

Viewpoint

Key Considerations for Designing Clinical Studies to Evaluate Digital Health Solutions

Elaina Bolinger^{1*}, PhD; Benoit Tyl^{2*}, MD

¹Integrated Evidence Generation & Business Innovation, Bayer AG, Berlin, Germany

²Integrated Evidence Generation & Business Innovation, Bayer HealthCare SAS, La Garenne Colombes, France

* all authors contributed equally

Corresponding Author:

Benoit Tyl, MD

Integrated Evidence Generation & Business Innovation

Bayer HealthCare SAS

10 Place de Belgique

La Garenne Colombes, F-92254

France

Phone: 33 6 80 29 07 79

Email: benoit.tyl@bayer.com

Abstract

Evidence of clinical impact is critical to unlock the potential of digital health solutions (DHSs), yet many solutions are failing to deliver positive clinical results. We argue in this viewpoint that this failure is linked to current approaches to DHS evaluation design, which neglect numerous key characteristics (KCs) requiring specific scientific and design considerations. We first delineate the KCs of DHSs: (1) they are implemented at health care system and patient levels; (2) they are “complex” interventions; (3) they can drive multiple clinical outcomes indirectly through a multitude of smaller clinical benefits; (4) their mechanism of action can vary between individuals and change over time based on patient needs; and (5) they develop through short, iterative cycles—optimally within a real-world use context. Following our objective to drive better alignment between clinical evaluation design and the unique traits of DHSs, we then provide methodological suggestions that better address these KCs, including tips on mechanism-of-action mapping, alternative randomization methods, control-arm adaptations, and novel end-point selection, as well as innovative methods utilizing real-world data and platform research.

(*J Med Internet Res* 2024;26:e54518) doi: [10.2196/54518](https://doi.org/10.2196/54518)

KEYWORDS

evidence generation; clinical robustness; clinical trials; digital health; solutions; digital health solutions; DHS; health care system; patients; patient; individuals; individual; healthcare system; control arm adaptations; randomization methods; real-world data; platform research

Introduction

Digital health solutions (DHSs) can improve health care access, patient equity, operational efficiency, and cost-effectiveness for health care organizations while delivering clinical outcomes for patients. Despite the rapid proliferation of DHSs, most lack convincing evidence supporting their clinical impact [1], preventing their uptake within the health care system and blocking subsequent development cycles required to realize their full clinical potential. Closing this evidence gap and producing strong evidence in a timely and cost-effective manner is critical for the establishment of trust in DHSs and imperative for their adoption and implementation—prerequisites for

unlocking digital health’s potential to truly transform health care [2].

Randomized controlled trials (RCTs) using placebo control groups; individual randomization; and strict, arguably artificial clinical settings are the gold standard for evaluating drug efficacy and safety [3]. But contrary to pharmacological treatments, in which efficacy is driven by the modulation of biology at the molecular level, DHSs belong to a specific interventional class known as “complex interventions” [4]. By definition, the impact of this type of intervention is not solely biological but also encompasses psychological, behavioral, and systems-level effects, indicating context-dependency [5]. These attributes may, for example, help explain why effect sizes for pharmaceutical products are generally higher in traditional RCTs

compared to research conducted under conditions reflecting real-world conditions [6,7], whereas DHS effect sizes tend to be higher when using designs that mimic real-world conditions [8]. The unique attributes of DHSs and the “complex” class of interventions they deliver may therefore necessitate a more tailored approach to their evaluation to measure their true impact.

In this viewpoint, we discuss our perspective on what must be considered when designing clinical evaluations of DHSs. We specifically aimed to (1) delineate a series of key characteristics (KCs) of DHSs that make them different from pharmacological agents and should be considered while designing their evaluation, and (2) offer methodological solutions to adapt research to the requirements of DHSs, addressing randomization; control-arm design; end-point selection; and nontraditional, innovative adaptations.

DHS Key Characteristics

KC1: DHSs Are Often Implemented at Health Care System and Patient Levels

While patients may generate most of the data, DHSs frequently involve other individuals such as health care professionals (HCPs) and caregivers, who must integrate the solutions into their workflows for the DHS to reach full effectiveness. The benefits of using DHSs frequently expand beyond patients to caregivers, HCPs, and the whole health care organization. For instance, implementing a remote patient monitoring DHS in a single health care center can have permeating systemic effects, with the product improving outcomes of the patients monitored as well as those of the other patients, either by decreasing readmission and then shifting clinical staff resources to preventive care or by improving health care efficiency (streamlined workflow, optimized patient engagement strategies, etc). Within an evaluation setting, HCPs in the control group may adapt their behavior via learning from those included in the intervention group and change their clinical strategies (ie, leading to contamination). Therefore, it is important to control not just at the patient level but also at the health care site level [9].

KC2: DHSs Are “Complex” Interventions

Owing to their multiuser designs and dependencies on real-world health care systems, the mechanism of action (MoA) of DHSs depends on the historical, situational, environmental, and psychological factors in which treatments are delivered. Contextual dependencies include but are not limited to: HCP characteristics; patient-HCP relationships; perceived intervention credibility; delivery modality; psychological state of the individual; and societal, economic, and cultural factors. Even seemingly small nuances, such as a patient having previously been asked about their health, can have a significant impact on their health-related behavior and consequently also on health outcomes [10-12]. Many of these contextual factors are unintentionally, artificially modified in pharmaceutical RCTs to maximize internal validity. For instance, the Hawthorn effect, which is an increase in engagement when people are observed [13], and other research participation effects [14] are controlled

between groups by using placebos. Yet this engagement is in fact a designed benefit of several DHSs; therefore, adding measurement points that would not occur in real-world situations to a DHS control group can render the control noninert. For example, measuring and regularly reporting blood pressure improves engagement and hypertension control [15], and similarly, reporting symptoms improves cancer self-management practices [16]. The artificial contexts and measurement conditions traditionally used in pharmaceutical RCTs can therefore limit the ecological validity of how the DHS is used. Consequently, conclusions drawn from such designs may lack external validity.

KC3: DHSs May Drive Multiple Clinical Outcomes Indirectly Through a Multitude of Smaller Clinical Benefits

DHSs are frequently “holistic” interventions that provide a broad spectrum of solutions to induce diverse changes, mainly behavioral, which then work as levers to deliver clinically meaningful outcomes. For instance, a remote patient monitoring platform for patients with diabetes that reduces time to treatment adjustment and thereby improves pharmaceutical adherence should consequently improve blood glucose regulation. Similarly, any improvement of sleep behavior by a sleep coaching DHS may also improve cardiovascular risk through better exercise, decreased weight, better eating habits, decreased cholesterol, and improved glucose regulation. While the magnitude of effect of the DHS on each of these end points may be limited, the cumulative effects may lead to a clinically significant benefit. Choosing a single primary clinical end point during evaluation therefore may not capture the complexity of the DHS’s intended purpose nor its true performance.

KC4: The MoA Driving Clinical Outcomes in DHSs Can Vary Between Individuals and Change Over Time Based on Patient Needs

Many DHS interventions include several separate, situationally activated components or features (HCP-facing interface, symptom monitoring, activity tracking, etc) that can bring unique benefits. For example, one patient might use feature A, which indirectly improves cholesterol levels; another might prefer feature B and experience an improvement in blood pressure, with smaller changes observed in cholesterol; and some patients might benefit from both. The effect size of the intervention for a single end point may therefore depend on how the product is used. Moreover, the second patient may begin to prefer feature A, but only later during use. The effect size of the intervention for a single end point can therefore be weakened when the DHS as a whole rather than specific feature use is considered as the intervention. What is more, many DHSs employ adaptive algorithms that are specifically designed to adjust the intervention over time. Depending on the DHS being considered, researchers might therefore anticipate that due to user-dependent variability in the MoA and emerging synergistic influences (ie, better product engagement following use-based personalization), the clinical impact may require a longer treatment period to be detected.

KC5: DHSs Develop Through Short, Iterative Cycles—Often Within a Real-World Use Context

Contrary to active pharmacological ingredients that cannot change, DHSs continuously evolve in terms of their functionality with each subsequent release. Because DHSs must meet their users' needs in the real world, an environment where many factors remain unknown [17], developers often perform accelerated, iterative tests to both discover product performance factors and adapt them before the next release. Due to the number of factors that must be discovered, how they can influence each other, and how influential certain factors can be (ie, a seemingly small adjustment of written content to an audio format might completely change engagement and therefore clinical outcomes), the solution often must stay moderately malleable, as product design hypotheses are validated over a continuum of nonclinical tests. Clinical researchers often opt to evaluate a “frozen” version of the product to optimize homogeneity in the intervention arm, but these approaches may limit external validity, as the marketed solution may drastically evolve over time and the versions tested by the RCT may be outdated at the time of study completion.

Recalibrating Research Design to Meet the Scientific Canvas of DHSs

Considering the KCs described above, alternative assessment strategies are needed to address the unique challenges of DHS evaluation while maintaining internal validity (unbiased study design), external validity (applicability to different contexts), and ecological validity (generalizability to real-world settings). The solutions discussed below are summarized in [Figure 1](#).

The first step to design a DHS evaluation is to take the time to delineate the MoA of the solution and the context. This includes a map of behavioral changes required from patients and all other individuals who are affected by the product (KC1), as well as the contextual environment of use (KC2). Given the limitations of an active control arm described above (KC2), the use of usual care—usually named treatment as usual (TAU)—may serve as an appropriate control. TAU corresponds to the actual routine care and may differ from the state-of-the-art, guideline-adherent clinical care. Therefore, a prerequisite for this approach would be a detailed understanding of the procedural standard of care, which includes, but is not limited to, *local* (ie, site-specific) procedures, treatment recommendations, information delivered (eg, health literacy counseling), and frequency of contacts with HCPs. This task is nontrivial, as evidenced by standard-of-care descriptions being generally underreported [18], and can pose exceptional challenges due to the number of factors in a real-world environment. Combining patient journeys, DHS user journeys, and HCP patient management protocols is a helpful first step to identify how the product induces change within its context of use. Process evaluation techniques [19] can also be employed to extract MoAs for complex interventions during preliminary implementation research [5]. With the MoA delineated in detail, one can begin designing an appropriate evaluation employing a fit-for-purpose scientific approach.

As many have suggested before [9,20], cluster randomization by health care site, cities, or region is a methodologically sound approach to control for contamination between arms while taking into account DHS benefits, which depend on factors extending beyond the treated patient (KC1). This design, however, can necessitate larger sample sizes, and site matching can be challenging. Pragmatic designs [9], which anchor RCTs in real-world settings and minimize research-related contact outside of TAU, can help maximize the ecological validity of both the experimental (+TAU) and TAU arms (KC2). Keeping the trial decentralized can minimize research-related contact while increasing trial accessibility and make the research population more representative of the target population (KC2). While DHSs, which digitally collect data by nature, are particularly suitable for such designs, limitations secondary to the reduced investigators oversight, including data quality and outcome assessment, persist.

Even with these tools in mind, designing a control arm with high internal validity remains challenging when evaluating DHSs. A reference DHS is an appropriate control but rarely exists. While some have argued that a “sham” app (ie, as defined by the Food and Drug Administration, an app without the “active” features anticipated to drive the MoA [21]) offers optimal internal validity, in that it is a form of “placebo” technology, this design can pose critical threats to external and ecological validity, particularly when the MoA itself is not fully understood. Providing even minimalistic digital support that is nevertheless seemingly endorsed by HCPs (as would be the case in a clinical trial) might still produce a similar, albeit muted, effect as the DHS. For example, a sham app delivering some health literature is not inert, as it can improve health literacy and consequently health-related behaviors. The use of a sham app as a control is therefore challenging unless it is possible to exclude all active ingredients of the MoA, including engagement and interaction with the real world, from the sham app. On the other hand, the use of TAU as a control can potentially inflate effect size with some interventions (KC2). The traditional “waitlist” design used in medical device evaluation may also be appropriate but can inflate effect sizes by decreasing behavioral activation in the control arm (in essence, inducing the opposite of the Hawthorne effect) [22,23]. Researchers considering having a waitlist TAU design as the comparator might use an appropriate *run-in period* before randomization, during which both study arms receive TAU to minimize the influence of treatment expectancy effects [24] within the comparator arm (KC2). Finally, it is important to consider how the nature of the intervention and the control arm may impact the risk of attrition in the study. Beyond minimizing attrition in general, it is crucial to ensure that the attrition rates between the active and control arms are balanced. Significant differences in dropout rates between both arms may lead to variations in the clinical characteristics of the remaining patients, potentially biasing the overall analysis.

DHS researchers may also explore alternate evaluation designs. The use of synthetic control arms built using real-world data sources such as electronic health records, claims databases, and disease registries may address certain limitations inherent to the conventional control group. However, the creation of these

synthetic arms needs to be considered carefully, as they pose their own set of challenges and are not devoid of their own limitations. In case of products already on the market, epidemiological methods could be used to examine causal inference in product registry data—again minimizing research participation effects. These approaches may in some cases both be more cost-effective and provide higher quality data on product performance.

Designing a trial that accounts for the holistic impact of DHSs is also critical for demonstrating performance and efficacy. Contrary to pharmacological agents that usually improve markedly a single biomarker (hemoglobin A_{1c}, C-reactive protein, etc), DHSs frequently improve modestly several biomarkers (physical activity, lipids, weight, etc; KC3). Here, rather than examining the effect on a single end point, researchers might employ a risk score that better reflects the holistic nature of the intervention. Win ratios are also an elegant way to analyze multiple, equally valuable end points [25]. Adaptive designs, such as those incorporating sample size re-estimation, hypothesis-adaptive, or response-adaptive designs [26], could also be both powerful and appropriate given the novel nature of many DHS interventions and the different use patterns and emergent interventions that can arise (KC4). Given the dynamic and personalized nature of DHSs, the evaluation

design should also carefully consider anticipated effect size at each time point during planning.

The above can greatly improve the evaluation of DHS clinical performance but do not address the practical issue of evaluating rapidly and continuously evolving DHSs (KC5). While there is no definitive way to address this, a few options exist beyond freezing the product. During evaluation planning, planned future modifications of the product can be scrutinized to determine whether they constitute major changes to the MoA or intended use and would therefore require protocol modification during study execution. By combining strict product design control practices with a Trials of Intervention Principles approach [27], researchers can aggregate groups of users who may have used product versions with minor design variations not anticipated to change the MoA, allowing the product to continue to organically develop (KC5). Platform trials that incorporate key product design and clinical value hypotheses in master protocols can be leveraged to have the successive versions of the DHS tested and the results of the substudies combined to maximize sample sizes, rather than performing several insufficiently powered successive trials, while keeping analyses anchored in clearly documented starting hypotheses and protocols (Figure 2) [20,28]. Likewise, factorial research designs can be used to investigate the efficacy of new features as they are added to the product [29,30].

Figure 1. Overview of key characteristics (KCs) of the digital health solutions (DHSs) and potential methodological solutions to improve their clinical assessment. EHR: electronic health record; MoA: mechanism of action. For a higher-resolution version of this figure, see [Multimedia Appendix 1](#).

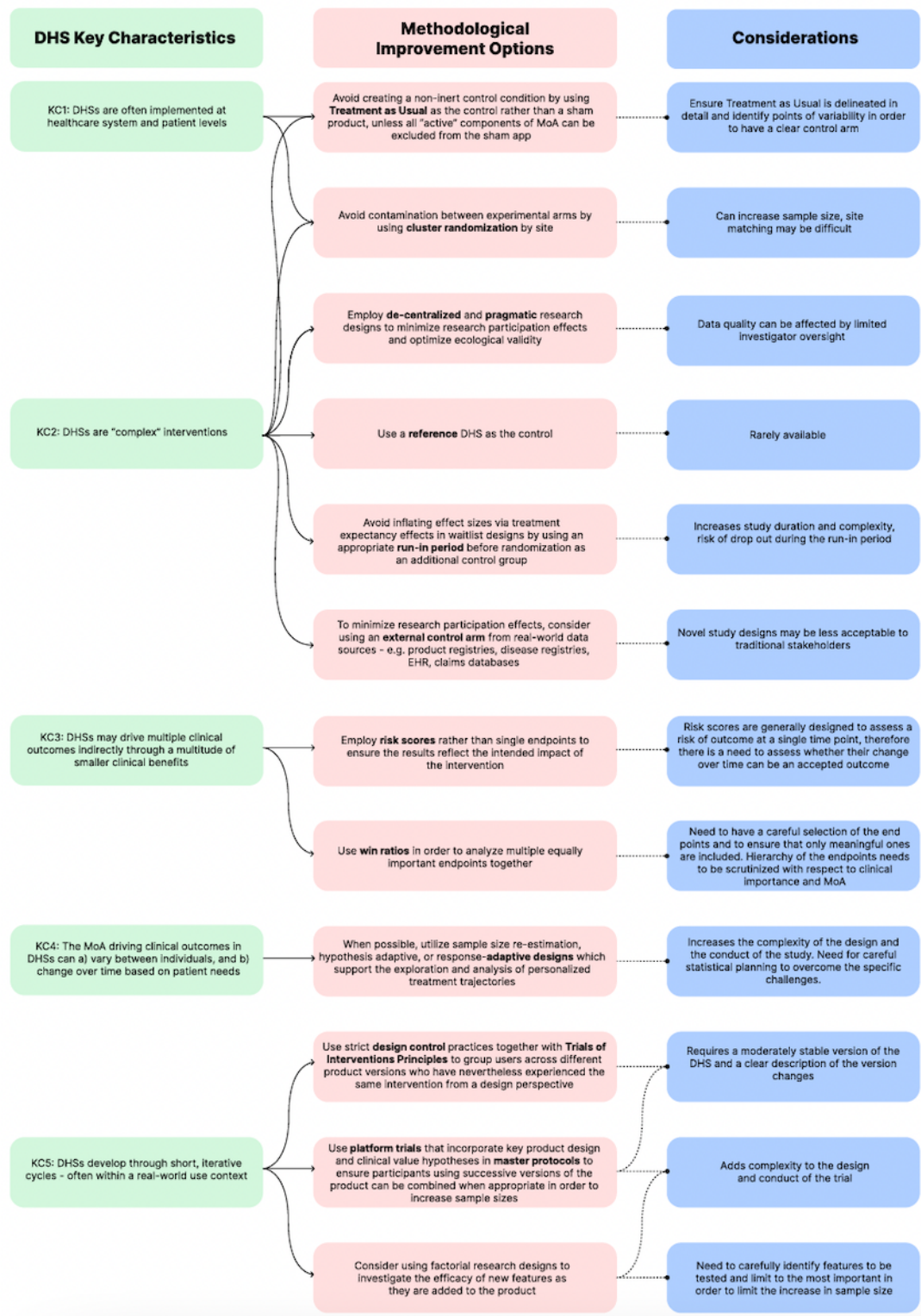
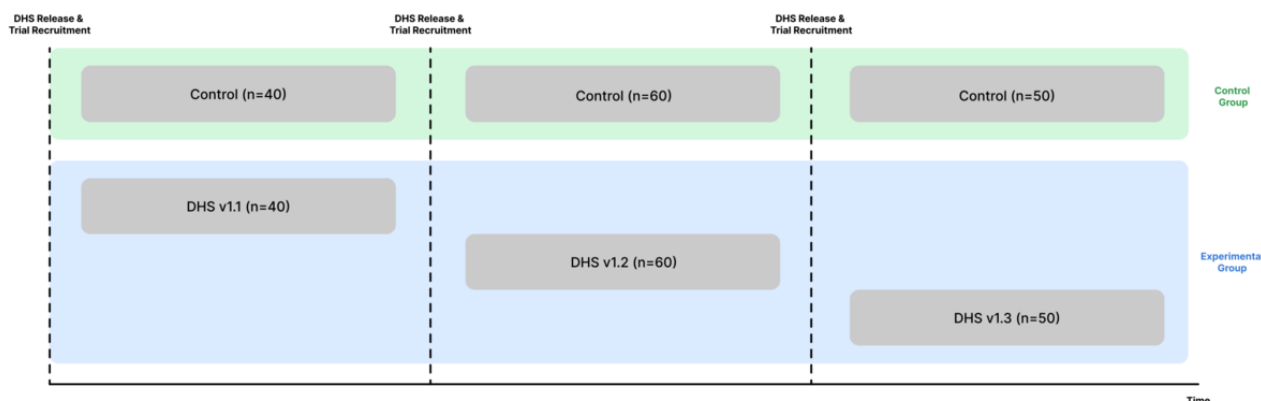


Figure 2. Example of design of a platform trial including 300 patients (150 per arm) enabling the evaluation of multiple versions of a DHS within one study. The figure illustrates the recruitment flow of the patients in the trial. Multiple versions of a DHS can be included in a single study when guidelines for identifying significant changes to study procedures and DHS solutions are delineated in advance. A matched control arm with similar recruitment procedures ensures comparability in the control group. DHS: digital health solution.



Conclusion

DHSs need to be rigorously evaluated to fulfill the needs of the various stakeholders, including HCPs, patients, regulators, and payers. In this viewpoint, we proposed that their unique characteristics and MoA create a critical need for tailored,

innovative approaches that move beyond traditional pharmaceutical RCTs in order to create fit-for-purpose, robust, and scientifically optimized evaluations. While no one-size-fits-all design exists, both researcher and stakeholders should embrace nontraditional methodologies that match the key characteristics of this emerging type of health care intervention.

Acknowledgments

We thank Jason Lott for his feedback and comments.

Authors' Contributions

Both authors equally contributed to the manuscript.

Conflicts of Interest

Both authors are full-time employees of Bayer.

Multimedia Appendix 1

Overview of key characteristics (KCs) of the digital health solutions (DHSs) and potential methodological solutions to improve their clinical assessment. EHR: electronic health record; MoA: mechanism of action.

[[PDF File \(Adobe PDF File\), 1793 KB-Multimedia Appendix 1](#)]

References

- Day S, Shah V, Kaganoff S, Powelson S, Mathews SC. Assessing the clinical robustness of digital health startups: cross-sectional observational analysis. *J Med Internet Res*. Jun 20, 2022;24(6):e37677. [[FREE Full text](#)] [doi: [10.2196/37677](https://doi.org/10.2196/37677)] [Medline: [35723914](https://pubmed.ncbi.nlm.nih.gov/35723914/)]
- Egermark M, Blasiak A, Remus A, Sapanel Y, Ho D. Overcoming pilotitis in digital medicine at the intersection of data, clinical evidence, and adoption. *Adv Intell Syst*. Sep 21, 2022;4(9):2200056. [doi: [10.1002/aisy.202200056](https://doi.org/10.1002/aisy.202200056)]
- Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. Feb 1986;89(2 Suppl):2S-3S. [doi: [10.1378/chest.89.2_supplement.2s](https://doi.org/10.1378/chest.89.2_supplement.2s)] [Medline: [3943408](https://pubmed.ncbi.nlm.nih.gov/3943408/)]
- Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med*. Nov 2016;51(5):843-851. [[FREE Full text](#)] [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
- Minary L, Trompette J, Kivits J, Cambon L, Tarquinio C, Alla F. Which design to evaluate complex interventions? toward a methodological framework through a systematic review. *BMC Med Res Methodol*. May 07, 2019;19(1):92. [[FREE Full text](#)] [doi: [10.1186/s12874-019-0736-6](https://doi.org/10.1186/s12874-019-0736-6)] [Medline: [31064323](https://pubmed.ncbi.nlm.nih.gov/31064323/)]
- Vonbank A, Drexel H, Agewall S, Lewis B, Dopheide J, Kjeldsen K, et al. Reasons for disparity in statin adherence rates between clinical trials and real-world observations: a review. *Eur Heart J Cardiovasc Pharmacother*. Oct 01, 2018;4(4):230-236. [[FREE Full text](#)] [doi: [10.1093/ehjcvp/pyy028](https://doi.org/10.1093/ehjcvp/pyy028)] [Medline: [30099530](https://pubmed.ncbi.nlm.nih.gov/30099530/)]

7. Kilcher G, Hummel N, Didden EM, Egger M, Reichenbach S, GetReal Work Package 4. Rheumatoid arthritis patients treated in trial and real world settings: comparison of randomized trials with registries. *Rheumatology (Oxford)*. Feb 01, 2018;57(2):354-369. [doi: [10.1093/rheumatology/kex394](https://doi.org/10.1093/rheumatology/kex394)] [Medline: [29149289](https://pubmed.ncbi.nlm.nih.gov/29149289/)]
8. Firth J, Torous J, Nicholas J, Carney R, Prata A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*. Oct 21, 2017;16(3):287-298. [FREE Full text] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
9. Guo C, Ashrafiyan H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions-a call for innovative evidence generation approaches. *NPJ Digit Med*. Aug 27, 2020;3(1):110. [FREE Full text] [doi: [10.1038/s41746-020-00314-2](https://doi.org/10.1038/s41746-020-00314-2)] [Medline: [32904379](https://pubmed.ncbi.nlm.nih.gov/32904379/)]
10. de Bruin M. Risk of bias in randomised controlled trials of health behaviour change interventions: evidence, practices and challenges. *Psychol Health*. Jan 09, 2015;30(1):1-7. [doi: [10.1080/08870446.2014.960653](https://doi.org/10.1080/08870446.2014.960653)] [Medline: [25223823](https://pubmed.ncbi.nlm.nih.gov/25223823/)]
11. Bishop FL, Fenge-Davies AL, Kirby S, Geraghty AW. Context effects and behaviour change techniques in randomised trials: a systematic review using the example of trials to increase adherence to physical activity in musculoskeletal pain. *Psychol Health*. Jan 09, 2015;30(1):104-121. [doi: [10.1080/08870446.2014.953529](https://doi.org/10.1080/08870446.2014.953529)] [Medline: [25109300](https://pubmed.ncbi.nlm.nih.gov/25109300/)]
12. Conner M, Godin G, Norman P, Sheeran P. Using the question-behavior effect to promote disease prevention behaviors: two randomized controlled trials. *Health Psychol*. May 2011;30(3):300-309. [doi: [10.1037/a0023036](https://doi.org/10.1037/a0023036)] [Medline: [21553974](https://pubmed.ncbi.nlm.nih.gov/21553974/)]
13. Berkhout C, Berbra O, Favre J, Collins C, Calafiore M, Peremans L, et al. Defining and evaluating the Hawthorne effect in primary care, a systematic review and meta-analysis. *Front Med (Lausanne)*. Nov 8, 2022;9:1033486. [FREE Full text] [doi: [10.3389/fmed.2022.1033486](https://doi.org/10.3389/fmed.2022.1033486)] [Medline: [36425097](https://pubmed.ncbi.nlm.nih.gov/36425097/)]
14. McCambridge J, Kypri K, Elbourne D. Research participation effects: a skeleton in the methodological cupboard. *J Clin Epidemiol*. Aug 2014;67(8):845-849. [FREE Full text] [doi: [10.1016/j.jclinepi.2014.03.002](https://doi.org/10.1016/j.jclinepi.2014.03.002)] [Medline: [24766858](https://pubmed.ncbi.nlm.nih.gov/24766858/)]
15. Milani RV, Lavie CJ, Bober RM, Milani AR, Ventura HO. Improving hypertension control and patient engagement using digital tools. *Am J Med*. Jan 2017;130(1):14-20. [doi: [10.1016/j.amjmed.2016.07.029](https://doi.org/10.1016/j.amjmed.2016.07.029)] [Medline: [27591179](https://pubmed.ncbi.nlm.nih.gov/27591179/)]
16. Darley A, Coughlan B, Maguire R, McCann L, Furlong E. A bridge from uncertainty to understanding: the meaning of symptom management digital health technology during cancer treatment. *Digit Health*. Jan 22, 2023;9:20552076231152163. [FREE Full text] [doi: [10.1177/20552076231152163](https://doi.org/10.1177/20552076231152163)] [Medline: [36714543](https://pubmed.ncbi.nlm.nih.gov/36714543/)]
17. Smits M, Ludden GDS, Verbeek P, van Goor H. How digital therapeutics are urging the need for a paradigm shift: from evidence-based health care to evidence-based well-being. *Interact J Med Res*. Oct 20, 2022;11(2):e39323. [FREE Full text] [doi: [10.2196/39323](https://doi.org/10.2196/39323)] [Medline: [36264624](https://pubmed.ncbi.nlm.nih.gov/36264624/)]
18. Ayling K, Brierley S, Johnson B, Heller S, Eiser C. How standard is standard care? exploring control group outcomes in behaviour change interventions for young people with type 1 diabetes. *Psychol Health*. Jan 20, 2015;30(1):85-103. [FREE Full text] [doi: [10.1080/08870446.2014.953528](https://doi.org/10.1080/08870446.2014.953528)] [Medline: [25118842](https://pubmed.ncbi.nlm.nih.gov/25118842/)]
19. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. Mar 19, 2015;350(mar19 6):h1258-h1258. [FREE Full text] [doi: [10.1136/bmj.h1258](https://doi.org/10.1136/bmj.h1258)] [Medline: [25791983](https://pubmed.ncbi.nlm.nih.gov/25791983/)]
20. Hemkens LG. Benefit assessment of digital health applications-challenges and opportunities [Article in German]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. Oct 15, 2021;64(10):1269-1277. [FREE Full text] [doi: [10.1007/s00103-021-03413-x](https://doi.org/10.1007/s00103-021-03413-x)] [Medline: [34524477](https://pubmed.ncbi.nlm.nih.gov/34524477/)]
21. 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*. Aug 18, 2022;4:823977. [FREE Full text] [doi: [10.3389/fdgh.2022.823977](https://doi.org/10.3389/fdgh.2022.823977)] [Medline: [36060538](https://pubmed.ncbi.nlm.nih.gov/36060538/)]
22. Cunningham JA, Kypri K, McCambridge J. Exploratory randomized controlled trial evaluating the impact of a waiting list control design. *BMC Med Res Methodol*. Dec 06, 2013;13(1):150. [FREE Full text] [doi: [10.1186/1471-2288-13-150](https://doi.org/10.1186/1471-2288-13-150)] [Medline: [24314204](https://pubmed.ncbi.nlm.nih.gov/24314204/)]
23. Bendtsen M, Ulfsdotter Gunnarsson K, McCambridge J. Effects of a waiting list control design on alcohol consumption among online help-seekers: protocol for a randomised controlled trial. *BMJ Open*. Aug 26, 2021;11(8):e049810. [FREE Full text] [doi: [10.1136/bmjopen-2021-049810](https://doi.org/10.1136/bmjopen-2021-049810)] [Medline: [34446493](https://pubmed.ncbi.nlm.nih.gov/34446493/)]
24. Rutherford BR, Wager TD, Roose SP. Expectancy and the treatment of depression: a review of experimental methodology and effects on patient outcome. *Curr Psychiatry Rev*. Feb 01, 2010;6(1):1-10. [FREE Full text] [doi: [10.2174/157340010790596571](https://doi.org/10.2174/157340010790596571)] [Medline: [24812548](https://pubmed.ncbi.nlm.nih.gov/24812548/)]
25. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. Jan 06, 2012;33(2):176-182. [doi: [10.1093/eurheartj/ehr352](https://doi.org/10.1093/eurheartj/ehr352)] [Medline: [21900289](https://pubmed.ncbi.nlm.nih.gov/21900289/)]
26. Mahajan R, Gupta K. Adaptive design clinical trials: methodology, challenges and prospect. *Indian J Pharmacol*. Aug 2010;42(4):201-207. [FREE Full text] [doi: [10.4103/0253-7613.68417](https://doi.org/10.4103/0253-7613.68417)] [Medline: [20927243](https://pubmed.ncbi.nlm.nih.gov/20927243/)]
27. Mohr DC, Schueller SM, Riley WT, Brown CH, Cuijpers P, Duan N, et al. Trials of intervention principles: evaluation methods for evolving behavioral intervention technologies. *J Med Internet Res*. Jul 08, 2015;17(7):e166. [FREE Full text] [doi: [10.2196/jmir.4391](https://doi.org/10.2196/jmir.4391)] [Medline: [26155878](https://pubmed.ncbi.nlm.nih.gov/26155878/)]

28. Subbiah V. The next generation of evidence-based medicine. *Nat Med.* Jan 16, 2023;29(1):49-58. [doi: [10.1038/s41591-022-02160-z](https://doi.org/10.1038/s41591-022-02160-z)] [Medline: [36646803](https://pubmed.ncbi.nlm.nih.gov/36646803/)]
29. Tarricone R, Petracca F, Ciani O, Cucciniello M. Distinguishing features in the assessment of mHealth apps. *Expert Rev Pharmacoecon Outcomes Res.* Aug 04, 2021;21(4):521-526. [doi: [10.1080/14737167.2021.1891883](https://doi.org/10.1080/14737167.2021.1891883)] [Medline: [33663324](https://pubmed.ncbi.nlm.nih.gov/33663324/)]
30. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *J Med Internet Res.* Jun 29, 2017;19(6):e232. [FREE Full text] [doi: [10.2196/jmir.7126](https://doi.org/10.2196/jmir.7126)] [Medline: [28663162](https://pubmed.ncbi.nlm.nih.gov/28663162/)]

Abbreviations

DHS: digital health solution
HCP: health care professional
KC: key characteristic
MoA: mechanism of action
RCT: randomized controlled trial
TAU: treatment as usual

Edited by T de Azevedo Cardoso; submitted 13.11.23; peer-reviewed by F Denis, M Savolainen; comments to author 14.02.24; revised version received 05.03.24; accepted 17.05.24; published 17.06.24

Please cite as:

Bolinger E, Tyl B

Key Considerations for Designing Clinical Studies to Evaluate Digital Health Solutions

J Med Internet Res 2024;26:e54518

URL: <https://www.jmir.org/2024/1/e54518>

doi: [10.2196/54518](https://doi.org/10.2196/54518)

PMID: [38885020](https://pubmed.ncbi.nlm.nih.gov/38885020/)

©Elaina Bolinger, Benoit Tyl. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 17.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.