Conjuring power from a theory of change: The PWRD method for trials with anticipated variation in effects

Timothy Lycurgus, Ben B. Hansen, Mark White

Many efficacy trials are conducted only after careful vetting in national funding competitions. As part of these competitions, applications must justify the intervention's theory of change: how and why do the desired improvements in outcomes occur? In scenarios with repeated measurements on participants, some of the measurements may be more likely to manifest a treatment effect than others; the theory of change may provide guidance as to which of those observations are most likely to be affected by the treatment.
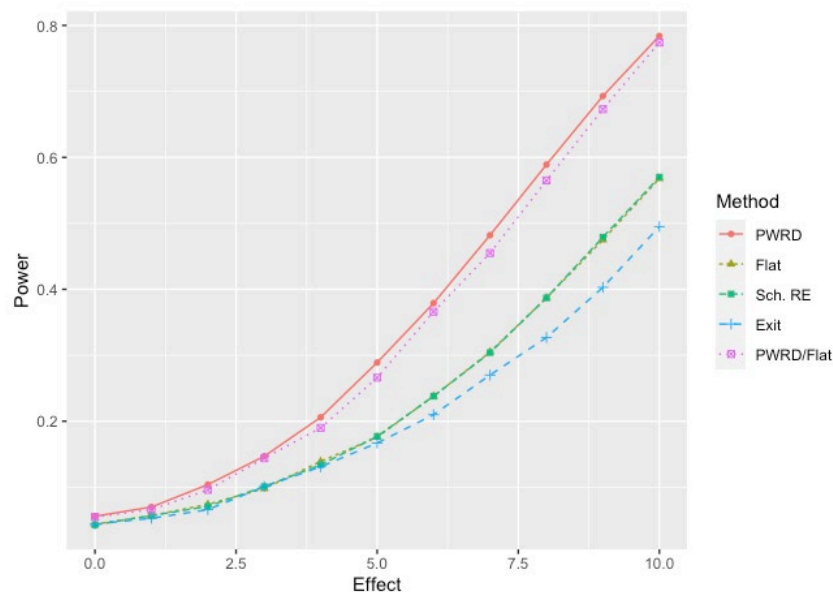

Figure 1: Power for the various methods across increasing effect sizes when the theory of change is correct.

In this paper, we present a method, power-maximizing weighting for repeated measurements with delayed-effects, i.e. PWRD aggregation, that converts the theory of change into a test statistic with improved asymptotic relative efficiency and thus, greater power to detect an effect when the theory of change holds.

PWRD Aggregation Weights

This paper is motivated by a large-scale early-elementary intervention that provides supplemental instruction to students who have fallen behind in learning to read. Notably, students only receive support once their performance has lagged so effects are delayed and scattered across the study population. Due to this, certain cohorts and occasions of follow-up have greater rates of exposure to the supplemental instruction. The theory of change provides guidance as to the expected exposure in each group; PWRD aggregation then uses the expected exposure to formulate weights to combine group-level effect estimates into an overall effect estimate in a manner that is power-optimal when the theory of change is correct.

How much power can be gained?

To demonstrate how PWRD aggregation works in comparison with commonly used alternatives, the paper uses a simulation study that mirrors the motivating intervention. When the theory of change holds, we find a substantial improvement in power (see the figure to the left). This remains the case both with and without spillover effects and across intraclass correlations typical in education research. When the theory of change proves incorrect, PWRD aggregation leads to a small loss of power, which was less than 3% under all effect sizes tested in the simulation. PWRD aggregation may also be used in tandem with standard estimation techniques with an appropriate multiplicity correction: when the theory of change holds, PWRD aggregation will provide an enhancement to power but when it does not, the standard technique will protect against a large loss in power.