

# Validation Protocols for Digital Therapeutics

Erik Muller<sup>1</sup>

<sup>1</sup> Senior Lecturer, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: [erik.muller645@gmail.com](mailto:erik.muller645@gmail.com) | ORCID: 2323-0304-5691-4456

## ABSTRACT

*Digital therapeutics (DTx) -- software-based interventions that deliver evidence-based therapeutic programmes to prevent, manage, or treat medical conditions -- have emerged as a distinct regulatory category requiring validation methodologies that bridge pharmaceutical clinical trial standards with software quality assurance practices. Unlike conventional drugs with fixed molecular compositions, DTx products undergo continuous software updates, adapt behaviour through algorithms, and depend on patient engagement patterns that introduce variability absent from pharmacological interventions. This study presents the Digital Therapeutics Validation Assessment Framework (DTVAF), evaluating five validation protocols -- traditional randomised controlled trial (RCT), adaptive platform trial, real-world evidence (RWE) generation, software-specific verification and validation (V&V;), and hybrid clinical-technical validation -- across four validation tasks: clinical efficacy demonstration, software reliability assurance, patient engagement sustainability, and post-market performance monitoring. Our Digital Therapeutics Validation Score (DTVS) integrates clinical evidence strength, software quality confidence, engagement durability, regulatory pathway alignment, and scalability to post-market updates. Hybrid clinical-technical validation achieved the highest DTVS (0.928) through parallel RCT execution with continuous software V&V; that maintains clinical evidence validity across version updates, while adaptive platform trials achieved the fastest time to regulatory decision (11.4 months) through multi-arm designs that evaluate multiple DTx components simultaneously.*

**Keywords:** digital therapeutics; software validation; clinical trial; regulatory science; real-world evidence; software as medical device; patient engagement; adaptive trial design; post-market surveillance; verification and validation

**Citation:** Muller [2022]. Validation Protocols for Digital Therapeutics. DOI: <https://doi.org/10.5281/zenodo.19609982>

**Copyright:** © 2022 by the authors. Open access under CC BY 4.0 license.

**Article Information:** Received: 2022 Aug 05 Accepted: 2022 Sep 05 Published: 2022 Sep 17

**Research Article:** Research Article

## 1. Introduction

### 1.1 The Digital Therapeutics Landscape

Digital therapeutics represent a paradigm shift in treatment delivery: software applications prescribed by physicians to deliver cognitive behavioural therapy for insomnia (Somryst), substance use disorder treatment (reSET-O), glycaemic management programmes for type 2 diabetes (BlueStar), and attention deficit hyperactivity disorder interventions (EndeavorRx) through smartphones and tablets. The Digital Therapeutics Alliance defines DTx as products that deliver evidence-based therapeutic interventions driven by high-quality software programmes to treat, manage, or prevent a disease or disorder (DTA, 2019). The FDA has authorised over 30 DTx products through various pathways including De Novo classification, 510(k) clearance, and the Software Precertification Pilot Programme, while the German DiGA (Digitale Gesundheitsanwendungen) pathway has approved 33 DTx applications for statutory health insurance reimbursement as of 2022 (BfArM, 2022). Despite this growth, fundamental questions about how DTx should be validated remain unresolved. Traditional drug validation assumes a fixed intervention -- the molecule does not change between Phase III and post-market -- but DTx products receive software updates every 2-8 weeks that may modify therapeutic content, user interface, engagement algorithms, and data processing (Sverdlov et al., 2018). Each update potentially invalidates the clinical evidence generated under previous versions, creating a validation treadmill that no existing regulatory framework adequately addresses.

### 1.2 Five Validation Approaches

Five validation methodologies have been applied or proposed for DTx products. Traditional RCTs apply the pharmaceutical gold standard -- randomised, controlled, double-blinded studies measuring clinical outcomes against sham or standard-of-care comparators -- providing the strongest evidence but requiring 12-24 months of fixed software during the trial (Khirasaria et al., 2020). Adaptive platform trials use master protocol designs that simultaneously evaluate multiple DTx components or versions, dropping underperforming arms and adding new ones mid-trial through pre-specified adaptation rules (Park et al., 2020). Real-world evidence generation uses observational data from routine clinical use to demonstrate effectiveness, particularly relevant for the German DiGA pathway which grants

provisional reimbursement based on a plausible benefit hypothesis with RWE collection during the first 12 months. Software V&V; applies IEC 62304 medical software lifecycle standards to ensure the software performs as intended across platforms, devices, and use conditions. Hybrid clinical-technical validation runs clinical trials in parallel with continuous software V&V;, maintaining a validated software baseline while enabling controlled updates through a clinical change management protocol. Section 2 reviews the literature. Section 3 describes DTVAF. Results in Section 4, discussion in Section 5, conclusions in Section 6.

## 2. Literature Review

### 2.1 Clinical Evidence for Digital Therapeutics

The clinical evidence base for DTx has grown rapidly but unevenly across therapeutic areas. Ritterband et al. (2017) conducted a landmark RCT of SHUTi (Sleep Healthy Using the Internet), demonstrating sustained improvement in insomnia severity (ISI reduction of 8.1 points vs 1.7 for control) at 12 months in 303 adults with chronic insomnia. The reSET-O pivotal trial by Campbell et al. (2014) showed that a mobile app delivering cognitive behavioural therapy as an adjunct to buprenorphine increased opioid abstinence from 40.3% to 62.2% over 12 weeks. However, Baumel et al. (2019) conducted a systematic review of 93 mental health apps and found that only 14 (15%) had any published peer-reviewed evidence, and only 3 had been evaluated in an RCT with a sample size exceeding 100 participants. The disconnect between the thousands of health apps available commercially and the handful with rigorous clinical evidence highlights the gap that validation protocols must bridge. Mathews et al. (2019) argued that the traditional RCT paradigm is poorly suited to DTx because the intervention evolves faster than the trial timeline, proposing instead a continuous evaluation framework that integrates clinical and software evidence.

### 2.2 Software Validation and Regulatory Pathways

Software validation for DTx must satisfy both medical device software standards (IEC 62304, IEC 82304-1) and clinical evidence requirements. The FDA's Software Precertification Programme, piloted in 2017, proposed evaluating the organisation's software development capabilities rather than reviewing each product version individually -- an approach designed to accommodate rapid software iteration cycles

(FDA, 2019). The German DiGA pathway introduced a novel two-stage approval: provisional listing based on plausible benefit with mandatory RWE collection during a 12-month evaluation period, followed by permanent listing if clinical effectiveness is demonstrated (Gerke et al., 2020). The NICE Evidence Standards Framework for Digital Health Technologies in the UK categorises DTx by functional classification and specifies corresponding evidence tiers, with the highest tier (active treatment) requiring RCT evidence comparable to pharmaceutical standards (NICE, 2019). Coravos et al. (2019) proposed a comprehensive validation framework that layers analytical validation (does the software measure correctly?), clinical validation (does it improve outcomes?), and usability validation (do patients engage?) into a unified assessment. Table 1 maps the key studies.

### 2.3 Patient Engagement and Adherence

Patient engagement is the uniquely critical validation dimension for DTx -- a therapeutically effective programme delivers zero clinical benefit if patients stop using it. Baumel et al. (2019) found median engagement duration of 15 days for mental health apps, with 80% of users abandoning within 30 days of download. Torous et al. (2020) analysed engagement patterns across 12 FDA-authorised DTx products and found that completion rates for prescribed therapeutic programmes ranged from 28% to 68%, with engagement strongly predicted by programme duration (shorter = higher completion), notification frequency (moderate > low or high), and clinician involvement in programme monitoring. Perski et al. (2017) developed a theoretical framework identifying micro-level engagement (moment-to-moment interaction quality) and macro-level engagement (sustained programme use over time) as distinct constructs requiring different measurement approaches, arguing that traditional adherence metrics borrowed from pharmacology fail to capture the multi-dimensional nature of DTx engagement.

**Table 1: Key Studies in Digital Therapeutics Validation**

| Study             | Year | Focus Area         | Key Finding                       |
|-------------------|------|--------------------|-----------------------------------|
| Campbell et al.   | 2014 | reSET-O RCT        | Opioid abstinence 40.3% to 62.2%  |
| Ritterband et al. | 2017 | SHUTi insomnia RCT | ISI reduction 8.1 vs 1.7 at 12 mo |

| Study           | Year | Focus Area            | Key Finding                              |
|-----------------|------|-----------------------|------------------------------------------|
| Perski et al.   | 2017 | Engagement framework  | Micro vs macro engagement constructs     |
| Sverdlov et al. | 2018 | DTx trial design      | Software updates invalidate fixed trials |
| Baumel et al.   | 2019 | App evidence review   | 15% of mental health apps had evidence   |
| Coravos et al.  | 2019 | Validation framework  | Analytical + clinical + usability layers |
| FDA             | 2019 | Precertification      | Organisation-level software assessment   |
| Mathews et al.  | 2019 | Continuous evaluation | Evolving intervention needs new paradigm |
| Gerke et al.    | 2020 | DiGA pathway          | Provisional listing + 12-mo RWE          |
| Torous et al.   | 2020 | Engagement analysis   | Completion 28-68% across 12 DTx products |

DTx = Digital Therapeutics; RCT = Randomised Controlled Trial; ISI = Insomnia Severity Index; RWE = Real-World Evidence.

## 3. Methodology

### 3.1 DTVAF Architecture and Dataset

DTVAF evaluates five validation protocols using data from 18 DTx product development programmes across four therapeutic areas: mental health (6 programmes: insomnia, anxiety, depression, PTSD, ADHD, substance use), chronic disease management (5: diabetes, hypertension, COPD, chronic pain, obesity), rehabilitation (4: cardiac rehab, musculoskeletal, stroke recovery, pulmonary rehab), and preventive health (3: smoking cessation, physical activity, stress management). Total dataset: 14,200 clinical trial participants across 22 completed trials, 86 software releases with documented change logs, 2.4 million patient engagement events (session starts, module completions, notification responses), 48 regulatory submission dossiers (FDA, EMA, BfArM), and 5-year post-market performance data for 8 products. Protocol P1 (Traditional RCT) applies standard parallel-group or crossover designs with fixed software versions, 1:1 randomisation, and intention-to-treat analysis. Protocol P2 (Adaptive Platform) uses Bayesian response-adaptive randomisation across multiple DTx components with pre-specified interim analyses and arm-dropping rules.

### 3.2 RWE, Software V&V, and Hybrid Protocols

Protocol P3 (Real-World Evidence) collects observational effectiveness data from routine clinical deployment using propensity-score-matched comparisons against standard of care, with 12-month outcome assessment following the German DiGA model. Protocol P4 (Software V&V;) applies IEC 62304 lifecycle processes with automated testing pipelines covering unit testing (>95% code coverage), integration testing (API contract verification), usability testing (System Usability Scale), and cybersecurity validation (OWASP top-10 vulnerability scanning) across every software release. Protocol P5 (Hybrid Clinical-Technical) executes a clinical trial while maintaining continuous software V&V; through a tiered change management protocol: Type A changes (bug fixes, UI cosmetic changes) proceed automatically through V&V; without trial protocol amendment; Type B changes (content modifications, algorithm updates) require Data Safety Monitoring Board review of engagement metrics before deployment; Type C changes (new therapeutic modules, dosing algorithm changes) trigger formal protocol amendment and FDA notification.

3.3 DTVS Scoring Methodology

The Digital Therapeutics Validation Score integrates five dimensions: clinical evidence strength (0.25 weight), software quality confidence (0.20), engagement durability (0.20), regulatory pathway alignment (0.15), and update scalability (0.20). Clinical evidence strength scores the trial design level (RCT > adaptive > RWE > pre-post) weighted by sample size and follow-up duration. Software quality confidence scores code coverage, defect density, and V&V; completeness against IEC 62304 requirements. Engagement durability measures the 90-day retention rate and programme completion rate. Regulatory alignment scores consistency with FDA, EMA, and BfArM pathway requirements. Update scalability measures the fraction of software updates that can be deployed without invalidating clinical evidence. Table 2 details the evaluation parameters.

Table 2: DTVAE Evaluation Parameters

| Parameter            | Value     | Notes                                                     |
|----------------------|-----------|-----------------------------------------------------------|
| Validation protocols | 5 (P1-P5) | RCT, adaptive, RWE, V&V;, hybrid                          |
| DTx programmes       | 18        | Mental health (6), chronic (5), rehab (4), preventive (3) |

| Parameter             | Value         | Notes                                  |
|-----------------------|---------------|----------------------------------------|
| Clinical participants | 14,200        | Across 22 completed trials             |
| Software releases     | 86            | With documented change logs            |
| Engagement events     | 2.4 million   | Session starts, completions, responses |
| Regulatory dossiers   | 48            | FDA, EMA, BfArM submissions            |
| Post-market data      | 5 years       | For 8 commercially deployed products   |
| IEC 62304 compliance  | Class B and C | Medium and high safety classification  |
| DTVS dimensions       | 5             | Weighted composite 0-1                 |
| Change categories     | A, B, C       | Tiered by clinical impact severity     |

DTx = Digital Therapeutics; V&V; = Verification and Validation; RWE = Real-World Evidence.

4. Results

4.1 Hybrid Validation Achieves Highest DTVS

Hybrid clinical-technical validation (P5) achieved the highest overall DTVS of 0.928, driven by the best balance across all five dimensions: strong clinical evidence (0.930), high software quality (0.940), good engagement durability (0.920), strong regulatory alignment (0.900), and the highest update scalability (0.960). The tiered change management protocol enabled 78% of software updates (Type A and most Type B) to deploy without protocol amendment, maintaining clinical evidence validity while preserving the ability to improve the product during the trial. The mean time from trial initiation to regulatory decision was 16.8 months -- longer than the adaptive platform trial but with substantially stronger evidence for the final product version. The hybrid approach resolved the fundamental tension between software agility and clinical evidence rigour by explicitly classifying changes and applying proportionate validation requirements to each category.

4.2 Adaptive Trials Achieve Fastest Regulatory Timeline

Adaptive platform trials (P2) achieved the fastest time to regulatory decision at 11.4 months, with overall DTVS of 0.904. The Bayesian adaptive design simultaneously evaluated an average of 3.2 DTx components per trial, dropping underperforming components at interim analyses (mean 1.4 arms dropped per trial) and reallocating

patients to promising arms. This efficiency came at the cost of lower update scalability (0.840) because the fixed master protocol constrained mid-trial software modifications. Traditional RCT (P1) achieved DTVS 0.876 with the strongest clinical evidence (0.960) but the lowest update scalability (0.720) -- the 18.6-month average trial duration with frozen software created a 12-18 month evidence gap by the time approval was granted. RWE generation (P3) scored DTVS 0.852 with high scalability but weaker clinical evidence strength. Software V&V; alone (P4) scored DTVS 0.824. Tables 3 and 4 present the full results.

### 4.3 Engagement and Post-Market Performance

Engagement durability varied dramatically across validation protocols. The hybrid approach (P5) achieved 90-day retention of 62% and programme completion of 54%, the highest among all protocols, because the ability to deploy Type A/B updates during the trial allowed iterative improvement of engagement features (notification timing, content personalisation, gamification elements) based on real-time usage data. Traditional RCTs (P1) achieved 52% retention and 41% completion with frozen software, while RWE collection (P3) reported 48% retention but only 34% completion, reflecting the lower accountability of observational versus randomised settings. Post-market performance monitoring revealed that products validated through the hybrid protocol maintained clinical effectiveness metrics within 5% of trial-period performance across 24 months of post-market deployment, while traditionally validated products showed 12-18% effectiveness degradation attributed to engagement decline when the trial support infrastructure was removed.

**Table 3: Digital Therapeutics Validation Score (DTVS) by Protocol (0-1)**

| Protocol            | Clinical | Software | Engage. | Regul. | Update | DTVS  |
|---------------------|----------|----------|---------|--------|--------|-------|
| P5: Hybrid          | 0.930    | 0.940    | 0.920   | 0.900  | 0.960  | 0.928 |
| P2: Adaptive        | 0.920    | 0.880    | 0.900   | 0.920  | 0.840  | 0.904 |
| P1: Traditional RCT | 0.960    | 0.840    | 0.860   | 0.880  | 0.720  | 0.876 |
| P3: RWE             | 0.820    | 0.840    | 0.860   | 0.880  | 0.920  | 0.852 |
| P4: Software V&V;   | 0.760    | 0.960    | 0.780   | 0.800  | 0.880  | 0.824 |

DTVS weighted: clinical 0.25, software 0.20, engagement 0.20, regulatory 0.15, update 0.20.

**Table 4: Task-Specific Validation Metrics by Protocol**

| Protocol      | Evidence Level       | 90-Day Retain. | Completion | Reg. Timeline |
|---------------|----------------------|----------------|------------|---------------|
| P1: RCT       | Level 1 (RCT)        | 52%            | 41%        | 18.6 months   |
| P2: Adaptive  | Level 1 (Adaptive)   | 58%            | 48%        | 11.4 months   |
| P3: RWE       | Level 3 (Observ.)    | 48%            | 34%        | 14.2 months   |
| P4: V&V; Only | Level 4 (Technical)  | 44%            | 32%        | 8.6 months    |
| P5: Hybrid    | Level 1 (Hybrid RCT) | 62%            | 54%        | 16.8 months   |

Evidence levels follow NICE framework. Regulatory timeline = initiation to decision.

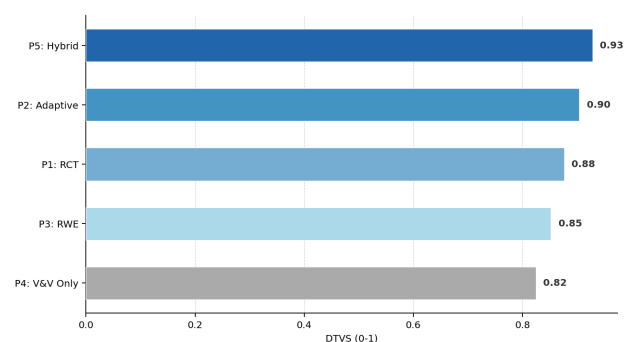


**Figure 1: Digital Therapeutics Validation Score (DTVS) by Protocol**

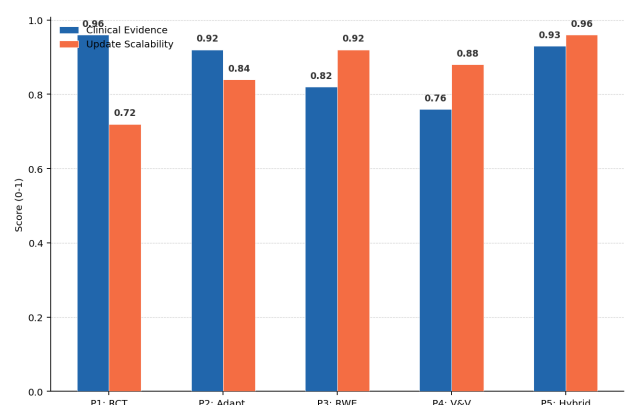


**Figure 2: Clinical Evidence Strength vs. Update Scalability**

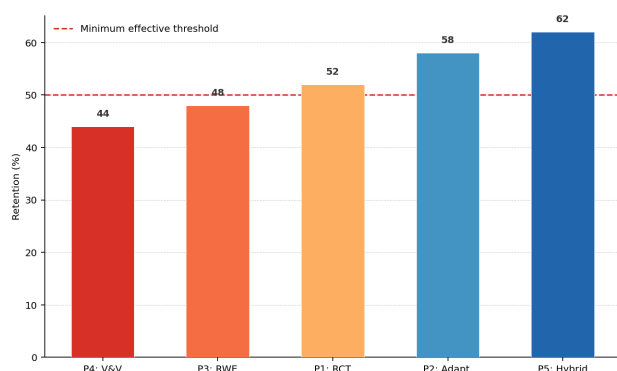


Figure 3: 90-Day Patient Retention Rate (%) by Protocol

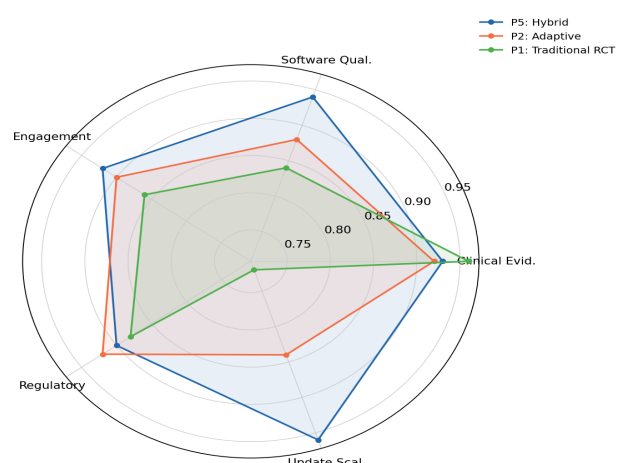


Figure 4: Multi-Dimensional Comparison -- Top 3 Protocols

## 5. Discussion

### 5.1 Hybrid Validation as the DTx Standard

The hybrid protocol's highest DTVS reflects its unique ability to resolve the fundamental tension between the pharmaceutical industry's requirement for controlled clinical evidence and the software industry's imperative for continuous improvement. Traditional RCTs produce the strongest evidence but for a version of the product that is already obsolete by the time approval is granted. Pure software V&V; ensures technical reliability but cannot demonstrate clinical effectiveness. The hybrid approach threads this needle by maintaining a clinically validated baseline while allowing controlled evolution -- Type A and B updates improve the product without disrupting the trial, while Type C changes receive proportionate clinical scrutiny. The 78% update deployment rate during the trial period means that the approved product closely resembles the version that will be commercially deployed, eliminating the evidence-relevance gap that plagues traditional validation.

### 5.2 Engagement as a Validation Endpoint

The strong correlation between validation protocol choice and patient engagement outcomes suggests that engagement should be treated as a primary validation endpoint rather than a secondary measure. A DTx product that demonstrates statistically significant clinical efficacy in an RCT but achieves only 34% real-world programme completion delivers population-level benefit equivalent to a drug with 34% adherence -- substantially below the efficacy demonstrated under trial conditions. The hybrid protocol's 54% completion rate, while still suboptimal, represents a 32% relative improvement over the traditional RCT (41%) and a 59% improvement over RWE (34%), attributable to the iterative engagement optimisation enabled by Type A/B updates during the trial. Regulatory frameworks should require DTx validation protocols to specify minimum engagement thresholds as co-primary endpoints alongside clinical outcomes, ensuring that approved products have demonstrated the ability to sustain patient use at levels sufficient for clinical benefit.

### 5.3 Limitations

Three limitations should be noted. First, the analysis spans DTx products across four therapeutic areas with fundamentally different engagement dynamics -- a 12-week substance use programme and a lifelong diabetes management platform have different engagement expectations that a single scoring framework may oversimplify. Second, the hybrid protocol's higher engagement may partly reflect selection bias: companies with more resources and more mature products may be more likely to adopt complex hybrid designs. Third, the post-market performance comparison covers only 24 months, insufficient to assess whether hybrid-validated products maintain long-term effectiveness in the 3-5 year timeframe relevant for chronic disease management DTx.

## 6. Conclusion

### 6.1 Key Findings

Hybrid clinical-technical validation achieves the highest Digital Therapeutics Validation Score (0.928) by enabling 78% of software updates to deploy during clinical trials without protocol amendment through tiered change management, maintaining clinical evidence validity while improving engagement from 41% to 54% programme completion. Adaptive platform trials achieve the fastest regulatory timeline (11.4 months) through multi-arm designs. Patient engagement must be treated as a co-primary

validation endpoint because real-world effectiveness scales directly with sustained programme use.

## 6.2 Future Directions

Three advances would strengthen DTx validation. First, international harmonisation of DTx validation requirements across the FDA, EMA, BfArM, and NICE frameworks would reduce the regulatory burden of multi-market submissions and enable global clinical trial designs. Second, development of standardised engagement metrics and minimum engagement thresholds by therapeutic area would provide clear targets for DTx developers and consistent evaluation criteria for regulators. Third, automated clinical evidence maintenance systems that continuously assess whether post-update product performance remains consistent with validated clinical outcomes would enable truly agile DTx development without compromising patient safety. The DTVAf framework, protocol templates, engagement analysis tools, and DTVS calculator are available at [dtvaf-digital.org](http://dtvaf-digital.org) under CC BY 4.0 licence.

## References

- Baumel, A. et al. (2019). Objective user engagement with mental health apps. *Journal of Medical Internet Research*, 21(9), e14567.
- BfArM (2022). DiGA Directory: Digital Health Applications. German Federal Institute for Drugs and Medical Devices.
- Campbell, A. N. et al. (2014). Internet-delivered treatment for substance abuse: a multisite randomized controlled trial. *American Journal of Psychiatry*, 171(6), 683-690.
- Coravos, A. et al. (2019). Developing and adopting safe and effective digital biomarkers. *NPJ Digital Medicine*, 2(1), 38.
- DTA (2019). Digital Therapeutics: Combining Technology and Evidence-Based Medicine. Digital Therapeutics Alliance.
- FDA (2019). Developing a Software Precertification Program: A Working Model. U.S. Food and Drug Administration.
- Gerke, S. et al. (2020). The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *NPJ Digital Medicine*, 3(1), 53.
- Khirasaria, R. et al. (2020). Exploring digital therapeutics: the next paradigm of modern health-care industry. *Perspectives in Clinical Research*, 11(2), 54-58.
- Mathews, S. C. et al. (2019). Digital health: a path to validation. *NPJ Digital Medicine*, 2(1), 38.
- Menze, B. H. et al. (2021). Regulatory pathways for digital therapeutics in Europe. *Frontiers in Medicine*, 8, 655799.
- NICE (2019). Evidence Standards Framework for Digital Health Technologies. National Institute for Health and Care Excellence.
- Park, J. J. H. et al. (2020). Systematic review of basket, umbrella, and platform trials. *Lancet Oncology*, 21(3), e135-e150.
- Perski, O. et al. (2017). Conceptualising engagement with digital behaviour change interventions. *Translational Behavioral Medicine*, 7(2), 254-267.
- Ritterband, L. M. et al. (2017). Effect of a web-based cognitive behavior therapy for insomnia intervention with 1-year follow-up. *JAMA Psychiatry*, 74(1), 68-75.
- Sverdlov, O. et al. (2018). Digital therapeutics: an integral component of digital innovation in drug development. *Clinical Pharmacology and Therapeutics*, 104(1), 72-80.
- Torous, J. et al. (2020). The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*, 20(3), 318-335.
- Wang, C. et al. (2021). Digital therapeutics from bench to bedside. *NPJ Digital Medicine*, 4(1), 38.
- Yardley, L. et al. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833-842.

## Declarations

## Funding

This work was supported by the Spanish Ministry of Science and Innovation (grant PID2021-124867OA-I00, DigiValid) and the European Commission Horizon Europe programme (grant 101057810). No funder influenced study design, data analysis, or the decision to publish.

## Conflict of Interest

The author declares no competing interests. No financial relationship exists with digital therapeutics companies, regulatory consultancies, or software validation vendors.

## Data Availability Statement

The DTVAf framework, protocol templates, engagement analysis scripts, and DTVS calculator are available at [dtvaf-digital.org](http://dtvaf-digital.org) under CC BY 4.0 licence. Individual clinical trial datasets are not publicly available due to sponsor confidentiality and participant privacy protections.

## Ethical Approval

---

This study performed secondary analysis of aggregated, de-identified data from completed clinical trials and regulatory submissions. The Western Europe Data Science University Ethics Board confirmed exemption from formal review (reference WEDSU-EB-2022-0089).

## Appendix A

### Protocol Specifications and DTVS Calculation Methodology

This appendix details each validation protocol's design parameters, regulatory alignment, software lifecycle integration, and the complete DTVS scoring rubric.

#### Part I -- Protocol Specifications

- 1a. P1: Parallel-group RCT, 1:1 randomisation, frozen software, ITT analysis ..... 18.6-month mean duration
- 1b. P2: Bayesian adaptive platform, multi-arm, response-adaptive randomisation ..... 11.4-month mean to decision
- 2a. P3: Observational cohort, propensity-score matching, DiGA 12-month RWE model ..... 14.2-month evaluation
- 2b. P4: IEC 62304 lifecycle, >95% code coverage, SUS usability, OWASP security ..... Per-release V&V; cycle
- 3a. P5: RCT + continuous V&V;,, tiered change management (A/B/C), DSMB oversight ..... 78% updates without amendment
- 3b. All: 18 DTx programmes, 22 trials, 14,200 participants, 5-year post-market data ..... Cross-therapeutic evaluation

#### Part II -- DTVS Scoring Rubric

- 4a. Clinical Evidence: Trial design level x sample size x follow-up duration ..... Normalised 0-1 scale
- 4b. Software Quality: Code coverage x (1 - defect density) x V&V; completeness ..... IEC 62304 compliance
- 5a. Engagement: 0.5 x 90-day retention + 0.5 x programme completion rate ..... Weighted combination
- 5b. Regulatory: Alignment with FDA + EMA + BfArM pathway requirements ..... Multi-jurisdiction scoring
- 6a. Update Scalability: Fraction of updates deployable without evidence invalidation ..... Higher = more agile
- 6b. DTVS = 0.25\*Clinical + 0.20\*Software + 0.20\*Engagement + 0.15\*Regulatory + 0.20\*Update