomarkers, (i.e., gait or heart rate), or may be novel in that they provide an entirely new domain of measurement (i.e., vocal biomarkers).²⁶ As with any biomarker, digital biomarkers require rigorous testing and validation.

Reported Outcomes

Reported outcomes, also referred to as clinical outcome assessments,²⁷ may also be defined based on the source of the assessment. Examples include:

» Clinician-reported outcomes. Clinician-reported outcome (ClinRO) is defined as "a measurement based on a report that comes from a trained health-care professional after observation of a patient's health condition." Most ClinRO measures involve a clinical judgment or interpretation of the observable signs, behaviors, or other manifestations related to a disease or condition. ClinRO measures cannot directly assess symptoms that are known only to the patient. ClinRO measures can include reports of particular clinical findings (i.e., presence of a skin lesion or swollen lymph nodes), clinical events (stroke, heart attack, death, hospitalization for a particular cause), or rating scales, such as the Hamilton Depression Rating Scale (HAM-D) for assessment of depression.²⁸ ClinROs reflect the clinician's assessment of the patient's symptoms. Depending on the purpose of the assessment, these may need to be conducted and contextualized appropriately (i.e., clinician's assessment of depression symptom severity versus patient's self-report).

- Patient-reported outcomes. Patient-reported outcome (PRO) is defined as "a measurement based on a report that comes directly from the patient (i.e., study participant) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else." A PRO can be measured by self-report or by interview provided that the interviewer records only the patient's response. Symptoms or other unobservable concepts known only to the patient can only be measured by PRO measures. PROs can also assess the patient perspective on functioning or activities that may also be observable by others. PRO measures include rating scales (i.e., numeric rating scale of pain intensity or Minnesota Living with Heart Failure Questionnaire for assessing heart failure) and count of events (i.e., patient-completed log of emesis episodes or micturition episodes).²⁹ PROs reflect the patient's experience and perspective. Depending on the purpose of the assessment, these may need to be contextualized appropriately (i.e., asking a patient with Alzheimer's disease to rate their memory or functioning versus their clinician or caregiver).
- » Observer-reported outcomes. Observer-reported outcome (ObsRO, which includes caregiver) is defined as "a measurement based on a report of observable signs, events or behaviors related to a patient's health condition by someone other than the patient or a health professional." Generally, ObsROs are reported by a parent, caregiver, or someone who observes the patient in daily life and are particularly useful for patients who cannot report for themselves (i.e., infants or individuals who are cognitively impaired). An ObsRO measure does not include medical judgment or interpretation. ObsRO measures include rating scales, such as Acute Otitis Media Severity of Symptoms scale (AOM-SOS), and Face, Legs, Activity, Cry, Consolability scale (FLACC), or counts of events (i.e., observer-completed log of seizure episodes).³⁰ ObsROs reflect the observer's experience and perspective on the patient's condition. Depending on the purpose of the assessment, these may need to be conducted and contextualized appropriately (i.e., asking a parent about their child's aggressive behaviors at school).
- » Performance outcomes. Performance outcome (PerfO) is defined as "a measurement based on standardized task(s) actively undertaken by a patient according to a set of instructions. A PerfO assessment may be administered by an appropriately trained individual or completed by the patient independently. PerfO assessments include measures of gait speed (i.e., timed 25 foot walk test using a stopwatch or using sensors on ankles), or measures of memory (i.e., word recall test)."³¹

Control Arm Considerations

Designing appropriate control arms for clinical studies can be a complex process. Although this section is non-exhaustive, it provides high-level considerations for DTx products.

Control Arm Design

For comparative clinical outcome studies, it is necessary to select or design a control arm. For DTx, this can be particularly challenging and requires careful consideration of the treatment delivery, its mechanism of action, and the type of evidence being collected, among other things.

Types of controls may include:

- » Placebo controls, such as sham applications that are designed to resemble an active therapy but do not possess a clinically active mechanism of action
- » Non-placebo controls, such as:
 - Waiting list
 - Active comparator (i.e., treatment as usual, the current standard of care, other digital application, an alternative treatment modality)
 - Historical control

An alternative study design that does not incorporate controls is a head-to-head trial, where the aim is to prove equivalence of two active interventions or a higher level of efficacy of one intervention over another.

Blinding Clinical Studies

Blinding, the concept that a participant in a trial does not know which arm of a study they are assigned to, is another challenge for DTx when conducting comparative trials. Different types and levels of blinding exist and careful consideration should be put into selecting the appropriate blind for each study. When blinding is incorporated into the study design, it is important to clearly define who can and will be blinded (i.e., researcher, patient, caregiver) and to include a plan to demonstrate how the integrity of the blinding process was maintained throughout the study. DTx manufacturers and HCDMs must reference local regulatory guidance requirements when establishing and conducting blinding protocols.

When claiming that a trial has been blinded, clinical trial sponsors are recommended to abide by research reporting standards (i.e., <u>Consolidated Standards of Reporting Trials, CONSORT</u>)³² to ensure that all appropriate blinding checks are conducted during trials. Many DTx trials select a blind-to-hypothesis approach over a blind-to-assignment approach. When digital controls are deployed, they should only be referred to as a sham application if sponsors are able to guarantee that participants remain appropriately blinded.

Part II: Clinical Evidence Types

DTx evidence dossiers may cover a spectrum of domains, spanning from product feasibility, usability, accessibility, and user engagement to real-world impact and effectiveness. DTx products must therefore use various study designs to assess the appropriate endpoints during the appropriate phase of the product's life cycle.

Product manufacturers may utilize the following clinical evidence types as part of a DTx evidence dossier:

- » Non-experimental, observational studies
 - **Descriptive.** Case report, case series, cross-sectional (descriptive or prevalence)
 - **Analytical.** Cross-sectional survey, case-control, cohort (prospective or historical)
 - Implementation pilot. Assess site-specific implementation capacity and value
 - Localization pilot. Assess cultural adaptation, language translation, linguistic accuracy/ validation, etc.
 - Prospective. Observational, cohort
- » Product analyses
 - Retrospective analyses. Chart reviews, medical/pharmacy claims, electronic medical records, other novel data sources
 - Expert reviews. Clinical practice guidelines, clinical pathways, health technology assessment agency evaluations, published systematic reviews
 - Coverage decision assessments and formulary reviews. External organization product evaluations, product indication reviews
 - Patient perspectives. Practical use of therapies, patient preference information (PPI)
 - Consumer studies. Medical affairs, marketing study

- » Experimental, interventional clinical studies
 - Non-controlled studies. Prospective single arm trial, open label trial, head-to-head comparative trial
 - Controlled trials. Non-randomized controlled trial, self-controlled study, crossover study
 - RCT. Unblinded/open-label, single-blind, double-blind, triple blind
- » RWD
 - Product performance and technical outputs
 - Patient-specific clinical and self-reported outcomes
 - End user and clinician engagement and satisfaction measures
- » RWE
 - Pragmatic clinical trials with real-world elements
 - RWE as a retrospective or prospective observational study
- » Systematic reviews of interventional, observational, or mixed methods studies
 - Meta analysis

The NICE evidence standards for digital health technologies recommends that "the choice of study design should be appropriate for the intended purpose of the DHT [digital health technology]. Randomized controlled trials would be preferable where this study design is appropriate. High quality, comparative real-world study designs may also be acceptable."

Source: NICE. Evidence standards framework for digital health technologies. August 9, 2022.³³

Additional forms of evidence generation that product manufacturers may utilize include HEOR studies (i.e., economic impact), implementation pilot studies (i.e., clinical workflow optimization), and localization pilot studies (i.e., linguistic validation, cultural adaptation).

Part III: Demonstrating Clinical Evidence Quality

Table 1 provides product manufacturers and HCDMs with recommendations for demonstrating and assessing the quality of clinical evidence developed for DTx products related to the five clinical outcome domains of (1) safety, (2) benefit, (3) durability and duration of response, (4) usability and accessibility, and (5) user engagement within the regulatory, payment, and clinical use and acceptance ecosystems.

Domain	Regulatory Pathway	Payment Pathway	Clinical Use and Acceptance
Safety	 » Short-term safety data, collected during a patient's use of the DTx or within a clinical trial, are essential for pursuing the regulatory pathway. » Long-term safety data, whose collection extends beyond the time period of a clinical trial, may be important for specific therapeutic areas. » Objective and subjective endpoints are used for collecting safety data to support DTx regulatory approval. » Interventional clinical trials and RWD/RWE clinical studies are performed for collecting safety data to support DTx regulatory approval. 	 » Safety data could be important to demonstrate that a DTx can reduce healthcare system burden and support payor goals of cost avoidance, as applicable. » Cost-analysis of supplemental safety data, such as a reduction in a certain type of adverse event, can be useful to support financial claims that payors desire. 	» Publishing clinical trial results in peer-reviewed journals that include appropriate safety data is important for gaining clinician acceptance.

TABLE 1. QUALITY OF EVIDENCE RECOMMENDATIONS

Domain	Regulatory Pathway	Payment Pathway	Clinical Use and Acceptance
Benefit	 » Efficacy data are essential for pursuing the regulatory pathway. » Effectiveness data are beneficial but may not be required by regulators for approval. » Engaging with regulators early in the process of study design to ensure data will be sufficient is highly recommended. » Regulators expect well- controlled investigations to support effectiveness claims, and in certain instances, other study designs may be sufficient. » If QoL claims are to be made, then it is required to collect QoL measures. » Validated QoL PROs are expected to provide the most reliable and acceptable level of evidence. 	 » Aligning clinical trial eligibility criteria with the target patient population is crucial in meeting payor requirements for the coverage of appropriate patients. » The ability to make strong clinical claims backed by robust evidence generated using high-quality study designs is important. The study design needs to include endpoints that are relevant to payors. » Demonstrating that a DTx can replace or delay the use of a more costly therapy is often attractive to payors. » Payors are concerned with issues related to relevance, quality, and interpretability of PROs when evaluating data from these instruments. 	 » Clinicians will understand and accept benefit claims when it comes from a trial with a rigorous study design, providing evidence of comparative effectiveness, and gold standard outcome measures. » Demonstrating a meaningful QoL measure is important for gaining clinician acceptance.
Durability and Duration of Response	» Long-term efficacy and effectiveness data may be requested by regulators. It is important to engage regulators early in the design process to determine what durability evidence they may request.	 » Payors prefer DTx products that have a durable treatment effect, as opposed to creating a blanket requirement for continued product use. Therefore, demonstrating the enduring effects of the DTx seeking a pathway to reimbursement. » The potential exists for negotiating with payors to provide incentives based on the results of prospectively designed RWD/RWE studies that demonstrate outcomes such as an overall reduction in healthcare costs over a period of time. 	» Clinician uptake is more likely when therapies are supported with RWE that demonstrate duration of therapy response.

•

4

6

Domain	Regulatory Pathway	Payment Pathway	Clinical Use and Acceptance
Usability & Accessibility	 » Evidence must meet acceptable levels of compliance and adherence to support claims. » Some regulatory pathways may require explicit usability testing (i.e., summative (validation) testing required by the FDA for software as a medical device (SaMD) products). 	» Dependent on robust RWE being available to support DTx product claims and demonstrate value.	» Demonstrating the ability for the DTx to fit within the existing clinician workflows is key to acceptance within the clinical community. This may include evidence to support how clinicians can easily access and interpret data generated by the DTx.
User Engagement	 » Evidence must meet acceptable levels of compliance and adherence to support claims. » It is recommended to involve end users in the trial design process, since involving patients could lead to higher levels of user engagement and better quality of evidence across categories. ³⁴ 	 Demonstrating patient engagement and interest in using the DTx is crucial, as this is often stated as a primary concern by payors (i.e., that even if the DTx works, patients might not use it as intended). Generating evidence to support patient engagement may be done as part of an interventional clinical trial (i.e., RCT), or gathered in the form of RWD and analyzed to produce RWE. Evidence supporting high treatment persistence rates may be beneficial, particularly if the therapeutic contains components known to have barriers to behavioral activation (i.e., cognitive behavioral therapy (CBT)). 	» Clinician uptake is more likely when DTx therapies are supported with RWE that demonstrate appropriate user engagement.

•

Part IV: Clinical Evidence Timing Recommendations

DTx clinical evidence dossiers are developed in phases, with studies often conducted during specific periods of the product development, launch, and post-market life cycle. Table 2 identifies the types, timing, and target outcomes of clinical evidence generation during four product development phases: pre-development, early development, late phase development, and post-marketing.

Development Phase	Purpose	Types of Study Designs	Evidence Collected	Pathway Relevance (in order of relevance)	
Pre- Development (exploratory, discovery phases)	e- evelopment coloratory, accovery ases) The main driver of this phase is identifying and improving understanding of the product's mechanism of actions (MoA) or active principles. * Review of existing clinical trial data and systematic reviews * Experimental clinical trials * Behavior change * Safety (short- term) * Benefit (efficacy)	Regulatory	Can support clinical evaluation (for regulatory and clinical association); early evidence of safety and performance		
			Clinical Practice Acceptance	Supports building trust in the product and an understanding of the product MoA	
				Payment Pathway	Builds an early evidence base for product claims

TABLE 2. TYPES, TIMING, AND TARGET OUTCOMES OF CLINICAL EVIDENCE GENERATION

4

Development Phase	Purpose	Types of Study Designs	Evidence Collected	Pathway Rel (in order of i	levance relevance)
Early Development (early feasibility, proof of concept phases)	This phase moves from basic science into productization: this will include validation of concepts, early safety and effectiveness data, as well as refining the therapeutic protocol (i.e., dosing, frequency of administration). It typically leverages shorter duration studies than late phase/ pivotal studies. Product design is also initially tested in this phase (i.e., usability, human factors), as well as generating initial evidence of value beyond	 » Experimental clinical trials » Product analysis » Non- experimental trials 	 » Behavior change » Safety (short- term) » Benefit (efficacy, effectiveness) » Usability and accessibility » User engagement » Compliance and adherence 	Clinical Practice Acceptance Regulatory Payment Pathway	Ecosystem acceptance and adoption, by increasing market confidence on the effectiveness of the product Supports wider adoption by building users' trust Informs future regulatory pathways Builds early evidence for product claims (i.e., safety, clinical utility)
	clinical outcomes (i.e., operational, experiential).				
Late Phase Development (pivotal, validation phases)	ase ment fon this phase includes developing core evidence for the product's regulatory submission. It is also of critical importance to consider the product's level of evidence of value and clinical utility to drive initial adoption and payor support immediately following regulatory approval.	 » Experimental clinical trials » Product analysis » Non- experimental trials 	tal » Safety (short- and mid- term) » Benefit (efficacy, effectiveness) » Usability and accessibility » User engagement » Compliance and adherence » Behavior change	Regulatory	Core evidence collected for product safety, effectiveness, and performance
				Payment Pathway	Main body of evidence developed for early conversations with payors
				Clinical Practice Acceptance	Evidence developed here will help drive adoption; regulatory approval contributes to this goal

•

Setting the Stage for a Fit-For-Purpose DTx Evidentiary Standard

6

Development	Purpose	Types of	Evidence	Pathway Rel	levance
Phase		Study Designs	Collected	(in order of I	relevance)
Post-Marketing (intended use, real-world use phases)	In this phase, data are continuously collected from real-world use and studies to further expand the body of evidence that supports product claims and commercialization strategies. This includes gathering insight into durability and long-term safety, engagement, and impact.	 » RWD » RWE » Experimental clinical trials » Product analysis » Non-experimental trials 	 » Safety (long- term) » Benefit (efficacy, effectiveness) » Usability and accessibility » User engagement » Compliance and adherence » Behavior change » Durability 	Payment Pathway Clinical Practice Acceptance Regulatory	Pivotal evidence to expand to market access and revenue generation (i.e., HEOR studies) Expands on evidence of utility and value to further expand adoption Support for validation of new commercial claims

•

DTx Evaluation Considerations

Digital therapeutics can be used as standalone therapies, in conjunction with, or in place of other clinically validated therapies. Based on their proximity to existing therapies, HCDMs therefore frequently evaluate DTx products using the same criteria and requirements as pharmaceuticals and other similar treatments. This perceived expectation for DTx products to meet the same requirements, in terms of types, quality, and volume of clinical evidence as pharmaceuticals ignores DTx's agile development processes, mechanisms of action, ability to generate real-time outcomes, ongoing iterative nature, lower potential risk profiles, and place in therapy.

DTx products therefore require a fit-for-purpose evaluation approach that incorporates aspects of existing pharmaceutical and medical device evaluation frameworks, but are designed to evaluate the safety, efficacy, and impact of DTx therapies more appropriately. Notably, a fit-for-purpose DTx product evidence evaluation framework, as initially proposed in this publication, does not weaken evidentiary requirements, but rather provides greater evidentiary strength and robustness that reflects how DTx products are designed and used in real-world settings.

Pharmaceuticals are recognized as an "embodied technology" for which the efficacy relates solely to the correct dose of the drug, which chemically interacts with the body's physiological systems. DTx products, on the other hand, use software-driven technologies to deliver their behavioral and physical impacts on end users, thus carrying different types and levels of risk than chemical-based products. For example, a counterfeit or low-quality DTx that relies on or disseminates incorrect, incomplete, or inconsistent therapeutic impact may be harmful to end users by delivering inadequate or inappropriate clinical interventions³⁵ (i.e., incorrect insulin dosing, inaccurate medical directions, insufficient infrastructure for mental health risk management). Although DTx products are not risk-free, they do carry different risks than pharmaceuticals, and should thus be evaluated appropriately.

Additionally, unlike pharmaceuticals, DTx products undergo incremental product modifications during the post-marketing phase. These product changes, or iterations, may impact many things, ranging from product functionality and bug fixes, to clinical and usability improvements that may impact the therapy's efficacy. Thus, DTx evaluation approaches that rely on RCTs and prospective studies often need to be paired with additional studies during the product's life cycle to account for the iterative nature of technological product design.³⁶ Where applicable, HCDMs should accept study designs that account for specific DTx characteristics, such as multi-phasic optimization strategies, sequential multiple assignment randomized trials, and micro-randomized trials developed with an adaptive design.³⁷

Discussion

To further enable patient access to DTx therapies across the global marketplace, HCDMs and product manufacturers require unified evaluation frameworks, ensuring consistency in DTx product assessment, regulation, reimbursement, and clinical scalability.

This publication sets the stage for a fit-for-purpose DTx evidentiary standard and provides a foundational set of expectations related to the types, quality, and timing of clinical trials necessary to evaluate and implement DTx products in real-world settings. The recommendations build on common elements of other therapeutic assessment frameworks, while specifically addressing DTx-specific considerations.

As more HCDMs at the local, national, and regional levels embark on evaluating DTx, it is important that evidence dossier requirements and frameworks increasingly move toward a harmonized, consistent set of expectations related to the type, timing, and quality of DTx studies consider to sufficiently demonstrate product safety, efficacy, and impact.

Next Steps

Future considerations to enable movement toward a fit-for-purpose DTx evidentiary standard include:

- Technical considerations related to DTx study pricing, use of remote study protocols, participant selection, study powering and sample size, etc.
- Clarifying RCT study applicability in the context of the DTx iterative process
- Generalizability of product studies across national jurisdictions
- Conducting appropriate RWE, HEOR, and product implementation pilot studies (i.e., clinical workflow optimization)
- Appropriate use of localization pilot studies (i.e., linguistic validation, cultural adaptation)

Conclusion

This publication provides a baseline set of expectations related to the types, quality, and timing of clinical trials necessary to evaluate and implement digital therapeutics in real-world settings.

Because digital therapeutics are regulated as medical devices, it follows that these products be evaluated using fit-for-purpose evaluation frameworks, as opposed to being required to meet pharmaceutical-specific clinical evidence criteria. HCDMs should adopt a fit-for-purpose evaluation approach that addresses the unique nature and innovation cycles of DTx products.

Developing a common expectation of what constitutes digital therapeutic clinical evidence sufficiency and building it into harmonized clinical evidence frameworks will lead to safer product use and prevent unnecessary delays in providing patients around the world with access to highquality, clinically validated DTx products.

ACKNOWLEDGEMENTS

The following individuals contributed their ideas and insights to this project:

- » Megan Coder, DTA
- » Emily Durden, Woebot Health
- » Ashleigh Golden, Woebot Health
- » John Hernandez, Google
- » Emilie Lopes-Fernandes, DTA
- » Vandana Menon, Akili Interactive
- » Dario Motti, Roche

- » Lisa Palko, Akili Interactive
- » Pooja Raghani, DTA
- » Adam Samson, The George Washington University
- » Whitney Stewart, Curebase
- » Simon Thomas, Freespira
- » Uma Vaidyanathan, Boehringer-Ingelheim

Setting the Stage for a Fit-For-Purpose DTx Evidentiary Standard

-

Annex

Acronyms

AOM-SOS	Acute Otitis Media Severity of Symptoms
BEST	Biomarkers, EndpointS, and other Tools
CFR	Code of Federal Regulations
Clin-RO	Clinician-reported outcomes
CONSORT	Consolidated Standards of Reporting Trials
COPD	chronic obstructive pulmonary disorder
DTx	digital therapeutic
FDA	U.S. Food and Drug Administration
FLACC	Face, Legs, Activity, Cry, Consolability
HAM-D	Hamilton Depression Rating Scale
HCDM	healthcare decision maker/s
HEOR	health economic outcomes research
HTA	health technology assessment
ICH	International Conference on Harmonization
ICHI	International Classification of Health Interventions
MoA	mechanism of actions
NICE	National Institutes for Health and Care Excellence (UK)
Obs-RO	Observer-reported outcomes
Perf-O	Performance outcomes
PPI	patient preference information
PRO	patient-reported outcome
QoL	quality of life
RCT	randomized controlled trial
RWD	real-world data
RWE	real-world evidence
SaMD	Software as a Medical Device
WHO	World Health Organization

ENDNOTES

- https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-ofconduct-for-data-driven-health-and-care-technology#have-a-clear-value-proposition
- ² https://www.nature.com/articles/s41746-020-0260-4#Sec25
- ³ https://www.tandfonline.com/doi/full/10.1080/14992027.2020.1736344#
- ⁴ https://www.ecfr.gov/current/title-21/chapter-l/subchapter-H/part-860/subpart-A/section-860.7
- ⁵ https://www.ncbi.nlm.nih.gov/books/NBK44024/; https://www.fda.gov/media/99769/download
- ⁶ <u>https://www.ncbi.nlm.nih.gov/books/NBK44024/</u>
- ⁷ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6785960
- https://www.astrazeneca-us.com/media/astrazeneca-us-blog/2018/clinical-trial-endpoints-in-cancer-research-four-terms-youshould-know-09242018.html#modal-historic-confirmation
- ⁹ https://www.fda.gov/files/medical%20devices/published/FDA-Perspectives-on--Human-Factors-in-Device-Development-(June-2012).pdf
- ¹⁰ https://ris.utwente.nl/ws/portalfiles/portal/40147417/FedericiBorsci2010Usabilityevaluation_Modelsmethodsandapplications.pdf
- ¹¹ https://arxiv.org/pdf/2205.01851.pdf
- ¹² https://dictionary.apa.org/behavior-change
- ¹³ https://www.eunethta.eu/wp-content/uploads/2018/01/Clinical-endpoints.pdf; www.fda.gov/files/drugs/published/Multiple-Endpoints-in-Clinical-Trials-Guidance-for-Industry.pdf; https://www.fda.gov/media/157635/download&sa=D&source=docs&ust= 1665155156023336&usg=AOvVaw1bbusLm_OnMSiqZxVdUIzZ
- ¹⁴ https://www.fda.gov/files/drugs/published/Multiple-Endpoints-in-Clinical-Trials-Guidance-for-Industry.pdf
- ¹⁵ https://www.fda.gov/patients/fast-track-breakthrough-therapy-accelerated-approval-priority-review/accelerated-approval
- ¹⁶ https://www.fda.gov/drugs/development-resources/surrogate-endpoint-resources-drug-and-biologic-development
- ¹⁷ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5854263/
- ¹⁸ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4087153/; https://www.ahajournals.org/doi/10.1161/circulationaha.108.768986# R1-768986
- ¹⁹ https://www.ahajournals.org/doi/10.1161/circulationaha.108.768986
- ²⁰ https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2768347#zoi200424r19
- ²¹ https://www.frontiersin.org/articles/10.3389/fdgth.2022.890081/full#B9
- ²² https://www.karger.com/Article/Fulltext/490413#ref20
- ²³ https://www.frontiersin.org/articles/10.3389/fphar.2019.01395/full#B10
- ²⁴ https://www.ncbi.nlm.nih.gov/books/NBK326791/
- ²⁵ https://www.nature.com/articles/s41746-022-00583-z
- ²⁶ https://www.frontiersin.org/articles/10.3389/fpsyt.2021.640741/full
- ²⁷ https://www.fda.gov/about-fda/clinical-outcome-assessment-coa-frequently-asked-questions#COADefinition
- ²⁸ https://www.ncbi.nlm.nih.gov/books/NBK338448/def-item/glossary.clinicianreported-outcome/
- ²⁹ https://www.ncbi.nlm.nih.gov/books/NBK338448/def-item/glossary.patientreported-outcome/
- ³⁰ https://www.ncbi.nlm.nih.gov/books/NBK338448/def-item/glossary.observerreported-outcome/
- ³¹ https://www.ncbi.nlm.nih.gov/books/NBK338448/def-item/glossary.performance-outcome/
- ³² https://www.researchgate.net/publication/12030511_The_Revised_CONSORT_Statement_for_Reporting_Randomized_Trials_ Explanation_and_Elaboration
- ³³ https://www.nice.org.uk/corporate/ecd7/resources/evidence-standards-framework-for-digital-health-technologiespdf-1124017457605
- ³⁴ https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-engagement-design-and-conduct-medicaldevice-clinical-studies
- ³⁵ <u>https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4576443/</u>
- ³⁶ <u>https://www.tandfonline.com/doi/full/10.1080/14737167.2021.1891883</u>
- ³⁷ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4732571/; https://mhealth.jmir.org/2016/3/e107



For more information, please contact **info@dtxalliance.org**

www.dtxalliance.org

