

Using Digital Twins to Improve Statistical Efficiency in Alzheimer’s Disease Trials with p-tau217

Authors
 Anoop Praturu, Coco Kusiak

Background

Emerging biomarkers offer new sources of prognostic and diagnostic information in Alzheimer’s disease (AD). In particular, the blood-based biomarker p-tau217 is increasingly incorporated into AD clinical trials as an exploratory or candidate endpoint, reflecting its biological relevance and less invasive measurement compared with cerebrospinal fluid or PET-based biomarkers¹. In parallel, digital twin–based covariate adjustment methods have previously been shown to reduce outcome variability and improve statistical power for clinical endpoints². However, the performance of these methods in trials incorporating emerging AD biomarkers has not been well characterized, particularly when such biomarkers are measured in only a subset of participants or at sparse follow-up visits. In later-phase AD trials, biomarkers such as p-tau217 are typically used as secondary endpoints to support interpretation of clinical outcomes and biological mechanisms, where improving their statistical power is valuable but constrained by limited biomarker sampling.

Methods

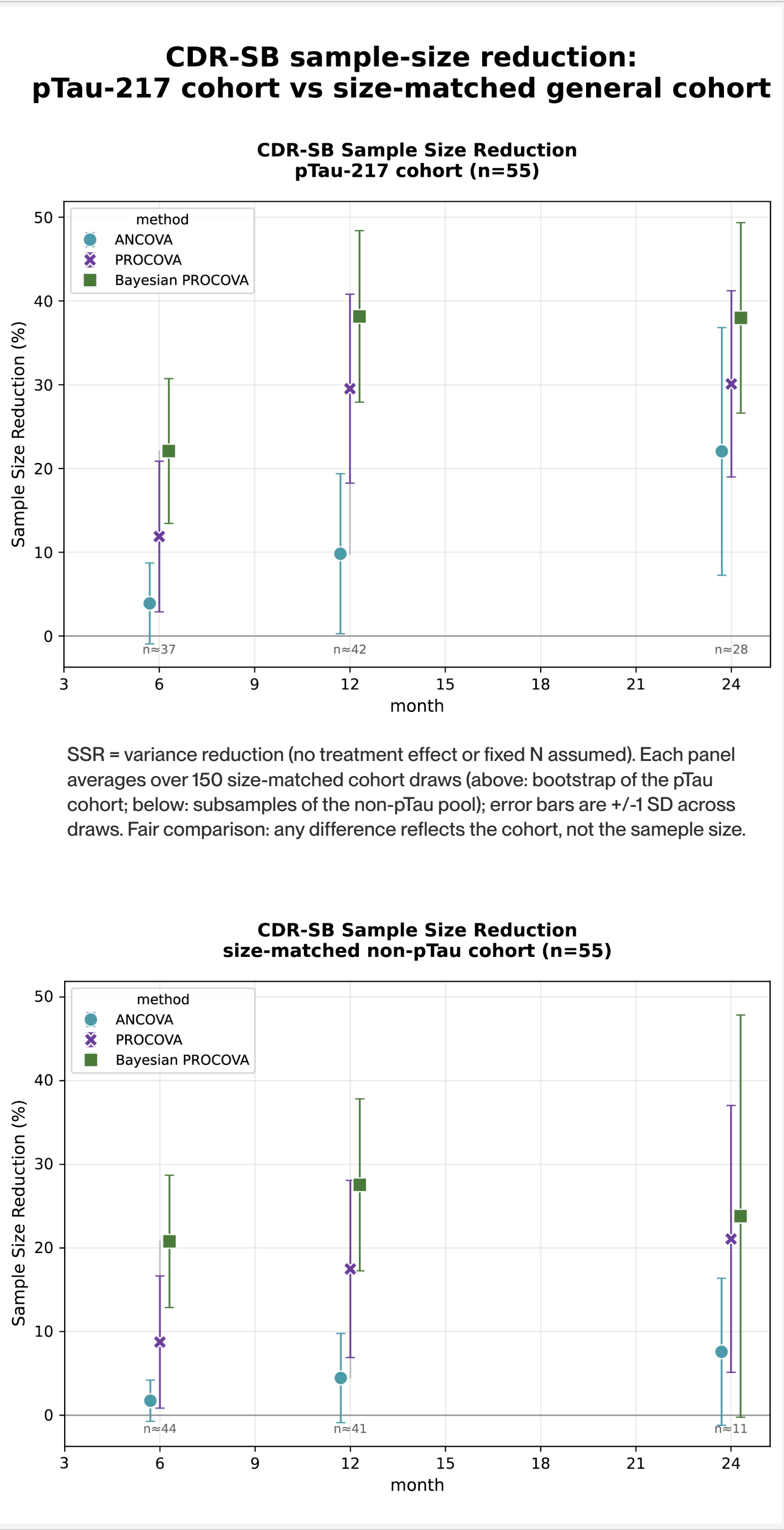
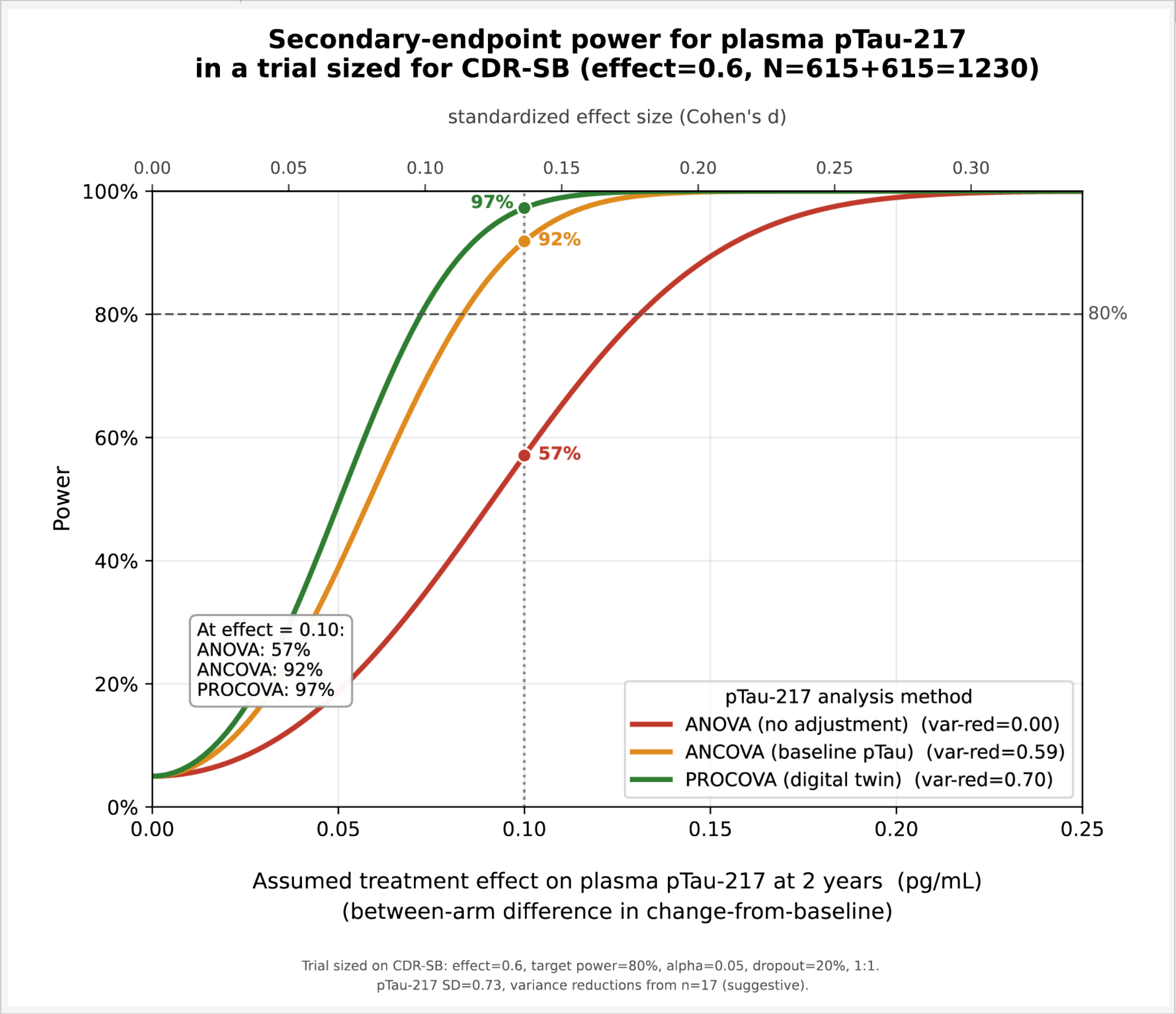
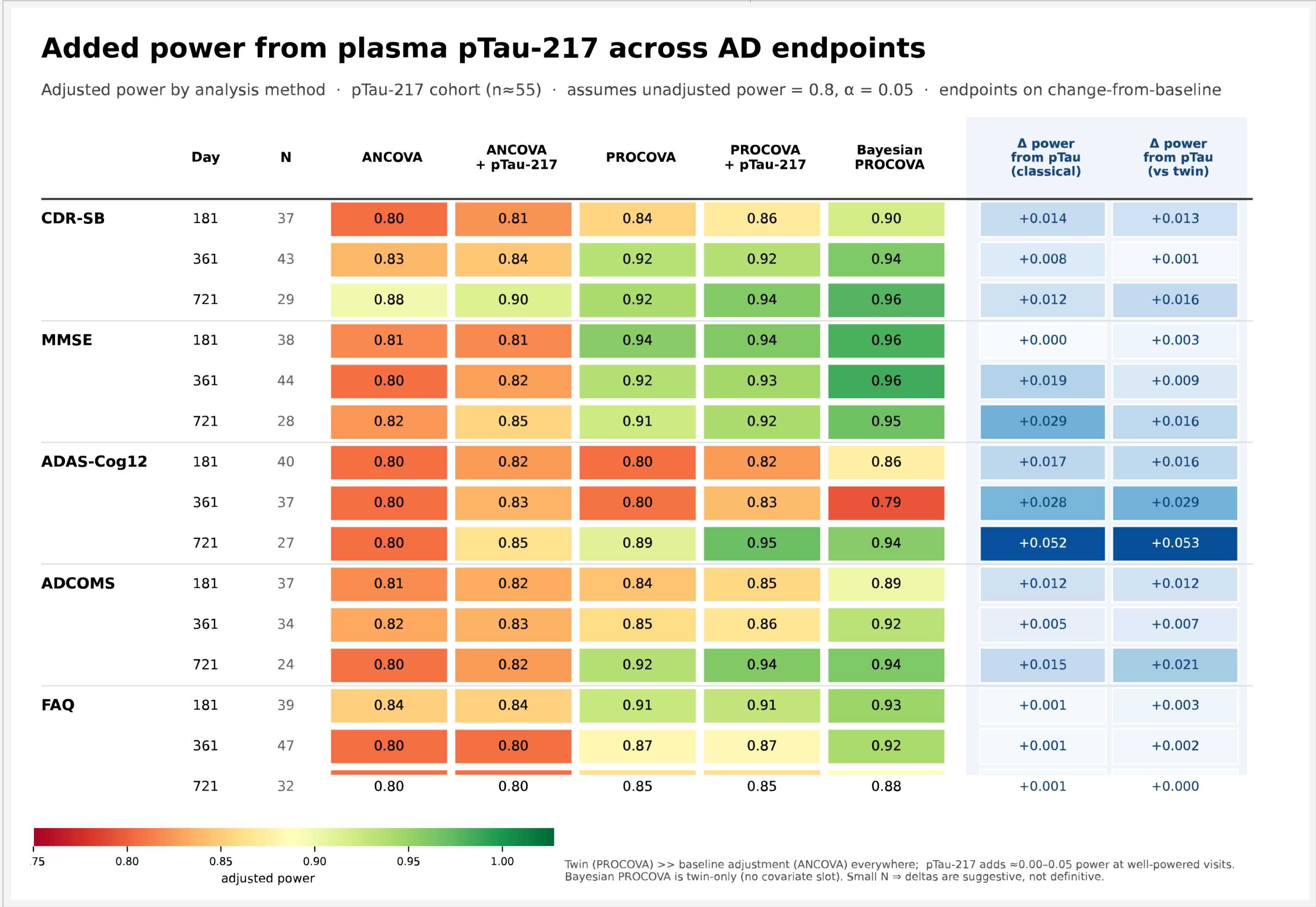
- We analyzed AD datasets containing longitudinal clinical outcomes alongside p-tau217 measurements.
- We considered a trial design powered on a clinical primary endpoint at two years—CDR-SB powered to 80% (yielding a fixed two-arm sample size of approximately 1,230 participants)—and evaluated the resulting statistical power for p-tau217 as a secondary endpoint at that fixed sample size under different analysis strategies: an unadjusted analysis (ANOVA), baseline covariate adjustment (ANCOVA), and digital twin–based adjustment (PROCOVA and Bayesian PROCOVA).
 - Our AD Digital Twin Generator is a machine learning model trained on historical data from prior clinical trials and observational studies to forecast expected placebo outcomes for trial participants [2]. Digital twin predictions were included as covariates in treatment effect estimation, and performance was assessed via variance reduction, the corresponding sample-size reduction, and power for p-tau217 as a secondary endpoint at fixed sample size.
 - To assess robustness in realistic biomarker-limited settings, we additionally compared CDR-SB sample-size reduction in the p-tau217–measured cohort against a size-matched cohort without p-tau217.

Results

- Across analyses incorporating p-tau217, digital twin–based covariate adjustment reduced outcome variability relative to standard adjustment methods.
- At the two-year timepoint, digital twin adjustment achieved a variance reduction relative to an unadjusted analysis of approximately 70% for p-tau217, compared with approximately 59% for baseline adjustment alone. This variance reduction translated into an estimated increase in statistical power for p-tau217 as a secondary endpoint, from 92% to 97%, at the sample size fixed by the CDR-SB primary endpoint.
 - For the CDR-SB primary endpoint itself, digital twins yielded estimated sample-size reductions of approximately 31% (PROCOVA) and 40% (Bayesian PROCOVA), versus approximately 21% for baseline adjustment, with the digital twin contributing the majority of the achievable reduction and p-tau217 adding modest further gains as a covariate.
 - These improvements were observed in settings with limited longitudinal biomarker support and were preserved in the p-tau217–measured cohort relative to a size-matched cohort without p-tau217, demonstrating that approaches incorporating AI-generated digital twins into analyses can effectively leverage emerging biomarker information as it becomes available.

Conclusion

As blood-based biomarkers such as p-tau217 become more prevalent in AD clinical trials, they present an opportunity to further improve trial efficiency. Digital twin–based covariate adjustment provides a practical framework for integrating emerging biomarkers into trial analyses—whether p-tau217 serves as a secondary endpoint or as an additional covariate—enabling improvements in statistical efficiency even when longitudinal biomarker data are sparse. This approach offers a scalable path toward more informative and efficient AD trial design as biomarker use continues to evolve.



References
 [1]
 [2]