

## PERSPECTIVE OPEN



# Selecting and describing control conditions in mobile health randomized controlled trials: a proposed typology

Simon B. Goldberg<sup>1,2</sup>✉, Shufang Sun<sup>3,4,5</sup>, Per Carlbring<sup>6</sup> and John Torous<sup>7,8</sup>

Hundreds of randomized controlled trials (RCTs) have tested the efficacy of mobile health (mHealth) tools for a wide range of mental and behavioral health outcomes. These RCTs have used a variety of control condition types which dramatically influence the scientific inferences that can be drawn from a given study. Unfortunately, nomenclature across mHealth RCTs is inconsistent and meta-analyses commonly combine control conditions that differ in potentially important ways. We propose a typology of control condition types in mHealth RCTs. We define 11 control condition types, discuss key dimensions on which they differ, provide a decision tree for selecting and identifying types, and describe the scientific inferences each comparison allows. We propose a five-tier comparison strength gradation along with four simplified categorization schemes. Lastly, we discuss unresolved definitional, ethical, and meta-analytic issues related to the categorization of control conditions in mHealth RCTs.

*npj Digital Medicine* (2023)6:181 ; <https://doi.org/10.1038/s41746-023-00923-7>

## INTRODUCTION

The World Health Organization defines mobile health (mHealth) as the “use of information and communication technologies in support of health and health-related fields”<sup>1</sup>. There has been immense research interest in this area in the past decade. For example, a recent review of mobile phone-based interventions for mental health included 145 randomized controlled trials (RCTs) and 47,940 participants, with most RCTs published in the last 10 years<sup>2</sup>. Many mHealth tools are being tested and, beginning in 2020, some have earned Food and Drug Administration (FDA) approval<sup>3,4</sup>. mHealth is projected to become a USD\$250 billion industry by 2026<sup>5</sup>.

RCTs are central for establishing efficacy in medicine, with the double-blind placebo-controlled RCT long being the gold standard<sup>6,7</sup>. The ability to blind patient and provider to group assignment allows a rigorous test of efficacy that accounts for non-specific factors such as expectancy. This method is ideal for testing pharmacological interventions where masking can be relatively easily done through placebo medications.

Double-blind placebo-controlled trials have also been used to test non-pharmacological interventions, such as surgery<sup>8</sup>. However, the translation of this design to behavioral interventions is not straightforward. It can be difficult to standardize interventions that include human interaction, as in psychotherapy<sup>9</sup>. Moreover, behavioral interventions commonly include multiple components that are not easily separated. Although mHealth interventions may or may not include human interaction, these approaches often include a combination of features (e.g., mood tracking, mindfulness)<sup>10</sup> which makes isolating active ingredients difficult without close attention to control conditions.

Several guidelines have been published for selecting and describing control conditions in behavioral interventions<sup>11–14</sup>. These guidelines describe a variety of control condition types and factors to be considered when selecting a control group

(e.g., trial phase and aim, need to balance internal validity and statistical power)<sup>11–13</sup>. However, there are no established guidelines for designing and categorizing control conditions for mHealth RCTs<sup>15,16</sup> (although some meta-analyses of the mHealth literature have tested this design feature as a moderator)<sup>17</sup>. Extant guidelines for behavioral interventions do not necessarily map neatly onto the mHealth literature where intervention and control conditions vary widely in form, intensity, and delivery setting. Currently, the same term—“control”—can be used across RCTs to refer to quite different control conditions<sup>18</sup>. “Treatment-as-usual” (TAU), “care-as-usual” and “usual care” are notoriously ambiguous designations. For example, TAU can be quite minimal (e.g., permission to pursue treatment outside of the study, referrals to non-study-related providers)<sup>19</sup> or fairly extensive (e.g., receipt of pharmacotherapy and/or psychotherapy)<sup>20–22</sup>. At times, studies may also refer to a group receiving TAU as a waitlist control<sup>23</sup>. Ambiguity defining and categorizing control conditions impacts attempts to synthesize trends across mHealth RCTs. Meta-analyses commonly combine substantially different control conditions, which complicates interpretation of results<sup>24</sup>.

As mHealth moves toward wider adoption within health systems<sup>3</sup>, there is a need for a consistently applied typology of control conditions. A lack of clarity can lead to misunderstanding and disappointment<sup>25</sup>. If control conditions are interpreted as a homogeneous group, interventions compared with stronger control conditions may be perceived as less effective<sup>18,26</sup>. Conversely, it may become clear relatively late in the evaluation process that interventions with promising pilot data do not outperform even minimally active controls<sup>27</sup>. Clarifying the universe of mHealth control condition types can help highlight the scientific questions each type can reasonably address and comparison strength each type provides<sup>12</sup>.

<sup>1</sup>Department of Counseling Psychology, University of Wisconsin–Madison, Madison, WI, USA. <sup>2</sup>Center for Healthy Minds, University of Wisconsin–Madison, Madison, WI, USA.

<sup>3</sup>Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, USA. <sup>4</sup>International Health Institute, Brown University School of Public Health, Providence, RI, USA. <sup>5</sup>Mindfulness Center, Brown University School of Public Health, Providence, RI, USA. <sup>6</sup>Department of Psychology, Stockholm University, Stockholm, Sweden. <sup>7</sup>Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA. <sup>8</sup>Harvard Business School, Boston, MA, USA.

✉email: sbgoldberg@wisc.edu

## CURRENT STUDY

Building on guidelines developed for behavioral interventions<sup>11–14</sup>, we propose a typology for selecting and defining control condition types in the context of mHealth RCTs. We highlight relevant dimensions of differentiation and provide a decision tree for identifying control conditions. We intend this typology to be comprehensive and capable of differentiating between the wide variety of control conditions appearing in the mHealth literature. For this reason, we define 11 distinct types. We clarify what scientific question each can answer and propose a five-tier comparison strength grading scheme. In addition, we propose simplified categorizations (from five to two categories) which may be more usable for meta-analysis. Lastly, we highlight open questions and future directions for control conditions in mHealth RCTs.

## PROPOSED TYPOLOGY

### Relevant dimensions

Our proposed typology considers five dimensions (Table 1): intended to be therapeutic, intensity, intensity match, masking/expectancy match, and received by both groups. Therapeutic intention is a critical first consideration<sup>28</sup>. Control conditions that are not intended to be therapeutic have an upper bound on the comparison strength they provide. Studies with non-therapeutic control conditions are still scientifically valuable as they can establish the impact of non-specific factors<sup>12</sup>, but they cannot provide the highest comparison strength. Control condition intensity is an important second consideration, particularly in the mHealth context where many focal interventions are self-guided and of potentially modest intensity. A third consideration is whether the intensity is matched between the focal intervention and the control condition. This can be challenging to establish, particularly given inconsistencies in the definition and reporting of engagement with mHealth interventions<sup>29,30</sup>. A fourth consideration is whether the treatment condition is masked and in such a way that expectancy is likely to be matched between focal and control conditions. This is arguably the characteristic feature of an ideal placebo control and any control conditions capable of clarifying intervention effects beyond non-specific factors<sup>31,32</sup>. It is important to acknowledge that clinical trialists, despite their best intentions, may or may not succeed in matching expectancy, and the development of psychotherapy placebos remains a challenging (and some might say impossible)<sup>9,33</sup> task. A fifth consideration is whether the control condition is received by both focal intervention and control groups. This is especially relevant in the case of TAU controls in which both focal intervention and control conditions receive the TAU<sup>23</sup>.

### Decision tree

Figure 1 displays a decision tree with the 11 control condition types. Table 1 provides examples from the literature for each type. A first decision point is whether a control condition is intended to be therapeutic<sup>28</sup>. Non-therapeutic control conditions are further defined based on intensity: none (No Treatment), low (Placebo-Minimal), and high (Placebo-Active). The comparison strength increases as intensity increases.

There are a wide variety of control conditions that are, at least to some degree, intended to be therapeutic. Within therapeutic controls, the first distinction is whether the control condition is also an mHealth intervention. mHealth RCTs focused on isolating the effects of a particular intervention ingredient may use a Dismantling Design (i.e., therapeutic ingredient subtracted) or an Additive Design (i.e., therapeutic ingredient added)<sup>34</sup>. These designs provide a strong comparison for the specific ingredients being tested, although alone they cannot be used to evaluate the effects of an intervention as a whole. mHealth RCTs interested in

comparing the efficacy of two or more mHealth interventions may use an mHealth Minimal Active control that is intended to be therapeutic but intended to be less potent than the focal intervention. The relative efficacy of two or more mHealth interventions can also be tested using a mHealth Comparative Efficacy control. mHealth Comparative Efficacy control conditions could be mHealth interventions with known efficacy (e.g., already FDA-approved)<sup>3</sup>, similar to comparisons conducted between novel pharmacotherapies and FDA-approved medications<sup>35</sup>.

mHealth RCTs often include non-mHealth intervention components. If both groups receive these elements, such control conditions are defined as TAU. However, TAU can vary in intensity, so this category is further graded into TAU-Minimal and TAU-Active. The determination of whether a control condition is TAU-Minimal or TAU-Active should be defined based on the context of standard care and the purpose of the focal intervention. For instance, for depression, TAU-Minimal could involve providing referrals to local clinicians with no additional follow-up<sup>19</sup>. An example of TAU-Active would be an active intervention received by all participants (e.g., antidepressants) to which mHealth is added as an adjunctive treatment<sup>36</sup>. For studying medication adherence among people living with HIV<sup>21,37</sup>, TAU-Minimal could be antiretroviral TAU (including the clinic's standard care such as regular physician visits and relevant referrals). TAU-Active could include antiretroviral TAU along with an adherence-focused support group (i.e., an added element above and beyond standard care).

Some designs compare non-mHealth care with mHealth, and therefore do not provide non-mHealth care to the mHealth arm. These designs can provide a rigorous test of the relative efficacy of mHealth. We define a two-level distinction based on whether the non-mHealth control is a frontline evidence-based treatment (non-mHealth EBT Active), providing the highest comparison strength<sup>38</sup>, or another active control (non-mHealth Other Active) that is intended to be therapeutic but is not a frontline EBT.

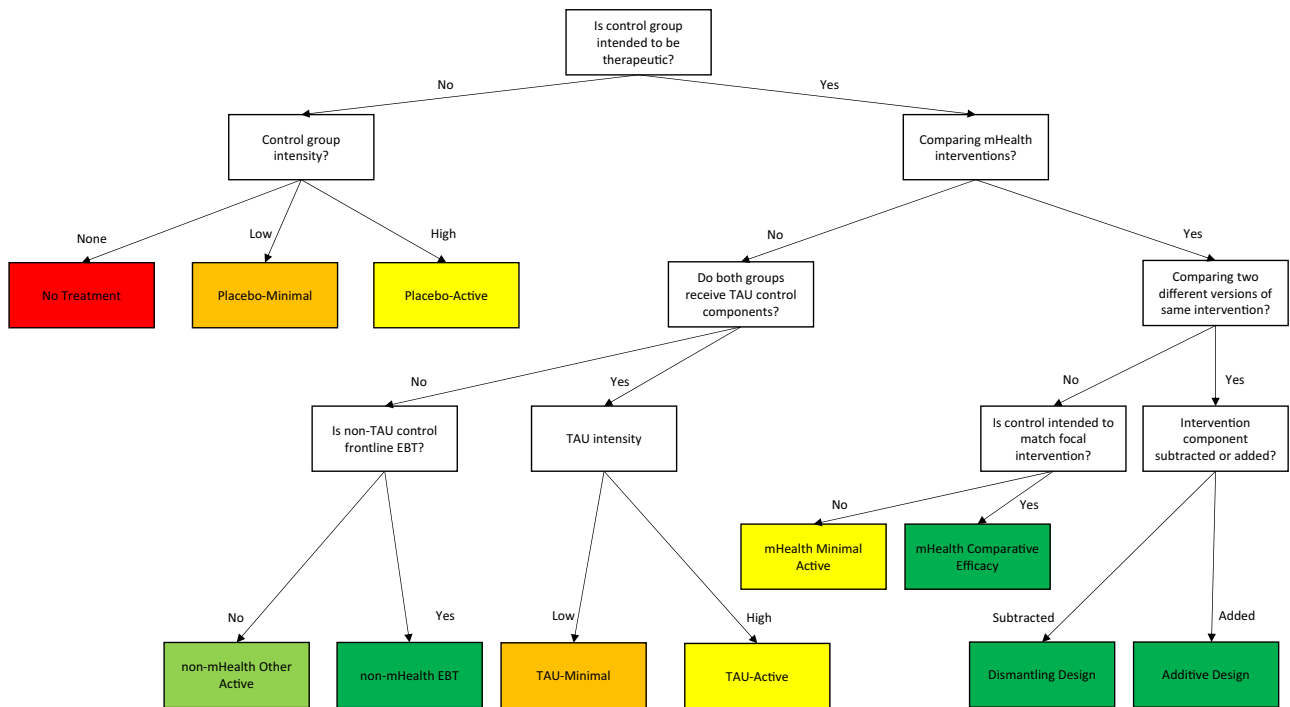
### Scientific inferences

Table 2 describes the scientific inferences possible for each control group. There is no single control condition that is appropriate for every study at every stage of the research process<sup>11,12,39</sup>. Indeed, as discussed in previous guidelines<sup>11,13</sup>, clinical trialists are faced with balancing controlling threats to internal validity (where stronger comparisons provide more confidence that observed differences between intervention and control groups are due to the intervention itself) with statistical power (where stronger comparisons are expected to produce smaller between-group differences that require larger sample sizes for detection). In early stages of research<sup>39,40</sup>, No Treatment controls may be very appropriate in order to evaluate intervention effects beyond the passage of time and related confounds (e.g., regression to the mean)<sup>41</sup>. At later stages of research, stronger comparisons are warranted<sup>13</sup>. Placebo-Minimal controls for some non-specific factors (e.g., minimal amounts of time, attention, or expectancy) while a Placebo-Active can theoretically control for all non-specific factors, particularly when time and expectancy are matched. As noted, Dismantling Design and Additive Design controls are ideal for clarifying the effects of an intervention without a key ingredient or with a key ingredient added, respectively. Scientific questions related to comparisons with other mHealth interventions are best addressed using mHealth Minimal Active controls or mHealth Comparative Efficacy controls. These controls can establish the degree to which a focal intervention outperforms interventions expected to be of minimal strength (mHealth Minimal Active) or similar strength (mHealth Comparative Efficacy).

Four control conditions can be used to evaluate the effects of mHealth in the context of non-mHealth interventions. Studies

**Table 1.** Eleven-category scheme with relevant dimensions of differentiation between focal intervention and control condition.

Control condition	Example	Intended to be therapeutic (yes/no)	Intensity (none/low/high)	Intensity match (no/maybe/yes)	Masking/expectancy match (no/maybe/yes)	Received by both groups (no/maybe/yes)
No Treatment	Waitlist vs. virtual reality smartphone app <sup>61</sup>	No	None	No	No	No
Placebo-Minimal	List-making app vs. meditation app <sup>62</sup>	No	Low	No	Maybe	No/maybe
Placebo-Active	Attention training app vs. attentional bias modification app <sup>63</sup>	No	High	Yes	Yes	No
TAU-Minimal	Antiretroviral psychoeducation + mood monitoring texts vs. antiretroviral psychoeducation + mood monitoring texts + personalized adherence reminder texts <sup>64</sup>	Yes	Low	No	No	Yes
TAU-Active	Antidepressants vs. antidepressants + CBT app <sup>36</sup>	Yes	High	No	No	Yes
mHealth Minimal	Psychoeducation control app vs. problem-solving therapy app <sup>65</sup>	Yes	High	Maybe	Yes	No
Active						
Dismantling Design	Mindfulness app with acceptance removed vs. mindfulness app <sup>66</sup>	Yes	High	Maybe	Maybe	No
Additive Design	Weight Watchers digital program vs. Weight Watchers digital program + just-in-time adaptive dietary lapse prevention smartphone app	Yes	High	Maybe	Maybe	No
mHealth Comparative Efficacy	Cognitive control app vs. problem-solving therapy app <sup>65</sup>	Yes	High	Maybe	Yes	No
non-mHealth Other	Weekend retreat alone vs. mindfulness and psychoeducation app <sup>67</sup>	Yes	High	Maybe	Yes	No
Active						
non-mHealth EBT	Clinician-guided online CBT vs. self-guided online CBT <sup>68</sup>	Yes	High	Maybe	Yes	No
mHealth mobile health, EBT evidence-based treatment (i.e., frontline intervention), CBT cognitive behavioral therapy.						



**Fig. 1 Eleven-category typology decision tree.** mHealth mobile health, EBT evidence-based treatment (i.e., frontline intervention), TAU treatment-as-usual. Color coding reflects comparison strength ranging from low (Class V, red) to high (Class I, darker green). The choice of control condition type must ultimately be guided by the scientific question motivating a given study, which may be strongly influenced by a variety of factors including the stage of research (i.e., Class V comparisons may be very appropriate at early stages in the development of an intervention)<sup>11,13,40</sup>.

focused on evaluating the effect of augmenting treatment with mHealth can use a TAU-Minimal or TAU-Active control. These control conditions allow evaluation of the effects of mHealth as an adjunctive to another treatment. The final two control conditions address the degree to which a mHealth intervention performs on par with a non-mHealth intervention. Non-mHealth Other Active controls clarify whether mHealth interventions outperform non-EBT non-mHealth interventions while non-mHealth EBT Active controls clarify whether mHealth interventions outperform frontline, non-mHealth interventions.

### Comparison strength

The choice of control condition heavily impacts the comparison strength within a study<sup>12</sup>. Table 2 displays a five-tier scheme for grading the comparison strength provided by the 11 control conditions. No Treatment controls provide the lowest tier (Class V). Class IV comparison strength can be derived from either Placebo-Minimal or TAU-Minimal, as both of these control conditions are expected to be either non-therapeutic or minimally therapeutic and received by both groups. They provide only slightly stronger comparison than No Treatment controls. Class III comparison strength can be derived from Placebo-Active, TAU-Active, or mHealth Minimal Active controls. The Placebo-Active control provides a stronger comparison than the Placebo-Minimal in its ability to match on a larger number of non-specific factors (e.g., expectancy). Likewise, TAU-Active sets a higher bar for demonstrating the effects of mHealth than TAU-Minimal, with both arms receiving an active therapeutic intervention. mHealth Minimal Active provides evidence of effects beyond a limited mHealth intervention which, although stronger than comparisons with somewhat or entirely inert controls (Class IV and Class V, respectively), is only a moderately strong comparison given mHealth conditions are not matched.

The highest classes of comparison strength (Class II and I) require therapeutic control conditions that may be matched in intensity and non-specific factors including expectancy. These comparisons may be most appropriate for the later stages of intervention testing<sup>39,40</sup>. Non-EBT non-mHealth controls provide strong comparisons (Class II), but weaker comparisons than frontline, non-mHealth EBTs. Dismantling Design and Additive Design controls provide very strong comparisons (Class I), but only for the components being subtracted or added and not for the intervention as a whole. The highest comparison strength (Class I) for the mHealth intervention as a whole requires comparisons with either a matched mHealth Comparative Efficacy control or a non-mHealth EBT control.

### Alternative coding schemes

The 11-category scheme is designed to characterize important differences between control conditions in mHealth RCTs. Nonetheless, this scheme is detailed and there may be advantages to having simpler schemes (e.g., for meta-analysis). Here we propose four simplified schemes (Table 3).

**Five-category.** A five-category scheme collapses across levels of placebo controls, TAU controls, mHealth controls, and non-mHealth controls. This scheme retains distinctions between controls intended to be therapeutic (TAU, mHealth controls, non-mHealth controls) and non-therapeutic (No Treatment, Placebo). Although this scheme can help describe the types of comparisons, it ignores variations in intensity and evidence base which may influence scientific inferences and effect sizes.

**Four-category.** This scheme also distinguishes between therapeutic and non-therapeutic controls, defined as non-specific and specific active controls<sup>42,43</sup>, while combining minimal treatment controls into a single group. The four categories include: No

**Table 2.** Scientific inferences and comparison strength provided by each control condition.

Control condition	Scientific inference	Comparison strength
No Treatment	Intervention effects vs. passage of time/regression to the mean	Class V
Placebo-Minimal	Intervention effects beyond some non-specific factors (minimal time match, minimal expectancy)	Class IV
Placebo-Active	Intervention effects beyond non-specific factors (time match, expectancy)	Class III
TAU-Minimal	Intervention effects beyond minimal TAU	Class IV
TAU-Active	Intervention effects beyond active TAU, as adjunctive treatment	Class III
mHealth Minimal Active	Intervention effects relative to a minimal similar intervention	Class III
Dismantling Design	Intervention effects with component removed	Class I (for component)
Additive Design	Intervention effects with component added	Class I (for component)
mHealth Comparative Efficacy	Intervention effects relative to a similarly intensive mHealth intervention	Class I
non-mHealth Other Active	Intervention effects equivalent to established non-mHealth intervention	Class II
non-mHealth EBT	Intervention effects equivalent to frontline non-mHealth intervention	Class I

Comparison strength ranges from low (Class V) to high (Class I).  
*mHealth* mobile health, *EBT* evidence-based treatment (i.e., frontline intervention).

Treatment, Minimal Treatment (Placebo-Minimal and mHealth Minimal Active), Non-Specific Active (Placebo-Active), and Specific Active (Dismantling Design, Additive Design, mHealth Comparative Efficacy, non-mHealth Other Active, non-mHealth EBT). One control condition that is challenging to include in this scheme is TAU-Active. While it may be defensible to view a minimal amount of treatment received by both focal and controls arms as essentially canceling each other out (i.e., yielding a control condition more similar to a No Treatment control than an active control), this case is harder to make for TAU-Active controls that includes intensive, therapeutic interventions received by both groups (e.g., antidepressants)<sup>36</sup>.

*Two-category: passive vs. active.* A first two-category scheme involves differentiating between Passive vs. Active controls. In this scheme, Passive controls include No Treatment and TAU-Minimal. Active controls include the remaining conditions. A downside of this scheme is it ignores whether a control is intended to be therapeutic. From our perspective, this scheme is not ideal, although, of note, it is used within the meta-analytic literature<sup>44</sup> and is arguably preferable to collapsing across all control conditions<sup>24,45</sup>.

*Two-category: therapeutic vs. non-therapeutic.* A second two-category scheme differentiates between control conditions intended to be therapeutic and not intended to be therapeutic<sup>28</sup>. We view this as the minimally viable distinction that should be made. However, it requires collapsing across a theoretically wide range of non-therapeutic (i.e., No Treatment, Placebo-Minimal, Placebo-Active) and therapeutic controls. This distinction is still likely superior to considering all control conditions the same.

## OPEN QUESTIONS AND FUTURE DIRECTIONS

The control condition coding schemes, ranging from simplistic two-category schemes to an 11-category scheme, are intended to highlight important differences between the variety of control conditions used in mHealth RCTs. At once, these schemes are all imperfect and are offered as an initial attempt to categorize the highly diverse mHealth literature. Here we discuss several open questions regarding categorizing control conditions in mHealth RCTs and highlight areas for future research.

## Fuzzy boundaries and edge cases

As an initial attempt, we have provided general guidance to distinguish between categories. However, it is currently unclear how several aspects should be handled, making the boundaries between categories fuzzy. One important issue is how precisely to determine whether focal and control interventions are matched on a particular dimension, especially in terms of intensity and expectancy. As noted, engagement with mHealth interventions can be highly variable<sup>29,30</sup>. Although focal and control interventions may both in theory be equally intensive, participants may not engage with them both equally. Similarly, expectancy may or may not actually be matched across conditions. Moreover, clinical trialists may reasonably disagree about how a particular control condition should be categorized. For example, one trial may describe ecological momentary assessment as “mood monitoring”<sup>46,47</sup> and consider it a mHealth Minimal Active control while another considers it a Placebo-Active, Placebo-Minimal, or even No Treatment control. Arguably, it is incumbent upon mHealth trialists to demonstrate that control conditions intended to be intensity and/or expectancy matched are indeed matched. These data may inform how a control condition is described. For example, a placebo condition that fails to demonstrate adequate intensity and expectancy match may be better described as Placebo-Minimal or No Treatment than Placebo-Active. The boundary between TAU-Minimal and TAU-Active may often be fuzzy in contexts where TAU is heterogeneous (i.e., patients receive differing amounts and/or type of treatment).

Several other aspects of modern mHealth RCTs do not fit neatly into the 11-category scheme. Adaptive designs in which the treatment that is delivered changes over time (e.g., sequential multiple assignment randomized controlled trial [SMART])<sup>48</sup> and factorial designs that evaluate multiple aspects of an mHealth intervention at once are challenging to categorize. Such studies collapse across various intervention elements in analyses to evaluate a range of comparisons within a single trial. For example, a SMART study may initially randomize participants to one of two app conditions (e.g., mHealth comparative efficacy control) but re-randomize participants who do not show symptom reductions after a set period of time to additional intervention components (e.g., coaching [non-mHealth Other Active]). Some analyses of SMART and factorial trials may involve a particular control condition type (e.g., analyses collapsing across conditions receiving a non-mHealth intervention component [non-mHealth Other Active]) while other analyses may involve another control condition type (e.g., analyses collapsing across conditions that

**Table 3.** Alternative coding schemes.

Control condition	Five-category	Four-category	Two-category active vs. non-active	Two-category therapeutic vs. non-therapeutic
No Treatment	No Treatment	No Treatment	Passive	Not therapeutic
Placebo-Minimal	Placebo	Minimal Treatment	Active	Not therapeutic
Placebo-Active	Placebo	Non-Specific Active	Active	Not therapeutic
TAU-Minimal	TAU	No Treatment	Passive	Therapeutic
TAU-Active	TAU	Unclear	Unclear	Therapeutic
mHealth Minimal Active	mHealth Comparison	Minimal Treatment	Active	Therapeutic
Dismantling Design	mHealth Comparison	Specific Active	Active	Therapeutic
Additive Design	mHealth Comparison	Specific Active	Active	Therapeutic
mHealth Comparative Efficacy	mHealth Comparison	Specific Active	Active	Therapeutic
non-mHealth Other Active	non-mHealth Comparison	Specific Active	Active	Therapeutic
non-mHealth EBT	non-mHealth Comparison	Specific Active	Active	Therapeutic

*mHealth* mobile health, *EBT* evidence-based treatment (i.e., frontline intervention), *Non-specific active* active control not intended to be therapeutic, *Specific Active* active control intended to be therapeutic.

did not receive a specific additional mHealth component [Additive Design]). Dismantling Designs, although categorized as intended to be therapeutic in our scheme, may have removed all theoretically therapeutic ingredients<sup>34</sup> and therefore be better categorized as Placebo-Active.

The contribution of study-related human support in mHealth RCTs also makes categorization challenging. Studies commonly involve interaction between participants and study staff, although it can be hard to evaluate the extent and nature of this interaction from published RCTs<sup>49</sup>. Clinical trialists clearly reporting details regarding human support within RCTs and the degree to which such interactions may have been therapeutic can aid in categorization (e.g., a No Treatment control with ongoing, therapeutic human support may be better categorized as TAU-Minimal).

### Ethics

One issue not addressed by our scheme, but of paramount importance, is the ethical dimension of study design. Selecting a control condition has ethical implications around access to care, scientific rigor, and bias. Particularly in under-resourced settings and when working with vulnerable populations, No Treatment control conditions may be inappropriate as they may involve withholding interventions with known therapeutic value<sup>50,51</sup>. In contexts where trust is limited between community members and academic research teams, providing an active control may be necessary for building trust and completing recruitment and enrollment goals for an RCT. In these contexts, it may be more appropriate to provide at least a minimal degree of intervention (e.g., TAU-Minimal), for example links to publicly available online resources<sup>52</sup>. As a program of research matures (e.g., moves through the stages of intervention development)<sup>11,39,40</sup>, increasingly rigorous and theoretically therapeutic control conditions may be warranted, both for ethical (e.g., beneficence and non-maleficence)<sup>53</sup> and scientific reasons (i.e., determining how a novel treatment compares with established treatments). As a general rule, clinical trials should only be run when there is the expectation that the treatment being studied may be beneficial (i.e., equipoise)<sup>54</sup>. The ethical use of digital health tools remains an active area of investigation, with frameworks emerging but none yet consistently adopted<sup>55,56</sup>.

### Application to meta-analysis

The proposed coding schemes are intended to be helpful for meta-analysts faced with categorizing the diverse range of control conditions used in mHealth RCTs. At a minimum, we hope this helps avoid considering control conditions to be a monolithic category<sup>57</sup>. Distinguishing between controls intended to be therapeutic and not intended to be therapeutic provides perhaps the simplest theoretically justified categorization. New meta-analyses sensitive to the impact of controls groups may provide more accurate effect size estimates and results of prior meta-analyses may need to be reexamined in light of these considerations.

### Limitations and future directions

As control conditions can dramatically impact effect sizes observed in RCTs of psychological interventions generally<sup>58,59</sup> and mHealth RCTs specifically<sup>44</sup>, considering control conditions to be a single category is typically not defensible. The framework we propose here may help guide the selection of a control group, but it cannot definitively identify the ideal one for any particular study. Like all frameworks, its value must be proven in its utility. Thus, several future directions follow. A valuable next step would be the development and adoption of detailed control condition categorization guidelines that can be used by both clinical trialists and meta-analysts. It would be helpful to convene a Delphi panel of global experts for this purpose, which ideally would result in a checklist that can be included with mHealth RCTs to clarify and justify choice of control condition. Relatedly, it may be helpful to identify and/or develop specific measures that can be used to help delineate between similar control conditions (e.g., demonstrating intensity and expectancy match for Placebo-Active and mHealth Comparative Efficacy controls). It would be valuable to code the existing mHealth RCT literature using the schemes proposed here. A future meta-analysis could evaluate the degree to which between-group effect sizes vary across categories. In theory, between-group effect sizes should become smaller as the comparison strength increases from Class V to Class I. As consensus is reached regarding the categorization of control conditions in mHealth RCTs, this information can be used by bodies charged with regulating these technologies (e.g., FDA)<sup>60</sup>, researchers seeking to adequately power mHealth RCTs, industry stakeholders involved in investing and marketing, and ultimately patients and clinicians deciding when to use a given mHealth tool.

Received: 11 December 2022; Accepted: 20 September 2023;  
Published online: 30 September 2023

## REFERENCES

- WHO. *mHealth Use of Appropriate Digital Technologies for Public Health*. <https://doi.org/10.1371/journal.pmed.1001362> (2018).
- Goldberg, S. B., Lam, S. U., Simonsson, O., Torous, J. & Sun, S. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. *PLoS Digit. Health* **1**, e0000002 (2022).
- FDA. FDA clears mobile medical app to help those with opioid use disorder stay in recovery programs. <https://www.fda.gov/news-events/press-announcements/fda-clears-mobile-medical-app-help-those-opioid-use-disorder-stay-recovery-programs> (2022).
- FDA. FDA permits marketing of first game-based digital therapeutic to improve attention function in children with ADHD. <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-first-game-based-digital-therapeutic-improve-attention-function-children-adhd> (2022).
- TBRC. The mobile healthcare market is expected to reach \$250. <https://www.globenewswire.com/news-release/2022/03/29/2412162/0/en/The-Mobile-Healthcare-Market-Is-Expected-To-Reach-250-Billion-Due-To-The-Rising-Prevalence-Of-Chronic-Diseases-Globally-As-Per-The-Business-Research-Company-s-Mobile-Healthcare-Glo.html> (2022).
- Hariton, E. & Locascio, J. J. Randomised controlled trials—the gold standard for effectiveness research. *BJOG* **125**, 1716 (2018).
- Friedman, L. M., Furburg, C. D., DeMets, D. L., Reboussin, D. M. & Granger, C. B. *Fundamentals of Clinical Trials* (Springer, 2015).
- Heger, P. et al. A systematic review and meta-analysis of physical exercise prehabilitation in major abdominal surgery (PROSPERO 2017 CRD42017080366). *J. Gastrointest. Surg.* **24**, 1375–1385 (2020).
- Wampold, B. E. & Imel, Z. E. *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work* 2nd edn (Routledge, 2015).
- Lagan, S., D'Mello, R., Vaidyam, A., Bilden, R. & Torous, J. Assessing mental health apps marketplaces with objective metrics from 29,190 data points from 278 apps. *Acta Psychiatr. Scand.* **144**, 201–210 (2021).
- Mohr, D. C. et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychother. Psychosom.* **78**, 275–284 (2009).
- Freedland, K. E. et al. The selection of comparators for randomized controlled trials of health-related behavioral interventions: recommendations of an NIH expert panel. *J. Clin. Epidemiol.* **110**, 74–81 (2019).
- Gold, S. M. et al. Control conditions for randomised trials of behavioural interventions in psychiatry: a decision framework. *Lancet Psychiatry* **4**, 725–732 (2017).
- Freedland, K. E. Purpose-guided trial design in health-related behavioral intervention research. *Health Psychol.* **39**, 539–548 (2020).
- Lutz, J., Offidani, E., Tarabonta, L., Lakhan, S. E. & Campellone, T. R. 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* **4**, 823977 (2022).
- Blease, C. Out of control: how to design digital placebos. *Curr. Treat. Options Psychiatry* 1–10 <https://doi.org/10.1007/S40501-023-00290-W> (2023).
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M. & Fuller-Tyszkiewicz, M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* **18**, 325–336 (2019).
- Wampold, B. E. et al. Evidence-based treatments for depression and anxiety versus treatment-as-usual: a meta-analysis of direct comparisons. *Clin. Psychol. Rev.* **31**, 1304–1312 (2011).
- Gao, M. et al. Targeting anxiety to improve sleep disturbance: a randomized clinical trial of app-based mindfulness training. *Psychosom. Med.* **84**, 632–642 (2022).
- Naughton, F. et al. Randomized controlled trial to assess the short-term effectiveness of tailored web- and text-based facilitation of smoking cessation in primary care (iQuit in Practice). *Addiction* **109**, 1184–1193 (2014).
- Lee, S. B. & Valerius, J. mHealth interventions to promote anti-retroviral adherence in HIV: narrative review. *JMIR Mhealth Uhealth* **8**, e14739 (2020).
- Keeler, J. L. et al. App-based food-specific inhibitory control training as an adjunct to treatment as usual in binge-type eating disorders: a feasibility trial. *Appetite* **168**, 105788 (2022).
- Graham, A. K. et al. Coached mobile app platform for the treatment of depression and anxiety among primary care patients: a randomized clinical trial. *JAMA Psychiatry* **77**, 906–914 (2020).
- Bendtsen, M., McCambridge, J., Åsberg, K. & Bendtsen, P. Text messaging interventions for reducing alcohol consumption among risky drinkers: systematic review and meta-analysis. *Addiction* **116**, 1021–1033 (2021).
- Smoktunowicz, E. et al. Consensus statement on the problem of terminology in psychological interventions using the internet or digital components. *Internet Interv.* **21**, 100331 (2020).
- Goldberg, S. B., Riordan, K. M., Sun, S. & Davidson, R. J. The empirical status of mindfulness-based interventions: a systematic review of 44 meta-analyses of randomized controlled trials. *Perspect. Psychol. Sci.* **17**, 174569162096877 (2022).
- Nassir Ghaemi, S., Sverdlow, O., Van Dam, J., Campellone, T. & Gerwien, R. A smartphone-based intervention as an adjunct to standard-of-care treatment for schizophrenia: randomized controlled trial. *JMIR Form. Res.* **6**, e29154 (2022).
- Wampold, B. E. et al. A meta-analysis of outcome studies comparing bona fide psychotherapies: empirically, 'all must have prizes'. *Psychol. Bull.* **122**, 203–215 (1997).
- Ng, M. M., Firth, J., Minen, M. & Torous, J. User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatr. Serv.* **70**, 538–544 (2019).
- Lipschitz, J. M. et al. Digital mental health interventions for depression: scoping review of user engagement. *J. Med. Internet Res.* **24**, e39204 (2022).
- Torous, J. & Firth, J. The digital placebo effect: mobile mental health meets clinical psychiatry. *Lancet Psychiatry* **3**, 100–102 (2016).
- Faria, V. et al. Do you believe it? Verbal suggestions influence the clinical and neural effects of escitalopram in social anxiety disorder: a randomized trial. *EBioMedicine* **24**, 179–188 (2017).
- Wampold, B. E., Frost, N. D. & Yulish, N. E. Placebo effects in psychotherapy: a flawed concept and a contorted history. *Psychol. Conscious. Theory Res. Pract.* **3**, 108–120 (2016).
- Bell, E. C., Marcus, D. K. & Goodlad, J. K. Are the parts as good as the whole? A meta-analysis of component treatment studies. *J. Consult. Clin. Psychol.* **81**, 722–736 (2013).
- FDA. The FDA's drug review process: ensuring drugs are safe and effective. <https://www.fda.gov/drugs/information-consumers-and-patients-drugs/fdas-drug-review-process-ensuring-drugs-are-safe-and-effective> (2022).
- Mantani, A. et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. *J. Med. Internet Res.* **19**, e8602 (2017).
- Mulawa, M. I., LeGrand, S. & Hightow-Weidman, L. B. eHealth to enhance treatment adherence among youth living with HIV. *Curr. HIV/AIDS Rep.* **15**, 336–349 (2018).
- Goodheart, C. D. et al. Evidence-based practice in psychology. *Am. Psychol.* **61**, 271–285 (2006).
- Onken, L. S., Carroll, K. M., Shoham, V., Cuthbert, B. N. & Riddle, M. Reenvisioning clinical science: unifying the discipline to improve the public health. *Clin. Psychol. Sci.* **2**, 22–34 (2014).
- Rounsaville, B. J., Carroll, K. M. & Onken, L. S. A stage model of behavioral therapies research: getting started and moving on from stage I. *Clin. Psychol. Sci. Pract.* **8**, 133–142 (2001).
- Bland, J. M. & Altman, D. G. Statistic notes: regression towards the mean. *BMJ* **308**, 1499 (1994).
- Goldberg, S. B. et al. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. *Clin. Psychol. Rev.* <https://doi.org/10.1016/j.cpr.2017.10.011> (2017).
- Galante, J. et al. Mindfulness-based programmes for mental health promotion in adults in nonclinical settings: a systematic review and meta-analysis of randomised controlled trials. *PLoS Med.* **18**, 1–40 (2021).
- Firth, J. et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* **16**, 287–298 (2017).
- Khoury, B., Sharma, M., Rush, S. E. & Fournier, C. Mindfulness-based stress reduction for healthy individuals: a meta-analysis. *J. Psychosom. Res.* **78**, 519–528 (2015).
- Webb, C. A., Swords, C. M., Lawrence, H. R. & Hilt, L. M. Which adolescents are well-suited to app-based mindfulness training? A randomized clinical trial and data-driven approach for personalized recommendations. *J. Consult. Clin. Psychol.* **90**, 655–669 (2022).
- Bakker, D. & Rickard, N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *J. Affect. Disord.* **227**, 432–442 (2018).
- Collins, L. M., Murphy, S. A. & Strecher, V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART). New methods for more potent eHealth interventions. *Am. J. Prev. Med.* **32**, S112–S118 (2007).
- Meyer, A., Wisniewski, H. & Torous, J. Coaching to support mental health apps: exploratory narrative review. *JMIR Hum. Factors* **9**, e28301 (2022).
- Boruch, R. F., Victor, T. & Cecil, J. S. Resolving ethical and legal problems in randomized experiments. *Crime & Delinquency* **46**, 330–353 (2016).

51. Pavone, I. R. Legal responses to placebo-controlled trials in developing countries. *Global Bioethics* **27**, 76–90 (2016).
52. NIMH. *Help for Mental Illnesses* (National Institutes for Health, 2022).
53. APA. Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code> (2022).
54. Freedman, B. Equipoise and the ethics of clinical research. *N. Engl. J. Med.* **317**, 427–431 (1987).
55. Nebeker, C., Bartlett Ellis, R. J. & Torous, J. Development of a decision-making checklist tool to support technology selection in digital health research. *Transl. Behav. Med.* **10**, 1004–1015 (2020).
56. Shen, F. X. et al. An ethics checklist for digital health research in psychiatry: viewpoint. *J. Med. Internet Res.* **24**, e31146 (2022).
57. Goldberg, S. B., Torous, J. & Sun, S. Let's decide what would be convincing, conduct randomized trials with rigorous comparison conditions, and report tests of moderation and publication bias in meta-analyses. *PLoS Digit. Health* **1**, e0000127 (2022).
58. Goldberg, S. B. et al. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. *Clin. Psychol. Rev.* **59**, 52–60 (2018).
59. Mohr, D. C. et al. Control condition design and implementation features in controlled trials: a meta-analysis of trials evaluating psychotherapy for depression. *Transl. Behav. Med.* **4**, 407–423 (2014).
60. Sim, I. Mobile devices and health. *N. Engl. J. Med.* **381**, 956–968 (2019).
61. Donker, T. et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. *JAMA Psychiatry* **76**, 682–690 (2019).
62. Howells, A., Ivtzan, I. & Eiroa-Orosa, F. J. Putting the 'app' in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *J. Happiness Stud.* **17**, 163–185 (2016).
63. Enock, P. M., Hofmann, S. G. & McNally, R. J. Attention bias modification training via smartphone to reduce social anxiety: a randomized, controlled multi-session experiment. *Cogn. Ther. Res.* **38**, 200–216 (2014).
64. Moore, D. J. et al. Individualized texting for adherence building (iTAB): improving antiretroviral dose timing among HIV-infected persons with co-occurring bipolar disorder. *AIDS Behav.* **19**, 459–471 (2015).
65. Arean, P. A. et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. *J. Med. Internet Res.* **18**, e6482 (2016).
66. Lindsay, E. K. et al. Mindfulness training reduces loneliness and increases social contact in a randomized controlled trial. *Proc. Natl. Acad. Sci.* **116**, 3488–3493 (2019).
67. Kahn, J. R., Collinge, W. & Soltysik, R. Post-9/11 veterans and their partners improve mental health outcomes with a self-directed mobile and web-based wellness training program: a randomized controlled trial. *J. Med. Internet Res.* **18**, e5800 (2016).
68. Dear, B. F. et al. Transdiagnostic versus disorder-specific and clinician-guided versus self-guided internet-delivered treatment for generalized anxiety disorder and comorbid disorders: a randomized controlled trial. *J. Anxiety Disord.* **36**, 63–77 (2015).

## ACKNOWLEDGEMENTS

S.B.G. was supported by the National Center for Complementary & Integrative Health of the National Institutes of Health (K23AT010879) and the Hope for Depression Research Foundation (Defeating Depression Award). S.S. was supported by the National Center for Complementary & Integrative Health of the National Institutes of Health (K23AT011173). J.T. was supported by a grant from the Baer Foundation.

## AUTHOR CONTRIBUTIONS

S.B.G. and J.T. conceptualized the study and wrote the first draft of the manuscript. All authors provided critical feedback on the manuscript and contributed to the final version of the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Simon B. Goldberg.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023