Cureus
Part of SPRINGER NATURE

Open Access Review Article

Integrating Evidence Generation and Commercialization in Digital Medicine: A Narrative Review

Review began 10/07/2026
Review ended 10/17/2026
Published 10/21/2026

© **Copyright** 2026
Lakhan. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118294

Shaheen E. Lakhan [1, 2, 3, 4, 5, 6]

1. Research and Development, Boricua Bio, San Juan, USA 2. Bioscience, Global Neuroscience Initiative Foundation, Boston, USA 3. Neurology, Western University of Health Sciences, Pomona, USA 4. Neurology, A.T. Still University School of Osteopathic Medicine in Arizona, Mesa, USA 5. Medicine, Morehouse School of Medicine, Atlanta, USA 6. Neurosciences, FIU Herbert Wertheim College of Medicine, Miami, USA

Corresponding author: Shaheen E. Lakhan, slakhan@gnif.org

Abstract

Digital medicine requires evaluation of clinical effects, delivery conditions, and the resources needed for use in practice. This narrative review examines how evidence requirements differ across clinical evaluation, regulatory authorization, purchasing, implementation, and postmarket assessment, using digital therapeutics as the principal illustrative application. Targeted searches of public literature and regulatory sources were supplemented with searches for additional studies of digital care delivery, economic outcomes, and implementation. Sources were selected for relevance to a defined evidence question; no systematic study enumeration or quantitative pooling was undertaken.

Published examples include aphasia rehabilitation, migraine prevention, major depressive disorder, schizophrenia, digital pain care, insomnia-related primary care costs, and implementation of substance use disorder treatment. The synthesis distinguishes randomized treatment effects from exploratory analyses, engagement from clinical benefit, regulatory status from purchasing evidence, and observed resource use from modeled economic consequences. The examples show why comparator choice, analysis population, human support, software version, and follow-up duration matter when translating findings to a different setting.

A nonsignificant primary analysis remains distinct from a favorable supportive analysis, and low uptake in routine care identifies an implementation problem without independently resolving efficacy. Original tables and conceptual figures organize these distinctions using established evidence-planning, implementation, economic, and measurement frameworks. The review does not rank products or establish that an evidence strategy causes commercial success. Evidence generation can inform commercialization when claims remain traceable to the intervention studied, the comparison performed, and the uncertainty that remains for the intended decision maker.

Categories: Therapeutics, Health Policy, Healthcare Technology
Keywords: commercialization, digital medicine, digital therapeutics, health economics, implementation science, integrated evidence planning, market access, patient engagement, real-world evidence

Introduction And Background

In this review, digital medicine encompasses clinical measurement and intervention using digital technologies. Digital therapeutics are software-based interventions with clinical evidence supporting their use to prevent, manage, or treat disease; their evaluation also considers the intended use, applicable regulatory oversight, and software version. Implementation research distinguishes the clinical effects of an intervention from the outcomes of introducing it into care, including adoption, feasibility, and sustainability [1]. The Nonadoption, Abandonment, Scale-up, Spread, and Sustainability framework further identifies interactions among the technology, intended users, organizations, and wider health system that influence continued use [2]. These distinctions explain why purchasing and sustained delivery require evidence on costs, workflow, support needs, and continued use in addition to clinical effects and regulatory status.

The published evidence supporting digital therapeutics varies across products and stages of development. Amyx et al. identified 436 studies for 45 digital therapeutics with US Food and Drug Administration (FDA) authorization decisions during 2019-2022; 46.7% had an identified pivotal randomized controlled trial [3]. The broad product definition limits extrapolation to any single class of behavioral application, and the 2019-2022 authorization cohort should not be interpreted as characterizing later or current authorization practice. The study nevertheless documents substantial variation in the amount and timing of publicly available evidence [3].

Integrated evidence planning addresses the relationship among development, regulatory evaluation, and

market access [4]. The Digital Therapeutics Real-World Evidence Framework organizes evidence generation around development decisions and iterative evaluation [5]. The National Institute for Health and Care Excellence (NICE) evidence standards framework includes clinical performance, economic impact, and deployment requirements for digital health technologies [6]. Together, these approaches provide a basis for examining the different questions posed by clinical, purchasing, and implementation decisions.

This review examines how evidence requirements differ across clinical evaluation, regulatory authorization, purchasing, implementation, and postmarket assessment in digital medicine. Commercialization is used here to describe the arrangements for purchasing, payment, delivery, and continued availability. The review focuses on evidence relevant to these arrangements; it does not assess company performance or test whether a particular evidence strategy produces commercial success. Selected clinical studies and independent delivery and economic evaluations are used to compare what each design establishes and what remains unresolved.

Review

Review methods

This targeted narrative review synthesizes public clinical studies, methodological literature, and regulatory documents. Initial searches used PubMed, publisher websites, FDA records, NICE guidance, web searches, and reference checking, without a publication start-date restriction. Supplemental searches through October 5, 2026, examined digital therapeutics and implementation, digital versus conventional pain care, and digital cognitive behavioral therapy for insomnia and healthcare costs. Search concepts also included integrated evidence planning, engagement, economic evaluation, and real-world evidence. This was an iterative, purposive search rather than a systematic search with exhaustive screening. The reporting approach was informed by the domains of the Scale for the Assessment of Narrative Review Articles, including a stated objective, description of the search, appropriate referencing, and scientific reasoning [7].

Sources were retained when they addressed an identifiable clinical, measurement, implementation, economic, or postmarket question and provided sufficient public information for the claim discussed. Four clinical examples were initially identified through the author's professional involvement and publications. Additional studies without the author's participation were selected to examine comparative care delivery, economic outcomes, and implementation. Author involvement was not sufficient for retention: tangential conceptual publications, company announcements, and study listings without results were excluded from the final synthesis. Both favorable and nonsignificant findings were considered. The examples are illustrative and do not represent a comprehensive sample of digital medicine.

For clinical examples, extraction focused on the population, comparator, sample size, treatment duration, principal outcome, analysis status, and limitations. Economic and implementation examples were examined for the setting, denominator, cost boundary, design, and limits of causal inference. Primary publications and official documents supported product-specific statements; bibliographic details were checked against publisher or indexed records. Where full text was unavailable, interpretation was restricted to the abstract and available regulatory documentation. No unpublished participant data, internal company analyses, confidential regulatory communications, or proprietary commercial information were used. No meta-analysis, formal risk-of-bias scoring, or duplicate independent screening was performed.

Intervention definition and intended decision

Complex-intervention guidance emphasizes the relationship among intervention components, context, and the research question [8]. For digital interventions, CONSORT-EHEALTH and the mobile health evidence reporting and assessment checklist identify reporting needs involving content, delivery, technology, human support, and use [9,10]. These features define what was evaluated. A study of software delivered with clinician contact therefore describes that combined intervention unless its design separately identifies the contribution of the software.

Clinical and purchasing evaluations address overlapping questions but apply different decision contexts. The Academy of Managed Care Pharmacy forum on digital therapeutics identified evidence assessment, product changes, privacy, professional awareness, and equitable access as concerns for adoption and coverage [11]. NICE similarly links the evidence assessment to the intended population, care pathway, value proposition, and deployment requirements [6]. These sources support describing the proposed use and delivery setting before interpreting an evidence package.

Measurement applications require a further distinction. The verification, analytical validation, and clinical validation framework, known as V3, evaluates whether a digital measurement technology is fit for its intended purpose [12]. V3+ adds usability validation [13]. These frameworks address the credibility and use of a measure; they do not by themselves demonstrate that a treatment delivered through the same device improves health.

Economic evaluation adds the analytical perspective, time horizon, resource use, and uncertainty, while

budget impact analysis adds the eligible population and expected uptake within a particular budget [14,15]. Table *1* maps these published considerations to illustrative adoption settings. The mapping is a synthesis of relevant questions, not a statement of universal reimbursement requirements.

Setting or intended use	Decision context	Evidence relevant to the decision	Interpretive boundary
Payer-funded therapeutic	Funding for a defined population	Comparative benefit, safety, eligible population, treatment mix, uptake, and budget impact [6,11,15].	Regulatory and purchasing assessments have different scopes; the relevant budget and setting must be specified [6,11,15].
Health-system service	Incorporation into an existing care pathway	Clinical outcomes, workflow, staff requirements, implementation cost, and sustainability [1,2,6].	Trial delivery conditions may differ from the local service [2,8].
Employer-funded care	A benefit offered to an eligible workforce	Clinical and work-related outcomes, eligible population, participation, and costs within the stated perspective [6,14,15].	Work-related consequences and healthcare expenditures are different economic outcomes [14].
Patient-funded access	Use under specified access and payment arrangements	Clinical evidence, usability, required support, and costs borne by the user [6,13,14].	The payer perspective does not necessarily capture patient costs [14].
Software with human clinical support	Delivery of a combined intervention	Description of components, contact, resources, and the comparator service [8-10].	An effect of the package does not isolate the software contribution [8].
Digital clinical measurement	Use of a measure for a defined clinical or research purpose	Verification, analytical validation, clinical validation, and usability validation [12,13].	Measurement validation does not itself estimate a therapeutic effect [12].

TABLE 1: Evidence considerations across intended adoption settings

Original synthesis of published evidence-planning, implementation, economic, and measurement methods. Citations in each row identify the supporting sources. The settings are illustrative and do not constitute universal coverage requirements or a ranking of evidence standards.

Complementary evidence domains

The reviewed frameworks separate clinical effects, implementation outcomes, economic consequences, and measurement validity [1,5,12,14,15]. Each domain supports a different inference. A technically valid measurement is not equivalent to a demonstrated treatment effect, and a favorable clinical comparison does not include an economic analysis unless costs and consequences were actually evaluated [12,14]. Table *2* summarizes these distinctions.

Domain	Evaluation question	Relevant evidence	Limit of inference
Clinical effects	What changes relative to the specified comparator?	Comparative clinical evaluation with defined outcomes and intervention components [8,9].	The inference applies to the comparison performed [8].
Intervention identity	What technology and support were delivered?	Version, content, access, infrastructure, and support descriptions [9,10].	A shared product name does not establish identical delivery [9,10].
Usability and use	Can users interact with the technology, and what use is recorded?	Usability validation, exposure measures, and attrition reporting [9,13].	Use measures and clinical outcomes are distinct [9].
Implementation	Can the intervention enter and remain in a service?	Adoption, feasibility, fidelity, cost, and sustainability measures [1,2].	Clinical effectiveness and implementation outcomes answer different questions [1].
Economic consequences	What costs and outcomes follow under the specified perspective?	Economic evaluation and budget impact analysis [14,15].	Modeled assumptions must be distinguished from observed inputs [14,15].
Monitoring	How does evidence evolve during continued use?	Repeated evaluation linked to development decisions and intervention description [5,9].	Monitoring data require a defined question and interpretable measures [5,9].

TABLE 2: Evidence domains and limits of interpretation

Original synthesis. The domains remain separate; no composite score, relative weighting, or validated readiness threshold is proposed. Each row cites the methodological basis for its content.

Dependencies between domains are explicit in established methods. Budget impact estimates depend on the population expected to receive treatment, the delivery arrangements, and the resources included [15]. Complex-intervention evaluation considers how context and implementation affect the intervention being studied [8]. Consequently, transferring findings to a new setting involves examining both the clinical comparison and the delivery conditions that contributed to it.

Intervention reporting is central to that assessment. CONSORT-EHEALTH addresses details such as versions, access arrangements, prompts, human assistance, and actual use; mERA adds attention to infrastructure and implementation context [9,10]. In this review, the domains remain separate rather than being combined into a product score. No weighting system or cross-product ranking was developed.

Comparator choice and applicability

A review of controls for behavioral and cognitive digital therapeutics describes trade-offs among waiting-list, usual-care, digital sham, and active-treatment comparisons [16]. A digital control may account for some expectations and device interaction, whereas an active-care comparator can address a treatment choice more directly. Interpretation depends on what both groups received; different comparators need not estimate the same treatment effect [16].

The Pragmatic-Explanatory Continuum Indicator Summary-2 tool examines whether trial choices match the intended purpose across domains including eligibility, recruitment, setting, organization, delivery, adherence, follow-up, outcomes, and analysis [17]. It provides a way to assess the relationship between trial conditions and intended care. Pragmatism concerns design and applicability; it is not a substitute for an appropriate comparison or reliable outcome assessment.

Reporting the comparator requires the same attention to support, content, access, and exposure as reporting the intervention [9,10,16]. Unequal clinician contact or technical assistance may be part of the treatment strategy being compared. Such differences limit an interpretation that attributes the entire observed contrast to one software component [8,16].

The following examples retain each report's outcome, analysis population, and status as a primary or exploratory analysis. They were selected to examine distinct inferential problems rather than to compare products directly and summarize the clinical evidence and separate the reported findings from the implications drawn in this review.

Aphasia rehabilitation: intervention and comparator

Constant Therapy-Research (The Learning Corp., Newton, MA, USA) was assessed in a virtual trial of post-stroke aphasia rehabilitation. Thirty-six participants were randomized, and 32 completed the study. Ten weeks of application-based practice were compared with workbook practice. The adjusted group-by-time estimate for the Western Aphasia Battery-Revised Aphasia Quotient favored the application by 6.43 points ($p < 0.01$) [18].

The report evaluated a research intervention with remote assessments and check-ins. Its small sample, short duration, and workbook comparator constrain interpretation [18]. The result does not constitute a comparison with an otherwise equivalent clinician-delivered rehabilitation service.

For adoption, the relevant inference is therefore bounded by the evaluated practice and support arrangements. Extending that inference to a different service requires attention to intervention components and context, as emphasized in complex-intervention methods [8].

Migraine prevention: adjunctive treatment and exploratory pooling

The FDA De Novo order for CT-132 (Click Therapeutics, Inc., New York, NY, USA), dated April 11, 2025, specifies preventive treatment of episodic migraine in adults aged 18 years and older as an adjunct to acute and/or other preventive migraine treatments [19].

A post hoc integrated analysis combined ReMMi-D and ReMMiD-C, totaling 678 participants. At 12 weeks, the between-group difference in monthly migraine days was -0.86 days (95% confidence interval (CI), -1.43 to -0.30; nominal $p = 0.003$). The smaller study's individual CI crossed zero; the predominantly White and female sample and short follow-up limit generalizability and durability conclusions [20].

The pooled finding describes adjunctive treatment and remains exploratory [19,20]. It does not establish medication replacement or a prospectively confirmed effect in every subgroup. Any economic extrapolation from symptom reduction requires separately identified resource-use inputs and assumptions [14,15].

Major depressive disorder: primary and supportive analyses

Rejoyn, developed as CT-152 (Otsuka America Pharmaceutical, Inc., Rockville, MD, USA, with Click Therapeutics, Inc., New York, NY, USA), was studied as an adjunct to antidepressant medication in Mirai. Of 386 randomized participants, 354 entered the primary modified intention-to-treat analysis. The six-week Montgomery-Asberg Depression Rating Scale difference was -1.78 points ($p = 0.0568$); the supportive all-randomized analysis yielded -2.12 points ($p = 0.0211$) [21].

The primary analysis was nonsignificant. The FDA record describes a primary significance threshold of 0.049 and clearance for adults aged 22 years and older receiving antidepressant medication, as an adjunct to clinician-managed outpatient care [22]. The supportive analysis remains distinct from the primary finding [21,22].

The FDA documentation also states that the device does not monitor clinical status or alert the prescriber [22]. This defines a delivery limitation relevant to clinical oversight; it does not establish the effectiveness of any particular monitoring arrangement.

Schizophrenia: efficacy and conditions of use

CONVOKE evaluated CT-155 (Click Therapeutics, Inc., New York, NY, USA, and Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany) with background antipsychotic treatment. The trial randomized 457 adults at 66 US sites. The 16-week between-group difference on the Clinical Assessment Interview for Negative Symptoms, Motivation and Pleasure subscale was -2.6 points (95% CI, -4.0 to -1.2; $p < 0.001$) [23].

Eligibility required smartphone and internet access, stable housing, and stable antipsychotic treatment. Engagement occurred with routine study-site contact, although the report specifies that sites did not provide engagement-enhancing or therapeutic support [23]. These conditions constrain extrapolation to less stable populations or different delivery settings.

The study reported no serious adverse event judged treatment-related [23]. That finding describes the observed trial experience and does not establish absence of uncommon harms during broader use.

Across the four reports, the comparator, analysis population, and follow-up define different evidentiary boundaries [18,20,21,23]. Table *3* retains those distinctions. Effect sizes measured on different disease-specific scales are not treated as a common ranking of clinical or commercial value.

Intervention	Design and population	Principal finding	Limit relevant to interpretation
Constant Therapy-Research	36 randomized; 32 completers; 10 weeks; workbook comparator [18].	Adjusted group-by-time WAB-AQ estimate: 6.43 points favoring the application; p < 0.01 [18].	Small trial of a research intervention; the comparator does not establish equivalence to clinician-delivered therapy [18].
CT-132	Post hoc pooling of two adjunctive migraine trials; N = 678 [20].	Monthly migraine-day difference: -0.86; 95% CI, -1.43 to -0.30; nominal p = 0.003 [20].	Exploratory pooling; short follow-up and limited demographic diversity [20].
CT-152/Rejoyn	386 randomized; primary modified intention-to-treat N = 354; six-week outcome [21].	MADRS difference: primary -1.78, p = 0.0568; supportive all-randomized -2.12, p = 0.0211 [21].	The supportive finding does not replace the nonsignificant primary analysis [21,22].
CT-155	457 randomized; 16 weeks; digital control with background antipsychotic treatment [23].	CAINS-MAP difference: -2.6; 95% CI, -4.0 to -1.2; p < 0.001 [23].	Selected stable population, technology-access requirements, and routine study contact [23].

TABLE 3: Selected clinical studies and their evidentiary boundaries

Original extraction and synthesis of the cited reports. Outcomes, controls, and populations differ; this table is not a comparative product ranking.

WAB-AQ: Western Aphasia Battery-Revised Aphasia Quotient; MADRS: Montgomery-Asberg Depression Rating Scale; CAINS-MAP: Clinical Assessment Interview for Negative Symptoms, Motivation and Pleasure subscale; CI: confidence interval

Digital pain care: evaluation of a combined service

Cui et al. randomized 140 adults with chronic low back pain to eight weeks of digital care or conventional in-person physiotherapy. Digital care combined exercise, education, cognitive behavioral content, and remote physiotherapist support [24]. This comparison evaluates care packages rather than an isolated software ingredient.

Both groups improved in disability, with no statistically significant between-group difference in change (median difference, -0.55; 95% CI, -2.42 to 5.81; p = 0.412). Dropout was lower in the digital group: 11/70 (15.7%) versus 24/70 (34.3%; p = 0.019) [24].

The study lacked participant and therapist blinding and an externally blinded outcome adjudicator; follow-up was short [24]. The reported between-group test did not demonstrate superiority; this result alone does not establish equivalence [24]. The result informs a comparison of delivery models but does not isolate the contribution of software or establish cost savings.

For nonrandomized service evaluations, the target-trial approach specifies eligibility, treatment strategies, assignment, time zero, follow-up, outcomes, and analysis [25]. This structure makes the intended causal comparison explicit. Residual confounding and incomplete measurement remain limitations when the required conditions cannot be reproduced in observational data [25].

Engagement, exposure, and population reach

Perski et al. describe engagement with digital behavior-change interventions as comprising observable use and subjective experience, influenced by content and context [26]. Application openings, completed activities, and perceived usefulness therefore represent different measurements. A report's definition of engagement determines what its results can establish [26].

Digital-intervention reporting guidance distinguishes use measures, attrition, and clinical outcomes [9]. More application activity does not by itself identify a treatment effect. In particular, restricting an outcome analysis to participants who complete treatment changes the analyzed population relative to the randomized comparison [9,25].

The target-trial framework highlights why treatment assignment, adherence, and follow-up cannot be treated interchangeably in causal analyses [25]. Associations between greater use and improvement may reflect differences in participants as well as intervention exposure. Such associations can inform further research but do not independently validate engagement as a surrogate clinical endpoint.

RE-AIM organizes evaluation around reach, effectiveness, adoption, implementation, and maintenance [27]. Its population perspective complements participant-level engagement analysis. In this review's synthesis, separating patients offered treatment, those initiating it, and those with observed outcomes makes the

denominator and missing follow-up explicit; this is an application of reach and reporting principles [9,27].

Economic evidence and the purchasing perspective

CHEERS 2022 specifies transparent reporting of economic evaluations, including perspective, comparators, time horizon, costs, outcomes, assumptions, and uncertainty [14]. ISPOR budget impact guidance addresses affordability within a specified population and budget [15]. Cost-effectiveness and budget impact answer different questions and can lead to different conclusions for the same intervention.

For digital technologies, NICE includes acquisition and implementation resources in the economic assessment [6]. The relevant cost boundary may include software access, equipment, staff training, integration, and ongoing support. Whether these costs are bundled or incurred separately affects their measurement; assumptions about uptake also affect the budget estimate [6,15].

A purposive US public-document study of prescription digital therapeutics identified heterogeneous financial terms and restrictive coverage policies alongside conditional clinical eligibility pathways [28]. These selected documents illustrate the distinction between documented access arrangements and realized patient access, without establishing national coverage prevalence [28].

The relationship between a clinical endpoint and an economic outcome must be specified within the analysis. CHEERS requires reporting the model, inputs, assumptions, and uncertainty; budget impact guidance considers the expected treatment mix and associated resource use [14,15]. A symptom-scale result therefore cannot be relabeled as an observed reduction in medical spending when resource use was not measured.

Sampson et al. examined the rollout of Sleepio using an interrupted time series of primary care costs in England. The analysis included 10,705 patients and reported an estimated total reduction of £71,027 over 65 weeks [29]. The study illustrates evaluation of costs using routine-care records rather than deriving a savings claim solely from a symptom endpoint.

The Sleepio evaluation was nonrandomized and assessed population-level costs; its authors distinguished the analysis from a full cost-effectiveness evaluation and did not include all rollout costs [29]. These boundaries limit transfer of the estimated reduction to a different payer, price, or delivery model. Reporting the study as an association with primary care costs preserves those limits.

Table *4* distinguishes clinical, utilization, cost-effectiveness, and budget impact claims. The table summarizes methodological requirements and interpretive limits; it does not prescribe coverage decisions or payment thresholds [14,15,25,29].

Claim or question	Evidence required	Major analytic consideration	Boundary of interpretation
Clinical improvement	Defined clinical outcome and comparator [8,9].	Outcome measurement and the comparison performed [8,9].	A clinical outcome is not an observed expenditure outcome [14].
Reduced healthcare utilization	Comparative resource-use data with aligned eligibility and follow-up [25].	Confounding, time zero, and outcome capture [25].	An uncontrolled before-after change does not alone identify the intervention effect [25].
Cost-effectiveness	Costs and consequences relative to alternatives [14].	Perspective, horizon, model, and parameter uncertainty [14].	A favorable estimate depends on the stated inputs and assumptions [14].
Budget impact	Population, uptake, treatment mix, and costs for the budget holder [15].	Adoption scenarios, resource availability, and sensitivity analysis [15].	Cost-effectiveness does not establish affordability [14,15].
Primary care cost association	Interrupted time series of routine costs, illustrated by Sleepio [29].	Pre-existing trends, concurrent changes, and the population analyzed [29].	The published analysis was not a full cost-effectiveness evaluation and did not include all rollout costs [29].
Delivery cost	Resources needed to deploy and support the technology [6,15].	Acquisition, implementation, and ongoing resource boundaries [6,15].	Software price alone may not represent the full resources assessed [6,15].

TABLE 4: Economic claims, evidence requirements, and interpretation

Original methodological synthesis. The table distinguishes types of claims rather than prescribing purchasing decisions. Citations identify the sources for specific design requirements and the empirical example.

Implementation studies and adoption in routine care

Effectiveness-implementation hybrid designs combine evaluation of clinical effectiveness and implementation, with different designs emphasizing the intervention, an implementation strategy, or both [30]. This provides a methodological basis for studying delivery alongside clinical outcomes. It does not make an uncontrolled rollout a randomized test of either effectiveness or implementation strategy.

In a two-clinic primary care pilot, Mogk and colleagues studied prescription digital therapeutics for substance use disorders. Of 18 patients prescribed an application, 10 activated it and eight completed at least one module. Clinicians described competing priorities, technical problems, and complex workflows; staff turnover also affected implementation [31].

The pilot's size, setting, and sequential introduction of support strategies limit causal and population-level conclusions [31]. Nevertheless, the observed losses between prescription, activation, and use identify delivery questions that are distinct from efficacy. Proctor's implementation outcomes and the NASSS framework provide complementary constructs for examining such service and organizational barriers [1,2].

Table *5* organizes implementation questions by stage, using established planning and implementation methods [4,5,30,31]. The stages summarize the literature and are not a validated launch-readiness instrument.

Stage	Question	Relevant evaluation	Interpretive boundary
Intervention development	What intervention is being evaluated and in what context?	Component, context, feasibility, and user assessment [5,8,13].	Feasibility does not independently establish clinical benefit [1,8].
Clinical evaluation	What effect is estimated by the selected comparison?	Defined comparator, outcomes, analysis, and intervention reporting [9,16,17].	Applicability depends on the trial population and delivery conditions [17].
Adoption assessment	What costs and delivery conditions accompany the proposed use?	Implementation measures and relevant economic evaluation [1,6,14,15].	The conclusion is conditional on setting, perspective, and assumptions [14,15].
Initial implementation	Where do referral, initiation, and treatment delivery fail?	Hybrid designs or focused implementation studies; reach and fidelity [27,30,31].	The cited pilot demonstrates local barriers, not a universal uptake rate [31].
Continued use	Do the evaluated intervention and delivery conditions remain applicable?	Iterative monitoring, intervention description, and sustainability assessment [2,5,9].	These methods do not constitute a validated prediction of commercial survival [2,5].

TABLE 5: Evaluation questions across development and implementation

Original synthesis of published development and implementation methods. Stages may overlap or recur. This table is not a validated launch-readiness instrument, organizational prescription, or prediction of commercial outcomes.

Figure *1* provides an original conceptual synthesis of the connection between a decision, unresolved uncertainty, and evaluation. It is informed by integrated evidence planning and decision-based development [4,5,8].

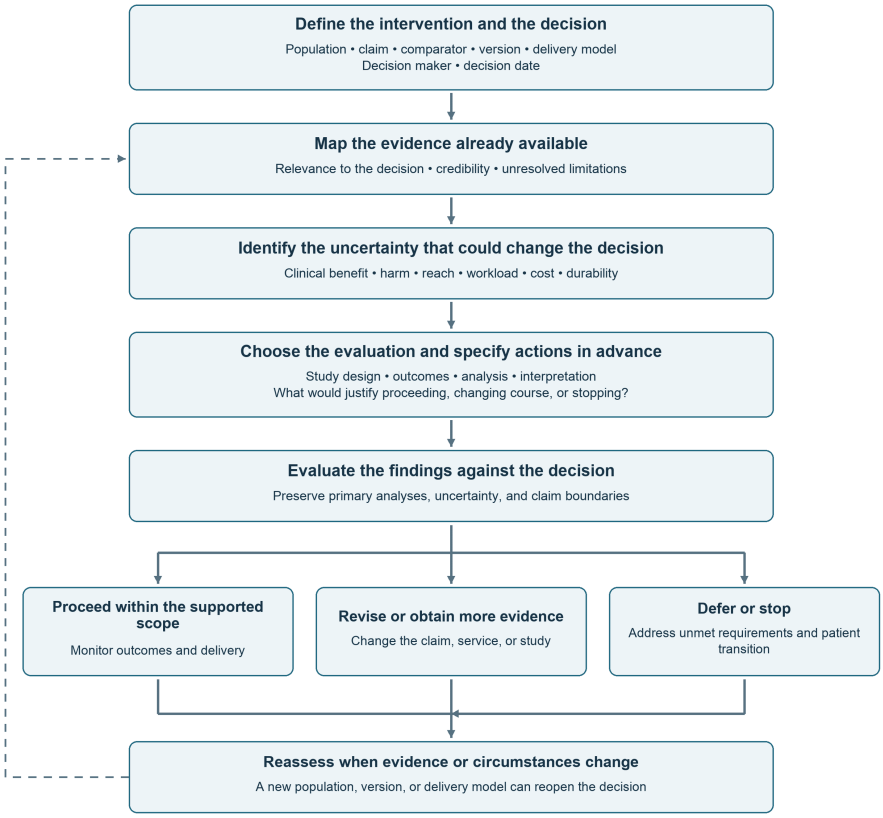


FIGURE 1: Connecting an adoption decision to evidence and action

Original conceptual synthesis informed by integrated evidence planning, decision-based digital therapeutic development, and complex-intervention evaluation [4,5,8]. The arrows organize the relationship among a defined use, unresolved uncertainty, evaluation, and reassessment. They do not represent an empirically validated algorithm or demonstrate that the process improves adoption or commercial outcomes. No patient data are depicted.

Figure *2* illustrates denominators along a care pathway. It applies reach, implementation, and digital-intervention reporting concepts; the schematic contains no observed patient data [1,9,27,31].

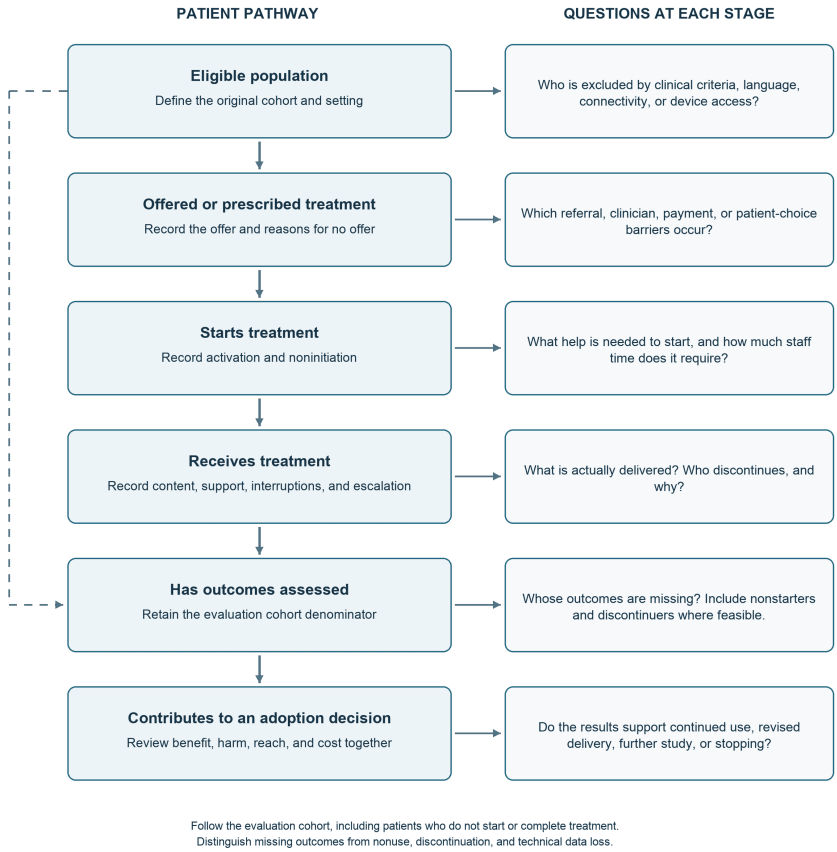


FIGURE 2: Following patients from eligibility to outcomes in routine care

Original conceptual schematic informed by implementation outcomes, digital-intervention reporting, RE-AIM, and an empirical primary care implementation study [1,9,27,31]. Each stage requires a stated denominator. Outcome ascertainment is distinct from treatment completion, and follow-up may include people who discontinue. Arrows describe a service pathway, not a formula for clinical benefit. No numerical patient observations or source figures are reproduced.

Postmarket evidence, safety, and software changes

The Digital Therapeutics Real-World Evidence Framework includes monitoring within an iterative development process [5]. CONSORT-EHEALTH also emphasizes intervention description and use [9]. Together, these approaches support interpreting postmarket observations in relation to the population, version, delivery conditions, and outcomes actually recorded.

The FDA identifies important limitations of the Manufacturer and User Facility Device Experience database: underreporting, incomplete or inaccurate information, uncertain causation, and lack of device-use denominators. These reports cannot independently establish incidence or support comparisons of adverse-event rates between devices [32].

Accordingly, sparse reports do not establish absence of harm, and an individual report does not establish that the device caused the event [32]. The CT-132 De Novo order identifies ineffective treatment, software failure, and use error among risks addressed through controls [19]. These examples support distinguishing technical events, treatment outcomes, and causal attribution when describing safety evidence.

FDA guidance on software changes addresses circumstances in which modification of an existing device requires a new 510(k) submission [33]. The guidance concerns a regulatory assessment of change. Separately, intervention-reporting methods require enough detail to understand the version and delivery arrangements evaluated in a study [9,10].

The implication for evidence interpretation is that a product name alone does not describe all relevant intervention features. When content, user interaction, or support differs from the evaluated intervention, applicability depends on the nature of the change and the evidence available [8-10,33]. This review does not assess undocumented changes to any named product.

Digital measurement and the scope of a clinical claim

V3 distinguishes verification of the sensor or technology, analytical validation of the derived measurement, and clinical validation for the intended use [12]. These are complementary steps. Evidence supporting one measurement task does not automatically establish performance in another population, setting, or purpose [12].

V3+ adds usability validation, recognizing that obtaining interpretable data depends on how intended users interact with the technology [13]. This is relevant both to measurement tools and to trials that rely on digital endpoints, although the validity of a measure remains distinct from the effect of an intervention.

For the present review, measurement evidence is included to clarify this boundary within digital medicine. Demonstrating that a measure is fit for a clinical purpose does not independently establish that acting on it improves outcomes; that is an additional evaluation question [8,12,13]. No diagnostic or therapeutic performance is inferred from a commercial partnership or product announcement.

Synthesis across evidence and adoption decisions

The literature supports treating an evidence package as a set of complementary evaluations. Integrated planning and decision-based development connect research questions across development, while implementation and economic methods specify additional outcomes and perspectives [1,4,5,14,15]. The synthesis presented here applies those established approaches rather than introducing a new validated framework.

The clinical examples show different limits of inference: a specific rehabilitation comparator, exploratory pooling, a nonsignificant primary analysis with a favorable supportive analysis, and efficacy in a selected stable population [18,20,21,23]. These differences remain material when describing the evidence to clinicians or purchasers.

The independent service, cost, and implementation studies extend the discussion beyond efficacy. They evaluate combined care delivery, a change in primary care expenditure, and losses between prescribing and treatment initiation [24,29,31]. None alone answers all three questions. Their relevance lies in the distinct uncertainty each design examines.

Across these examples, matching evidence to a commercialization decision means specifying the proposed use, identifying the comparison that was actually made, and assessing whether the outcome and setting address the decision at hand [4-6,17]. This is an interpretation of the reviewed methods, not evidence that this process improves revenue, accelerates launch, or predicts commercial survival.

The principal research gap identified by this synthesis is the connection between clinical effects and delivery, resource use, and sustained access in the same intended setting. Hybrid designs and explicitly specified observational comparisons offer established approaches to studying parts of that connection [25,30]. Their suitability depends on the uncertainty and the feasibility of a credible comparison.

Limitations

This review used a purposive narrative search and did not estimate the completeness of retrieval. Author participation in several clinical programs informed initial example selection and creates potential selection and interpretive bias. Additional studies without author participation broaden the evidence considered but do not eliminate that bias. The review is not an independent comparative assessment of the named products.

The examples are concentrated in behavioral interventions, neurological and psychiatric care, and US or UK settings. Treatment periods, comparators, outcomes, and analysis populations differ [18,20,21,23,24,29,31]. These differences preclude a common effect estimate or inference that one evidence package is superior across conditions.

Not all sources were available in full text, and no formal risk-of-bias assessment was performed. Publicly available evidence cannot establish the content of internal investment decisions, confidential negotiations, or unpublished company analyses. The review therefore makes no causal claim about commercial success or failure and does not interpret missing public evidence as proof that evidence does not exist.

The tables and figures are original syntheses of published methods and selected findings. They have not been evaluated as prediction models, scoring systems, or decision instruments. Prospective assessment would be required before claiming that their use improves study selection, adoption decisions, or patient outcomes.

Conclusions

Evidence generation and commercialization in digital medicine intersect at decisions about clinical use, purchasing, delivery, and continued availability. The reviewed literature distinguishes clinical benefit, regulatory status, engagement, implementation, economic consequences, and measurement validity. Evidence in one domain can inform another, but the required inference depends on the comparator, population, intervention version, support model, and outcome assessed.

The clinical examples demonstrate why primary and supportive analyses, exploratory pooling, and trial conditions must remain visible when evidence is translated into an adoption discussion. Independent care-delivery, cost, and implementation studies show the additional questions that arise in routine settings. An integrated evidence package makes these relationships explicit while preserving the limits of each source. The resulting synthesis can organize evaluation around an intended decision; whether it improves adoption or commercial outcomes remains an empirical question.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Shaheen E. Lakhan

Acquisition, analysis, or interpretation of data: Shaheen E. Lakhan

Drafting of the manuscript: Shaheen E. Lakhan

Critical review of the manuscript for important intellectual content: Shaheen E. Lakhan

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** Lakhan SE declare(s) employment and stock/stock options from Click Therapeutics, Inc. He is a former Chief Medical and Scientific Officer. Lakhan SE declare(s) employment from Constant Therapy. He is a former VP of R&D and Head of Clinical. Lakhan SE declare(s) personal fees and employment from Fern Health. He is a Former Head of Clinical. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

ChatGPT (OpenAI, San Francisco, California, USA; GPT-6 Astra) was used to assist with manuscript drafting and revision, including improving clarity, organization, and flow; correcting grammar, spelling, and punctuation; proofreading; and checking reference details, formatting, and citation consistency. The sole author reviewed and revised the manuscript and assumes full responsibility for the accuracy and integrity of the final content, including the data, analyses, interpretations, references, and conclusions.

References

1. Proctor E, Silmere H, Raghavan R, et al.: Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health. 2011, 38:65-76. 10.1007/s10488-010-0319-7
2. Greenhalgh T, Wherton J, Papoutsi C, et al.: Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res. 2017, 19:e367. 10.2196/jmir.8775
3. Amyx M, Alebouyeh F, Phi NT, Ravaud P, Tran VT: Mapping the evidence supporting FDA-authorized digital therapeutics, before and after authorization. NPJ Digit Med. 2026, 9:618. 10.1038/s41746-026-02737-9
4. Lee WC, Blanchette C, Pokras S, Shaikh J, Miller J: The evolution and future of integrated evidence planning . Expert Rev Pharmacoecon Outcomes Res. 2025, 25:855-62. 10.1080/14737167.2025.2497876
5. Kim M, Patrick K, Nebeker C, et al.: The digital therapeutics real-world evidence framework: an approach for guiding evidence-based digital therapeutics design, development, testing, and monitoring. J Med Internet Res. 2024, 26:e49208. 10.2196/49208
6. Evidence standards framework for digital health technologies: how to meet the standards . (2022). Accessed: October 4, 2026: https://www.nice.org.uk/corporate/ecd7/chapter/how-to-meet-the-standards.
7. Baethge C, Goldbeck-Wood S, Mertens S: SANRA: a scale for the quality assessment of narrative review articles. Res Integr Peer Rev. 2019, 4:5. 10.1186/s41073-019-0064-8
8. Skivington K, Matthews L, Simpson SA, et al.: A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ. 2021, 374:n2061. 10.1136/bmj.n2061
9. Eysenbach G: CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res. 2011, 13:e126. 10.2196/jmir.1923

10. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, Labrique A: Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ. 2016, 352:i1174. 10.1136/bmj.i1174
11. AMCP partnership forum: the evolving role of digital therapeutics . J Manag Care Spec Pharm. 2022, 28:804-10. 10.18553/jmcp.2022.22093
12. Goldsack JC, Coravos A, Bakker JP, et al.: Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med. 2020, 3:55. 10.1038/s41746-020-0260-4
13. Bakker JP, Barge R, Centra J, et al.: V3+ extends the V3 framework to ensure user-centricity and scalability of sensor-based digital health technologies. NPJ Digit Med. 2025, 8:51. 10.1038/s41746-024-01322-2
14. Husereau D, Drummond M, Augustovski F, et al.: Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. BMJ. 2022, 376:e067975. 10.1136/bmj-2021-067975
15. Sullivan SD, Mauskopf JA, Augustovski F, et al.: Budget impact analysis-principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value Health. 2014, 17:5-14. 10.1016/j.jval.2013.08.2291
16. Lutz J, Offidani E, Taraboanta L, Lakhan SE, Campellone TR: Appropriate controls for digital therapeutic clinical trials: a narrative review of control conditions in clinical trials of digital therapeutics (DTx) deploying psychosocial, cognitive, or behavioral content. Front Digit Health. 2022, 4:823977. 10.3389/fdgth.2022.823977
17. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M: The PRECIS-2 tool: designing trials that are fit for purpose. BMJ. 2015, 350:h2147. 10.1136/bmj.h2147
18. Braley M, Pierce JS, Saxena S, et al.: A virtual, randomized, control trial of a digital therapeutic for speech, language, and cognitive intervention in post-stroke persons with aphasia. Front Neurol. 2021, 12:626780. 10.3389/fneur.2021.626780
19. US Food and Drug Administration: De Novo classification order: CT-132, DEN240064 . (2025). Accessed: October 4, 2026: https://www.accessdata.fda.gov/cdrh_docs/pdf24/DEN240064.pdf.
20. Lakhan SE, Petrova M, Snipes C, et al.: Integrated analysis of two randomized controlled trials of a digital therapeutic for episodic migraine prevention. NPJ Digit Med. 2026, 9:573. 10.1038/s41746-026-02675-6
21. Rothman B, Slomkowski M, Speier A, et al.: A digital therapeutic (CT-152) as adjunct to antidepressant medication: a phase 3 randomized controlled trial (the Mirai study). J Affect Disord. 2025, 388:119409. 10.1016/j.jad.2025.119409
22. US Food and Drug Administration: 510(k) premarket notification: Rejoyn, K231209 . (2024). Accessed: October 4, 2026: https://www.accessdata.fda.gov/cdrh_docs/pdf23/K231209.pdf.
23. Pratap A, Lakhan SE, Snipes C, et al.: A digital therapeutic intervention for negative symptoms of schizophrenia: a randomized clinical trial. JAMA Netw Open. 2026, 9:e2635753. 10.1001/jamanetworkopen.2026.35753
24. Cui D, Janela D, Costa F, et al.: Randomized-controlled trial assessing a digital care program versus conventional physiotherapy for chronic low back pain. NPJ Digit Med. 2023, 6:121. 10.1038/s41746-023-00870-3
25. Hernán MA, Robins JM: Using big data to emulate a target trial when a randomized trial is not available . Am J Epidemiol. 2016, 183:758-64. 10.1093/aje/kwv254
26. Perski O, Blandford A, West R, Michie S: Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med. 2017, 7:254-67. 10.1007/s13142-016-0453-1
27. Glasgow RE, Vogt TM, Boles SM: Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health. 1999, 89:1322-7. 10.2105/ajph.89.9.1322
28. Lakhan SE: Coverage policies, benefit routes, and public financial terms for prescription digital therapeutics in the United States: a cross-sectional health policy and market access study. Cureus. 2026, 18:e117911. 10.7759/cureus.117911
29. Sampson C, Bell E, Cole A, Miller CB, Marriott T, Williams M, Rose J: Digital cognitive behavioural therapy for insomnia and primary care costs in England: an interrupted time series analysis. BJGP Open. 2022, 6:e0146. 10.3399/BJGPO.2021.0146
30. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C: Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care. 2012, 50:217-26. 10.1097/MLR.0b013e3182408812
31. Mogk J, Idu AE, Bobb JF, et al.: Prescription digital therapeutics for substance use disorder in primary care: mixed methods evaluation of a pilot implementation study. JMIR Form Res. 2024, 8:e59088. 10.2196/59088
32. US Food and Drug Administration: about Manufacturer and User Facility Device Experience (MAUDE) database. (2025). Accessed: October 5, 2026: https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-faciliti....
33. US Food and Drug Administration: deciding when to submit a 510(k) for a software change to an existing device. (2017). Accessed: October 4, 2026: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/deciding-when-submit-510k-software-change-ex....