Questions and Answers

Mind the Gap — Deconstruction of the Type 2 Hybrid Effectiveness-Implementation Study Design that Uses Two Randomized Controlled Trials

June Stevens, Ph.D.

March 20, 2024

Q: What's the difference between SPIRIT and CONSORT guidelines?

A: SPIRIT addresses the preparation of a protocol and CONSORT addresses the preparation of a manuscript reporting results from a trial. Both focus on randomized trials. Neither has the explicit goal of guiding study rigor, but nevertheless, in the elaboration papers they address methodologic issues that impact rigor. The following quote from the SPIRIT 2013 statement that addresses the relationship between SPIRIT and CONSORT: "The SPIRIT 2013 Statement mirrors applicable items from CONSORT 2010 (Consolidated Standards of Reporting Trials) (43). Consistent wording and structure used for items common to both checklists will facilitate the transition from a SPIRIT-based protocol to a final report based on CONSORT."

Q: Do you ever measure participant adherence and/or participation as an implementation outcome?

A: In implementation, we often have multiple levels of participants. We often measure reach in those implementing the intervention - which includes documenting who participated in implementation strategies (the number, the percent of the total, and representativeness of those intended to participate). We also assess reach of the intervention (the number, the percent of the total, and representativeness of those intended to participate) who receive the intervention. Often the reach of intervention is just a first time reach and does not include attendance - which is maybe more of an intervention exposure measure and patient adherence to intervention measure.

Q: In the SxPc group, because implementation is not the focus, would it be better to use just the Pc to make it truly a control group for the intervention outcome? How robust can the intervention interpretation be given that Sx is intertwined with implementation? How can we determine if Sx and Sc are equivalent? How can we test it? For the SxPc arm of the study, how will you determine whether any lack of observed effectiveness is due to the intervention program versus due to the implementation strategy? For example, if Sx yields poor adoption or fidelity, couldn't that yield misleading evidence in the comparison of Pa versus Pc?

A: Thank you for these questions. I appreciate that these concepts are complex, and I have struggled with these specific concerns. I will summarize my current thinking below.

In my explanations I have addressed the simplest case of somewhat generic studies. Actual studies have a wide range of goals and investigators choose designs that best meet the goals of their specific study. I give guidance for a simple design that can be expanded and edited to the investigators' needs. So, I do not address all the "what if's". Also, at this very early stage of creating and presenting this new paradigm I cannot claim that this is the only way to do a rigorous DRCT. After thinking hard about the design for over a year, I believe this approach has strong utility.

My goals were to design a study that could use RCT's to both test the impact of the implementation strategies on implementation outcomes and test the effect of the intervention on health outcomes. Further, I assumed that SaPa denotes the package that would be disseminated if proven superior. It was not a goal to make the implementation exposure in the control arm of the intervention (Sx) equal to Sc or Sa. I instead made the assumption that the interpretation of the test of the intervention would include the caveats that, as it was delivered, the intervention was (or was not) successful at changing the health outcome more than the control in participants with certain characteristics, in a certain setting, in a given era and when administered in a specific way. The generalizability of the results will depend on numerous other factors in addition to the study design.

If the active intervention is evidence based, i.e. has been shown to work, it will come to the investigator packaged with implementation elements. Those elements need to be retained when the evidence based intervention is repeated. In our group we call those implementation strategies part of the evidence based intervention. I call the test of the intervention a test of the implemented interventions.

It is up to the investigators to figure out what control group (SxPc) will yield a comparison to the active intervention that will produce information that best promotes the public's health, public health/medical practice and disparities reduction. That comparison may or may not be one in which Sx and Sa produce equal levels of implementation of Pc and Pa respectively. However, investigators often consider that equivalence and provide ***at least*** *an attention control in both Sx and Pc such that the Hawthorne effect is avoided, i.e. the premise that any form of attention can change behavior (perhaps because of social desirability bias). If the investigator REALLY wants to know if the active intervention is better than the control, they should avoid a situation in which the implementation of Pc is so weak that comparisons to Pa are corrupted. It is in the implementation strand that weaknesses of the control implementation (Sc) are meant to be exposed.*

It is an option to use usual care or "nothing new" as the control intervention arm (SxPc). If an intervention impacting the health outcome is already being delivered in the relevant setting Sx and Pc could be the already present implementation strategies and intervention components. In any event, I call this usual care or nothing new implementation and intervention set SxPc rather than just Pc as suggested in the question above. SxPc can be no activity if the investigator thinks that the best choice, but they should consider if an attention control would yield a stronger interpretation.

I come back to the principle that the investigators should choose SxPc to yield the comparison to the active implementation and intervention arm (SaPa) that will produce the information and the interpretation that best promotes health, public health/medical practice and disparities reduction. Further, that choice will often be independent of the decision process of choosing Sc. The comparison of the interventions in a DRCT is a comparison of the conducted interventions. If disseminated and spread (which is, after all, the ultimate goal), the intervention package exported will retain both the implementation strategies (Sa) and the intervention components (Pa).