Bridging AI Innovation and Clinical Trial Data in Psychiatry:

Mathematical Limits and Clinical Consequences

Dr. Joseph Geraci, PhD

Disclosures

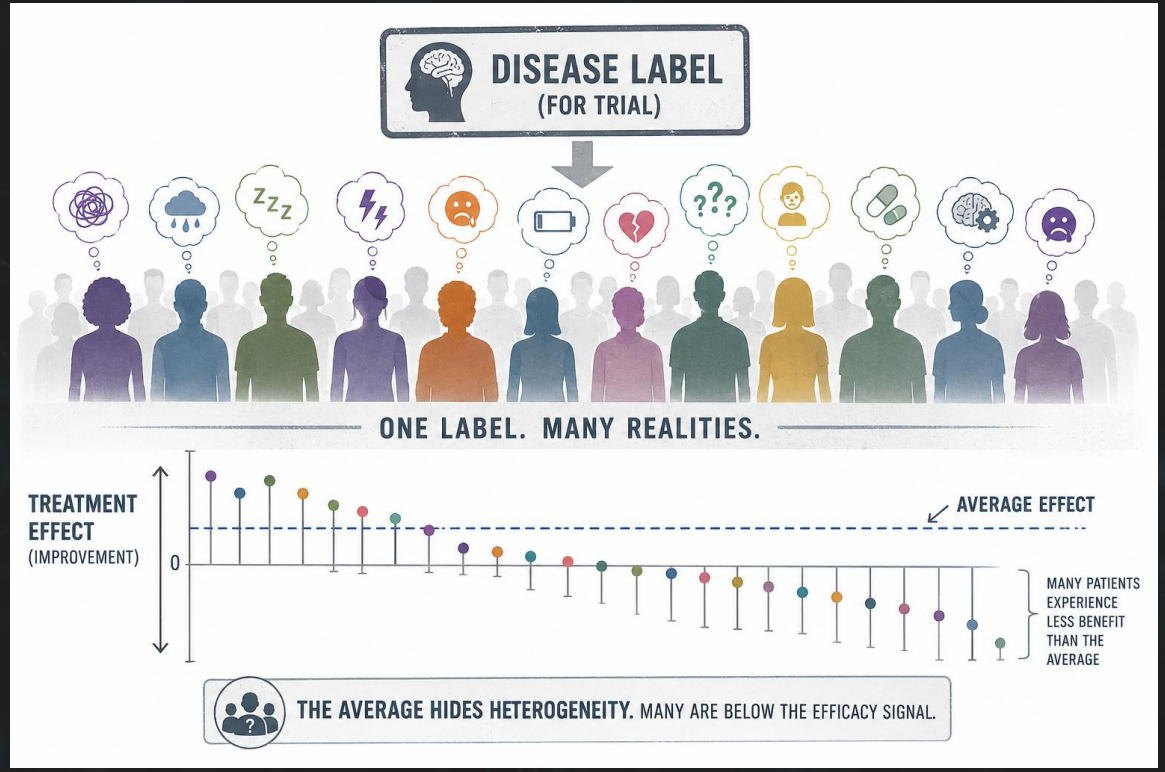
Joseph Geraci is the Founder, Chief Technical & Scientific Officer, and Board Member at NetraMark , a publicly traded corporation.

I am affiliated with CAMH, Queen’s university, and The University of California, San Diego.

Clinical Reality of AI for Psychiatric Trials

One of the hardest areas for AI in clinical trials:

- The diagnostic label does not define a single disease biology or a single treatment-response pattern.
- The average patient is often a mathematical convenience, not a clinical reality because clinical trials are complex systems



Core thesis:

AI for psychiatric trials must be designed to discover trial-specific patient structure.

It cannot only summarize prior trials, literature, or broad medical knowledge.

Heterogeneity is the First AI Challenge

Psychiatric diagnoses contain multiple hidden patient types

Patients can meet multiple trial criteria, yet profoundly differ in:

- Symptom Structure
- Comorbidity
- Medication History
- Cognition
- Placebo Susceptibility
- Treatment Response

The AI Problem:

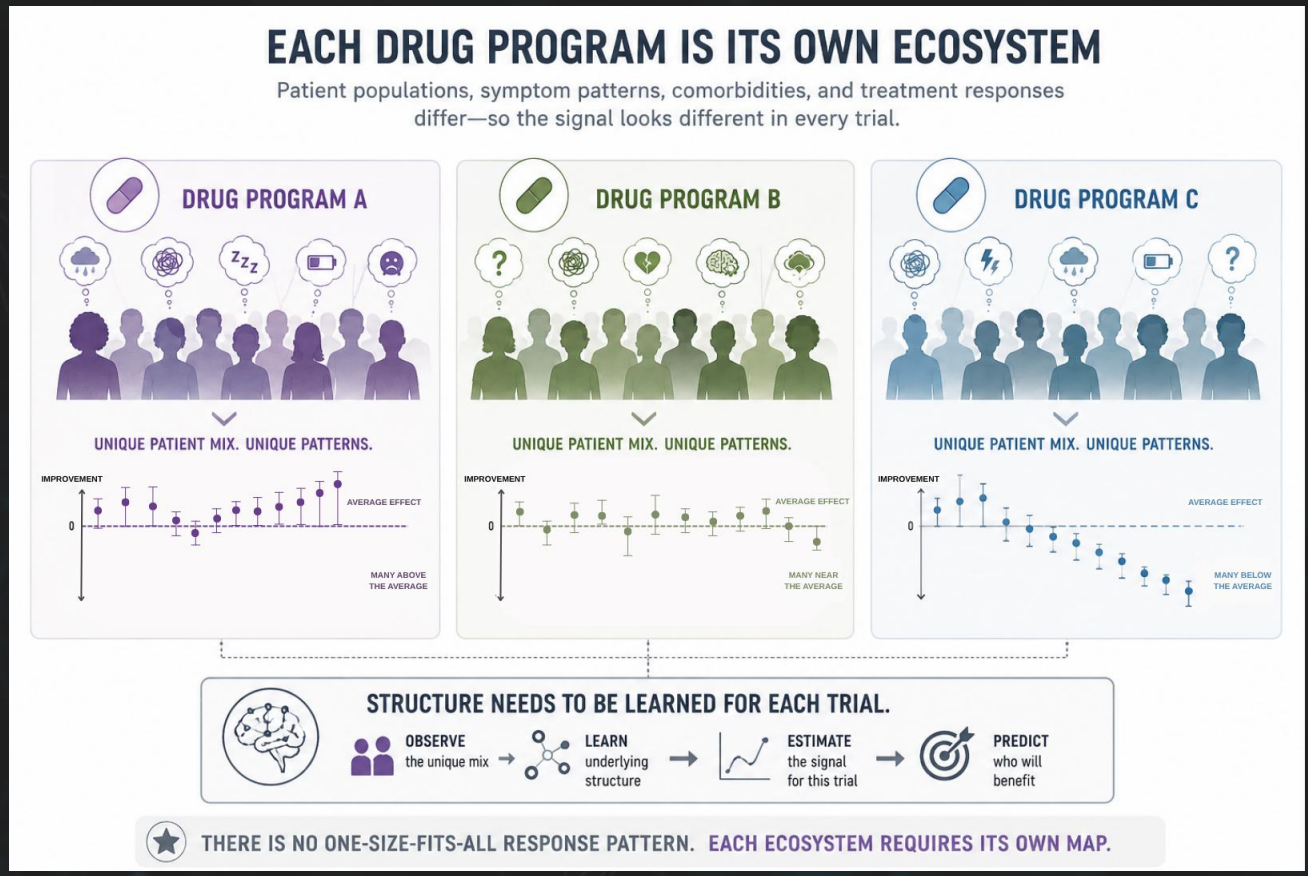
Not simply predicting response in the average patient.
The AI problem is *discovering which hidden patient subgroups are driving response*.

Every Drug Program is its own Complex System

Not just another spreadsheet → A newly created clinical ecosystem

Trial Internal Structure shaped by:

- Inclusion Criteria
- Site Effects
- Endpoint Behavior
- Concomitant Medications
- Placebo Dynamics
- Patient Composition
- Baseline Severity



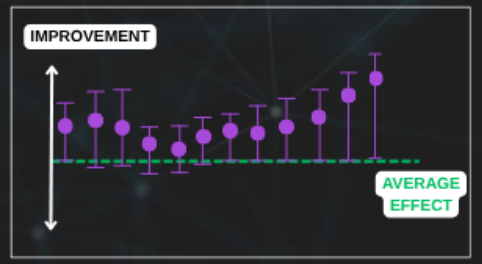
Prior training can provide context, but it cannot determine the patient geometry of a new trial.
That structure must be discovered from the trial itself.

The Mathematical Reason Signal can be Hidden

High-dimensional data can make the average patient look real, but in psychiatry, the average patient may be exactly the patient *who does not exist*.

The trial becomes high-dimensional as we collect more:

- ➢ Symptoms
- ➢ Biomarkers
- ➢ Scales
- ➢ Imaging Features
- ➢ Labs
- ➢ Clinical History
- ➢ Digital Measures



Global summaries can become *deceptively stable*.

Clinical consequence:

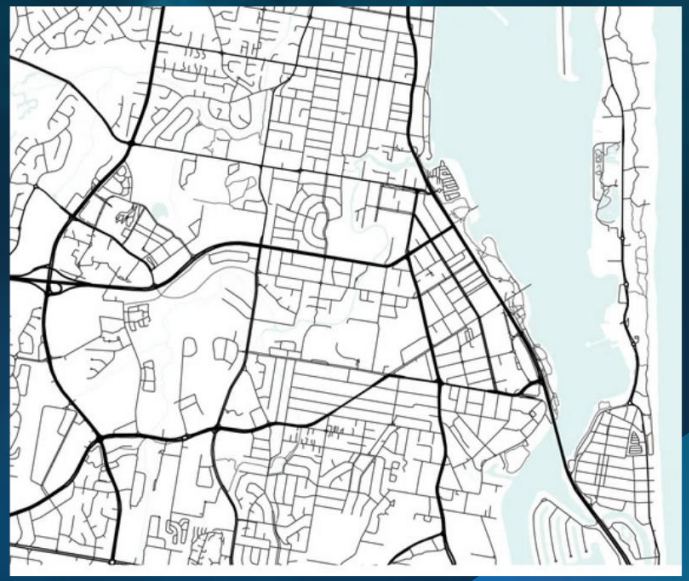
A trial can look negative overall while still containing a reproducible treatment-relevant subgroup.

Real Discovery Example: MDD Trial Relmada 1017-202

Would you use the map of a similar city to help you navigate while traveling?

Miami, Florida, USA

Southport, Gold Coast, Australia



MDD Trial

Not a post-hoc story.

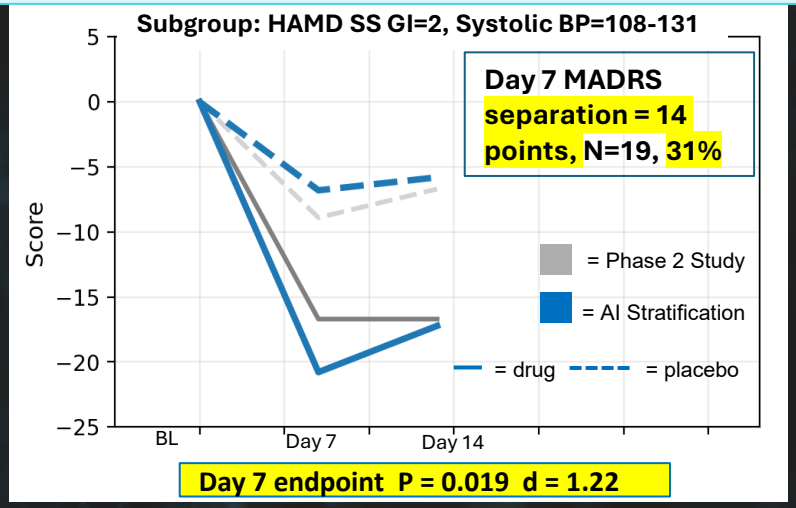
Psychiatric AI needs trial-specific subgroup discovery, stability testing, and replication, not only whole-cohort prediction or literature-based reasoning.

Signal Discovery & Validation

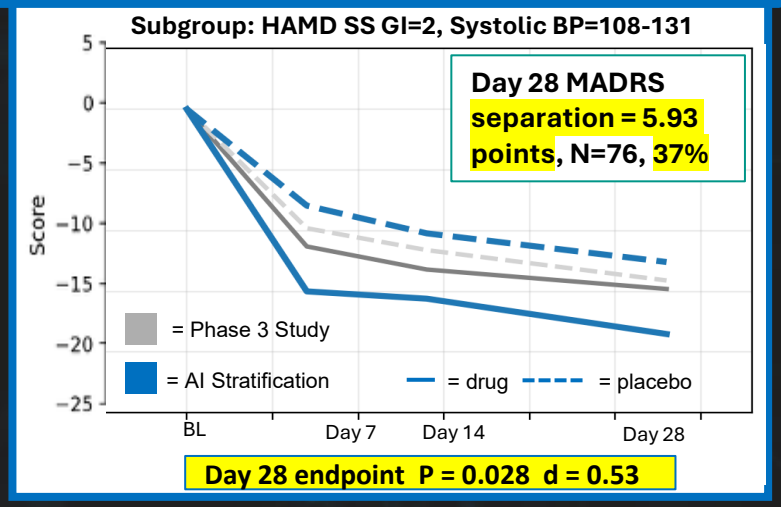
If the signal lives in a subgroup, then the average treatment effect can mislead us.

- Whole-cohort analysis did not fully reveal treatment-relevant structure.
- A signal emerged in a defined patient population when the data was mathematically interrogated for hidden subgroups.
- The finding replicated.

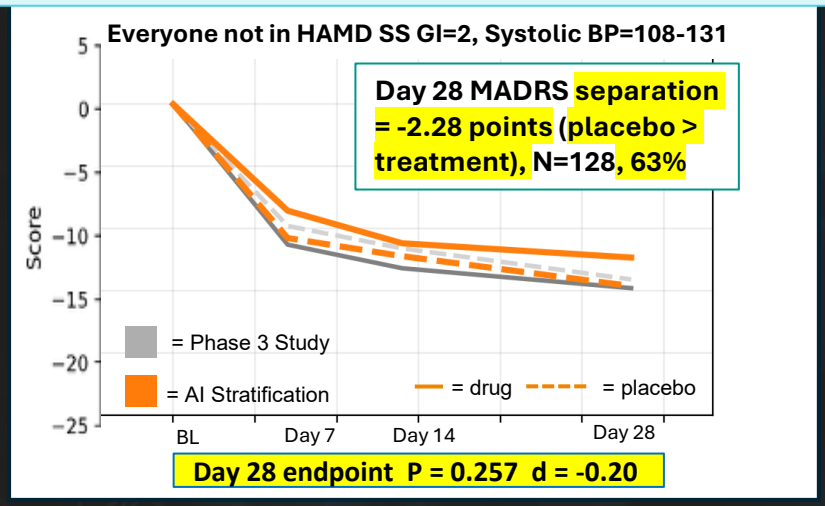
Phase 2 Study, Subgroup Identified



Phase 3 Study, Subgroup Replicated



Phase 3 Study, Patients Not in Subgroup



The Central Issue:

The treatment signal was not absent or in the average patient…it was localized in the right patient subgroup.

Clinical Trials Are Complex Systems: So What?

- If clinical trials are complex systems, then **prediction** based on the "average" prior knowledge is a category error.
- In a regime where the mathematical geometry is constrained by high dimensionality: The "average patient" becomes a mathematical phantom.
- Stability is found not in global summaries, but in **discovered local structures**.

Think of all the factors that put pressure on a patient's response

- Large Pretrained Models are great but they learn about reliable averages across large data from the past.
- Each clinical trial is a microcosm of out of distribution events that the large models have not seen.

LLMs are not enough. We can't just rely on huge models to handle the hard problem of small data. A trial is a new, highly sensitive source of data from human beings in a completely new clinical program, and it demands specialized methods that work alongside large models to uncover its real geometry.

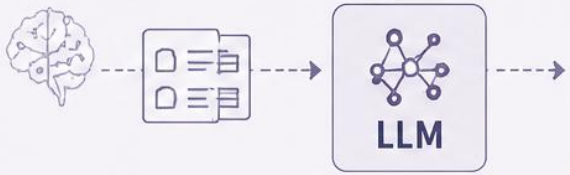
1. PSYCHIATRIC TRIAL POPULATIONS ARE INHERENTLY HETEROGENEOUS

One label. Many realities.

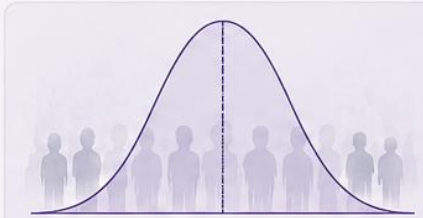
DIVERSE BIOLOGIES, SYMPTOMS, HISTORIES, ENVIRONMENTS, AND COMORBIDITIES.

2. LLMs ARE LIMITED IN WHAT THEY CAN FIND

They summarize what's known, not what's true for each patient.



- Finds common patterns
- Identifies known associations
- Produces average insights



RESULT:
AVERAGE-BASED INSIGHTS THAT OBSCURE WHO TRULY BENEFITS—AND WHY.
RISK OF MISLEADING CONCLUSIONS.

3. AI FOR PSYCHIATRIC TRIALS NEEDS TO BE DONE RESPONSIBLY

- Designed for psychiatric complexity and heterogeneity
- Goes beyond known patterns to discover hidden structure
- Finds the patients driving true efficacy signals
- Built with clinical rigor, interpretability, and transparency
- Developed in close partnership with clinical experts

BUILT TO REVEAL WHAT MATTERS IN PSYCHIATRIC TRIALS.

HETEROGENEITY IS THE RULE. AVERAGES HIDE THE TRUTH.
RESPONSIBLE, PURPOSE-BUILT AI REVEALS IT.