# You Paid for the Information in the Study,
# Don’t Aggregate it Away in Analysis

*Two independent demonstrations that standard clinical scoring practices mathematically discard statistically significant, clinically interpretable information.*

---

**Auqxana** — White Paper, June 2026
Anderson M. Rodriguez, *CEO / Head of R&D*
anderson@auqxana.com

Every clinical trial that uses an ordinal outcome scale — the ALSFRS in ALS, the UPDRS in Parkinson's, the ADAS-Cog in Alzheimer's, the HAM-D in depression — collects detailed, item-level data across multiple functional domains. Every patient, at every visit, answers each question on a graded scale. This data is both slow and expensive to collect, but the calculus on collection is that it is worth it because this data encodes real structure about how disease progresses and how patients respond to treatment.

Published work by Auqxana's founder Anderson M. Rodriguez (2026) demonstrates conclusively: standard analytic pipeline then aggregates away key signal in this expensive data when items are summed into a total score.

This discarded information is not an abstract technical concern; sum scoring longitudinal data demonstrates a provable, quantifiable loss of clinically significant information.

Examples are numerous and include linear models fit to resulting trajectories, or drug treatment arms being compared by slope. The distributional structure within each domain — which items are declining, at what rate, *with what timing relative to each other* — is gone before the analysis begins.

This white paper presents two independent demonstrations, in two unrelated clinical contexts, that the information discarded by standard scoring practices contains signal that matters.

---

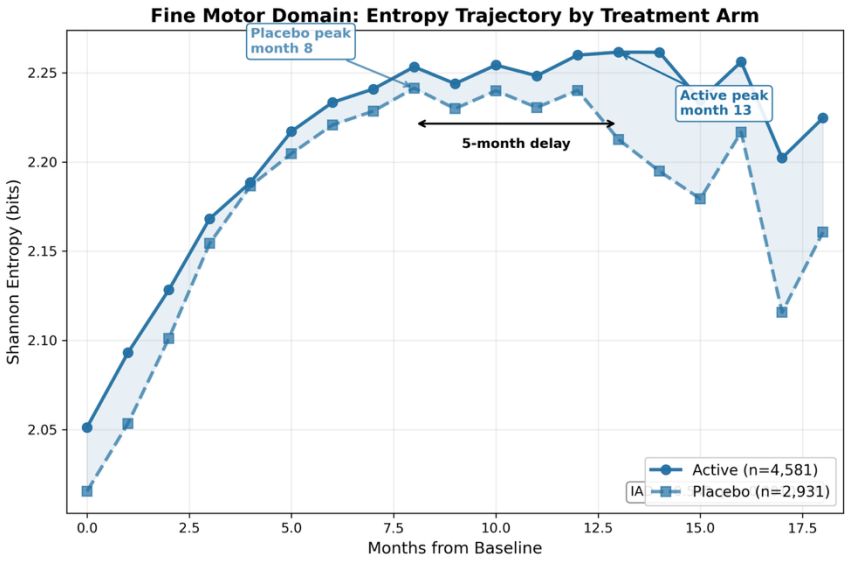
## Case 1: ALS clinical trials — a five-month signal undetectable by conventional endpoint analysis

The PRO-ACT database contains item-level ALSFRS data from over 8,800 ALS patients across multiple completed Phase II/III trials, stratified into Active and Placebo arms. The standard endpoint for these trials compares the linear slope of total ALSFRS scores between arms.

When Shannon entropy is computed instead from the item-level score distributions within each of the four ALSFRS functional domains — Bulbar, Fine Motor, Gross Motor, and Respiratory — the Active and Placebo populations diverge visibly and systematically. In the Fine Motor domain, the placebo population reaches maximum distributional spread at month 8 and reverses. The drug-treated population does not reach the same point until month 13. The active arm delays this transition by five months.

A permutation test shuffling patient-level arm labels confirms the total divergence across all four domains at $p < 0.001$ — 7.5 standard deviations above the null mean. Not one of 1,000 random shuffles produced divergence as large as what the data contain.

This signal does not exist as a measurable quantity in the total-score linear-slope framework. A five-month shift in the timing of distributional spread is not a slope difference. The standard endpoint cannot represent it, and therefore cannot detect it. The compression step eliminates the structure before the comparison is ever made.




This finding does not prove that any specific drug worked. The PRO-ACT database anonymizes compounds, and the pooling of heterogeneous trials introduces compositional confounds. What it does establish, empirically, is that the information the standard endpoint discards is not empty. The distributional structure assumed to be uninformative contains a substantial between-arm signal.

*Rodriguez, A.M. (2026). Shannon Entropy Trajectories Reveal Between-Arm Distributional Structure Invisible to Standard Endpoint Analysis in Pooled ALS Clinical Trials.* medRxiv. 7.5σ permutation result; press coverage. Based on the method developed in *Rodriguez, A.M. (2026). An Entropy-initiated Coupled-trait ODE Framework for Modeling Longitudinal Cohort Dynamics.* PLOS One.

Analysis code: github.com/amr28693/als_arm_entropy ;
https://github.com/amr28693/ECTO_walkthrough_2026

---

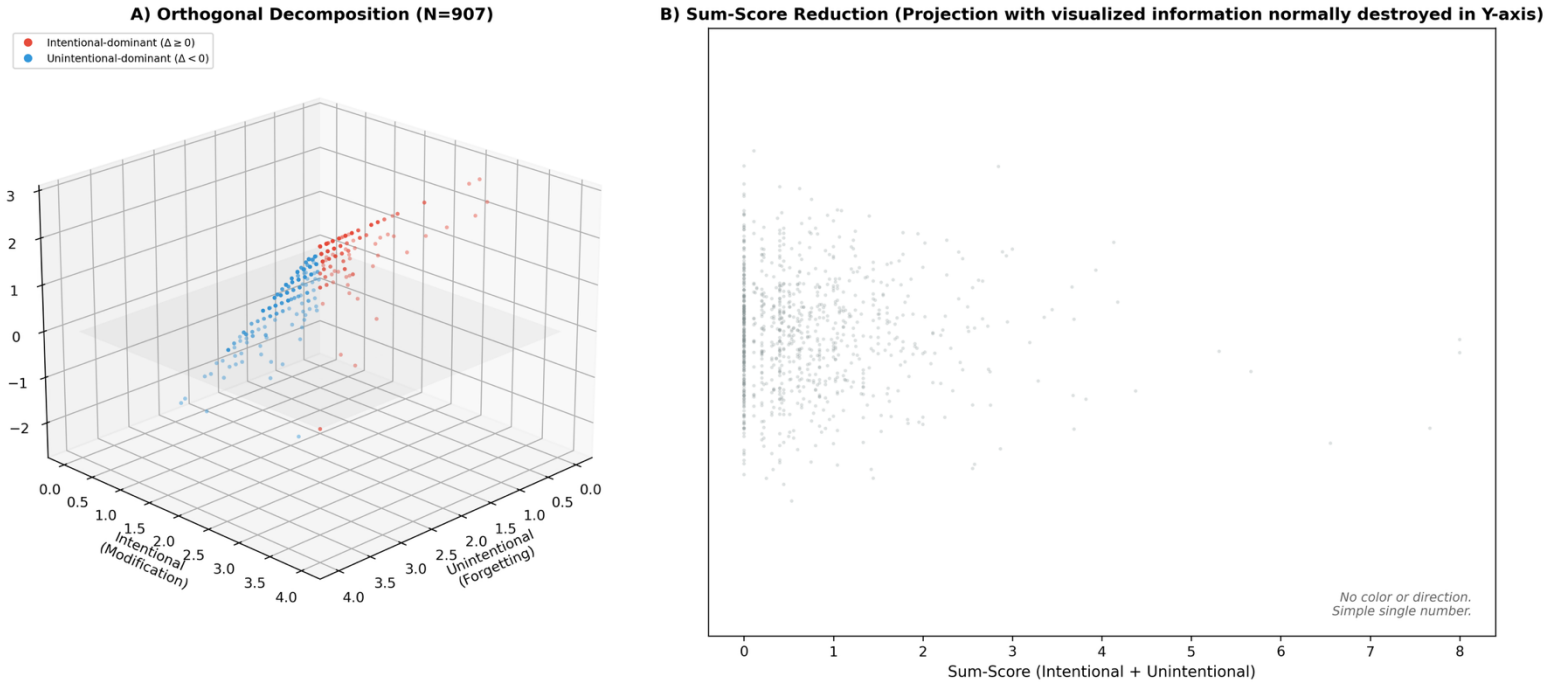
## Case 2: Medication adherence — sum score erasure

The most widely used medication adherence instruments produce a single sum score that treats intentional dose modification and unintentional forgetting as interchangeable contributions to the same number. A patient who deliberately stops a medication because of side effects and a patient who simply forgets to take it receive the same score. They need fundamentally different clinical responses.

In a secondary analysis of 907 adults with neurological disorders from the NeuroGerAd study, the intentional and unintentional adherence sub-factors of the SAMS instrument were decomposed into orthogonal components: a sum (total nonadherence, analogous to the standard score) and a difference (the Differential Adherence Index, Δ = intentional − unintentional, capturing nonadherence directionality).

The sum and the difference share 0.6% of their variance. They are effectively independent measurements of the same patient. One measures how much while the other measures why.

Diagnosis groups differed significantly on the difference ($F = 2.63$, $p = .03$) but not on the sum ($F = 1.48$, $p = .21$). Movement disorder patients, whose cognitive decline impairs the ability to remember medication, showed a forgetting-dominant profile. Neuromuscular patients, whose cognition is preserved, showed a balanced profile consistent with deliberate modification. This distinction — which determines whether the patient needs a pill organizer or a conversation with their physician — is invisible to any instrument that reports only a sum score.




The mathematical operation that recovers it is a 45-degree rotation of the existing measurement space. It costs nothing and requires no new data collection. It has been available, latent in the existing instruments, since they were first validated.

*Rodriguez, A.M. (2026). Differential Adherence: Orthogonal Decomposition of Medication Nonadherence Reveals Clinically Significant Heterogeneity Destroyed by Sum-Score Instruments.* Pre-printed.
Analysis code: github.com/amr28693/orthogonal_differential_adherence

---

## The question for your data

Both cases share the same underlying signal loss: premature mathematical compression of ordinal data into a single scalar, performed before the analysis, aggregates away structure that turns out to matter.

This is not unique to ALS or to medication adherence. Any clinical instrument that sums ordinal items across functionally distinct domains is performing the same operation. The UPDRS sums across motor, non-motor, and activities of daily living. The ADAS-Cog sums across memory, language, and praxis. The EDSS compresses eight functional systems into one disability score. In every case, the question is the same: does the distributional structure discarded by the summation contain signal relevant to treatment, prognosis, or patient stratification?

The analytic tools to answer that question exist now. They are computationally-lite, model-independent, easily interpretable, and applicable to any dataset where item-level ordinal scores are retained — which is nearly all of them.

---

## Auqxana — Clinical Data Analysis

Auqxana provides information-preserving analysis of clinical trial and outcome study data. Our methods detect treatment-arm divergence, subgroup heterogeneity, and domain-specific dynamics in ordinal outcome scales — signal that standard total-score analyses discard by design.

**For inquiries:**
Anderson Rodriguez, CEO / Head of R&D
anderson@auqxana.com
+1 (832) 928-0762

**Publications:** Rodriguez (2026), PLOS ONE (published); Rodriguez (2026), medRxiv (ALS entropy, 7.5σ result, press coverage).