Machine learning-based methodology optimizes composite endpoints for signal detection in clinical trials

Submitter Jonathan Walsh

Affiliation Unlearn.AI

SUBMISSION DETAILS

I agree to provide poster pdf for attendee download. Yes

Methodological Issue Being Addressed We propose a novel machine-learning based methodology to improve signal detection in the analysis of composite endpoints commonly used in clinical trials evaluating neurodegenerative diseases.

Introduction Composite endpoints are commonly used in randomized clinical trials evaluating investigational drugs in neurodegenerative diseases like Alzheimer's disease (AD), Amyotrophic Lateral Sclerosis (ALS), Huntington's disease (HD) and Parkinson's disease (PD). These composite endpoints, commonly used in clinical trials, are also generally required by FDA and other regulatory agencies as primary endpoints to assess approval for investigational drugs, e.g. Clinical Dementia Rating Scale Sum of Boxes (CDR-SB), Revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R), Unified Huntington's Disease Rating Scale (UHDRS), Movement Disorder Society - Unified Parkinson's Disease Rating Scale (MDS-UPDRS). However, these composite endpoints have well-recognized limitations that stem from the common practice of weighting all components equally, irrespective of their impact on the life of the patient. For example a global call to action has been recently taken in the ALS field to improve the ALSFRS-R composite endpoint (Genge et al., 2024).

Methods Methods from portfolio theory in finance (Arnold, 2002) may be used to optimize composite endpoints. The efficient frontier is a convex optimization that assigns weights to each item in a composite score. The approach can be used to maximize the amount of progression for a given target variance, and overall identify a set of optimal item weights for the composite scores that yields the maximum signal to noise. With mild assumptions, this point in turn maximizes the power for the composite score as an endpoint in a clinical trial.

In this work, we generated digital twins of patients using their baseline data from clinical trials to optimize the signal-to-noise ratio of composite endpoints. Digital twins are comprehensive longitudinal predictions of control outcomes that are generated from a pre-trained machine learning model. For common study populations and endpoints, the efficient frontiers were constructed using the predictions of the expected per-item progression and covariance structure between items. The optimal weights that maximize signal-to-noise for items were obtained for each cohort and endpoint, and also compared to direct optimization on observed outcomes. This approach can also determine which endpoints in particular benefit from optimization. We applied this method to CDR-SB, ALSFRS-R, UHDRS, and MDS-UPDRS, common primary endpoints used in AD, ALS, HD, and PD clinical trials, respectively.

Results The signal-to-noise ratio improved across endpoints, with the largest gains in HD and PD. The signal-to-noise ratio for UHDRS and MDS-UPDRS improved by as much as 50%, leading to significant gains in power. AD and ALS had smaller gains, approximately 10-15%, and depended on the particular cohort analyzed.

Conclusion Application of digital twins generated from a machine learning model can improve the signal detection in composite endpoints by reweighting each component according to its impact on the patient's outcomes and disease progression. This novel methodology is ideally suited for early-stage randomized clinical trials such as Phase 2 proof-of-concept studies.

Co-Authors

Jonathan Walsh[1], Run Zhuang[1], Coco Kusiak[1], Luis Olmos[1]

[1] Unlearn.AI

Keywords

Keywords
composite endpoints
clinical trial design
machine learning

Guidelines I have read and understand the Poster Guidelines

Disclosures All authors are equity-holding employees of Unlearn.AI.

Related Tables and Supporting Materials <blank>