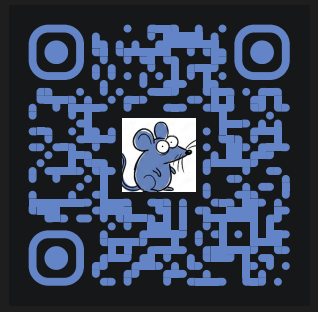
SCI Annual Meeting, 2026
Symposium Session

CATE Estimation in Complex Longitudinal Data

An Introduction

Ashley I Naimi, PhD | Dept of Epidemiology
Professor | Emory University
ashley.naimi@emory.edu
https://ainaimi.github.io/



EMORY | ROLLINS SCHOOL OF PUBLIC HEALTH

Acknowledgements

- This (in progress) work benefitted from discussion with several colleagues:
  - Jennifer Kawwass, Austin Schirmer, and Heather Hipp at Emory Reproductive Center
  - Audrey Gaskins at Emory Epidemiology
  - Edward Kennedy at CMU

- Please feel free to report errors and/or send comments to me: ashley.naimi@emory.edu

Overview

An Applied Problem

A Staring Point with G-Estimation of SNMMs

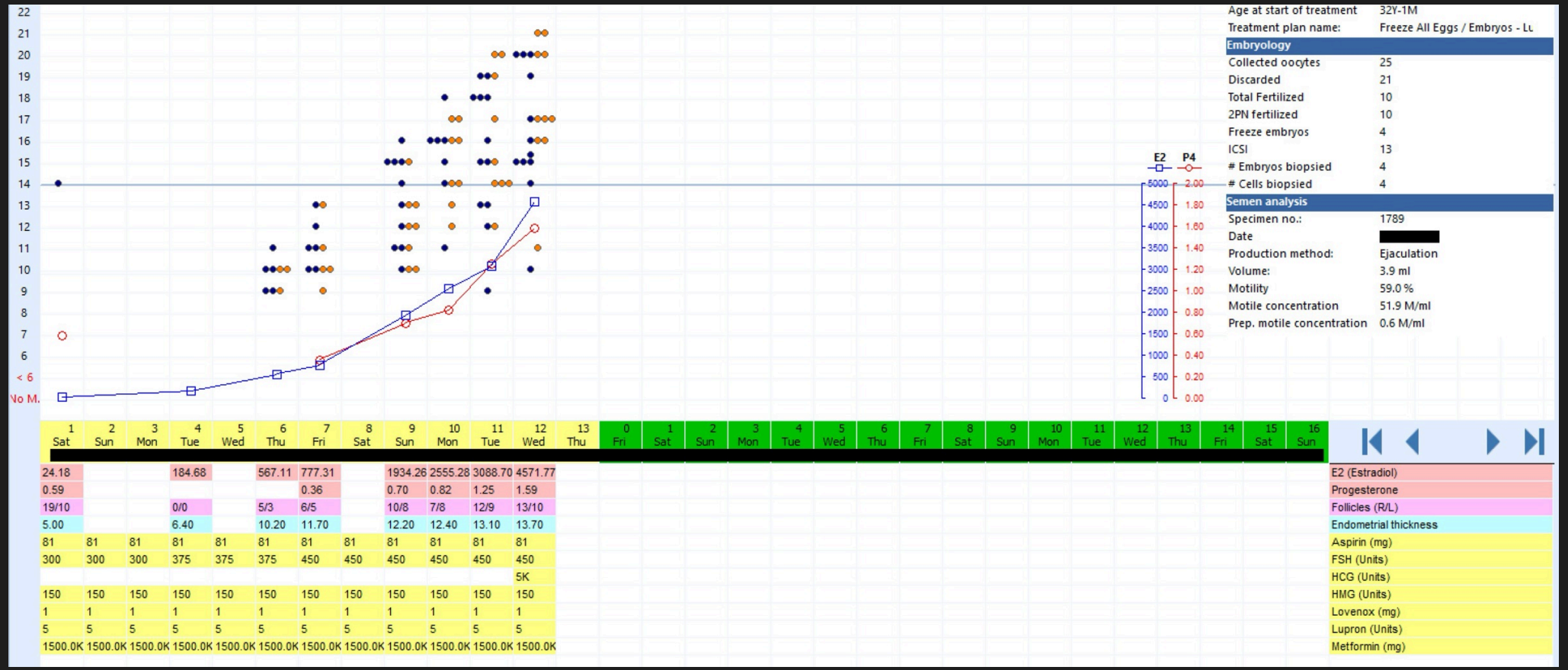
Three Approaches

Some Questions

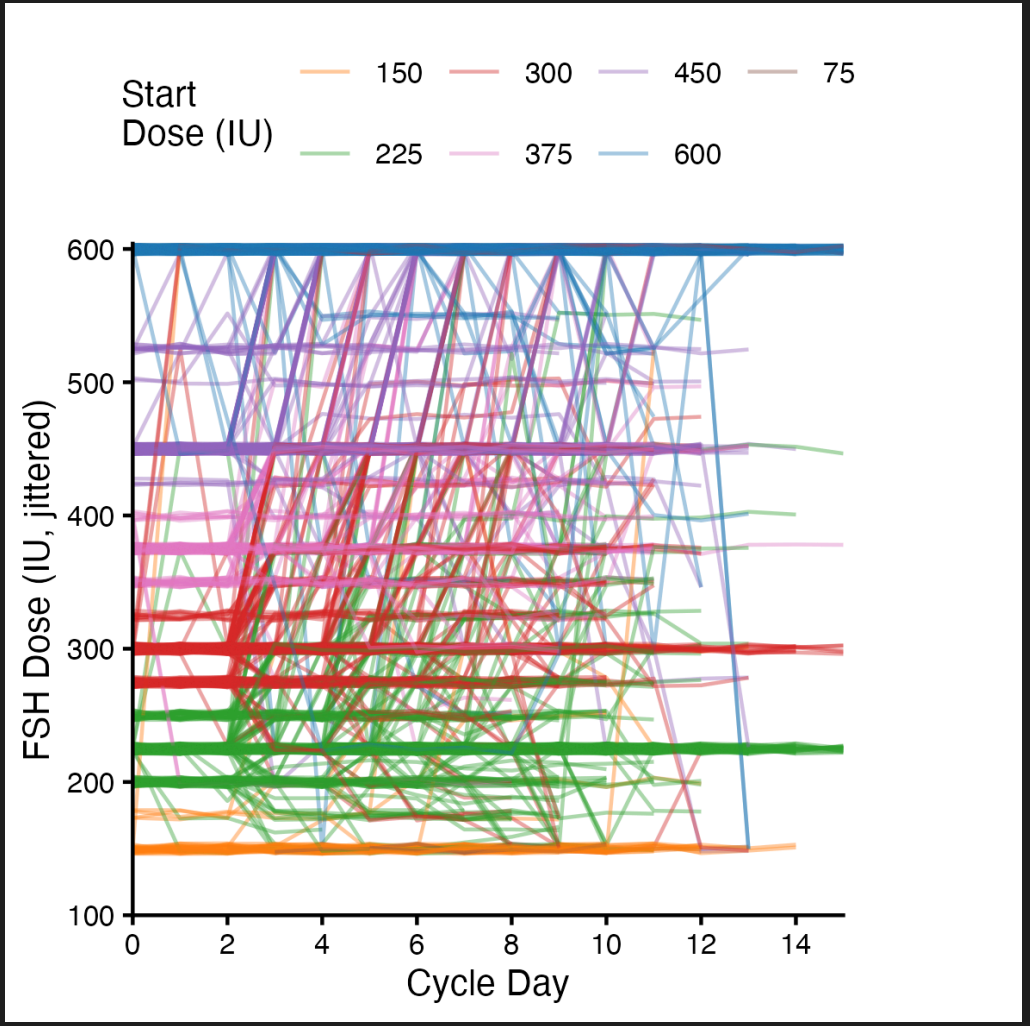
IVF: A Repeated, High-Stakes Treatment Decision

- Infertility affects roughly 15% of couples worldwide
- IVF: Daily gonadotropin injections to stimulate follicle growth + Trigger Shot
- Clinicians make dose decisions roughly every **2 days** based on:
  - Hormone Level Measures (e.g., Estradiol, Progesterone)
  - Follicle count and size (right + left ovary)
  - Baseline and Prior Daily Doses
  - Other
- But, in the US, fewer than 40% of IVF cycles result in live birth.
- Failure is expensive for the patient, at $15K to $30K per cycle

A Patient Example



Treatment Variation



What Do We Want to Know?

- Person-Day Specific Effects:

> For a specific patient with a dose and covariate history H through day k, what would happen to their outcome(s) if we changed the dose on day k alone?

- Person Specific Optimal Regimes:

> For a specific patient with a specific dose and covariate history H through day k, what changes to their dose on day k would optimize the outcome(s) of interest?

- Clinic Specific Policies:

> For {new patients, returning patients with unexpected prior failures} can/should there be clinic specific time-varying dose protocols that optimize outcomes?

How Can We Answer These Questions?

Robins (1994) introduced the **blip function** encoded in a SNM, which formalizes the person-day CATE

For example, if we let:

- $Y^{\bar{a}_k, \underline{a}^*_{k+1}}$ represent the end-of-cycle oocyte retrieval count under
  - a intervention history up to day k, denoted $\bar{a}_k$
  - a referent dose future from day $k+1$, denoted $\underline{a}^*_{k+1}$

$$\gamma_k(h_k; \psi) = E\Big[Y^{\bar{a}_k, \underline{a}^*_{k+1}} - Y^{\bar{a}_{k-1}, \underline{a}^*_k} \;\Big|\; H_k = h_k \Big]$$

where $h_k = (\bar{Z}_k = \bar{z}_k, \bar{A}_{k-1} = \bar{a}_{k-1})$ is the full patient history through day k.

For example, on the third ($k = 2$) day of COS, we can specify a linear SNMM for a "blip" increment as:

$$\gamma_2(h_2; \psi) = (\psi_1 + \psi_2 z_0 + \psi_3 z_1 + \psi_4 z_2 + \psi_5 a_1 + \psi_6 a_0)(a_2 - a_2^*)$$

(but...)

G-Estimation

Robins also introduced g estimation, a moment based estimating function for ψ.

Find $\hat{\psi}$ such that the "blip-free" outcome

$$Y^{\text{free}}(\psi) = Y - \sum_k \gamma_k(H_k; \psi)$$

is mean independent of treatment residual $(A_k - E[A_k \mid H_k])$, given history.

At each k, solve:

$$\sum_i \left[d^*(A_{ik}, H_{ik}) - \hat{E}\{d^*(A_k, H_{ik}) \mid H_{ik}\}\right]\left[Y_i^{\text{free}}(\psi) - \hat{E}\{Y_i^{\text{free}}(\psi) \mid H_{ik}\}\right] = \mathbf{0}$$

where $d^*(A_{ik}, H_{ik})$ is a p-dimensional function of current treatment and history.

Problem 1: It's Complicated

For decades, authors (e.g., Stijn Vansteelandt) have been advocating for more use in applied settings, but the concepts are genuinely demanding:

- SNMs and blip functions are very unfamiliar objects
- The construction of Y^{free}, and the choice of d^* require understanding before implementation
- Estimating equations are not standard for many applied researchers

Software situation remains poor:

- `gesttools` in R: no longer on CRAN and appears non-operational
- To my knowledge, no well-maintained, general-purpose package exists in R or Python
- Practitioners who want to use these methods largely have to code from scratch

Problem 2: Fully Conditional Structural Model

G-estimation requires **correct specification of the SNMM**:

- Must correctly model covariates that modify the effect
- The "double robustness" property is not like AIPW or TMLE of the ATE.
- There is some interesting work on interpreting ψ in misspecified SMMs, but not clear:
  - How this generalizes to time-varying settings
  - How to convey these "projection" parameter interpretations to clinical colleagues

What We'll Discuss Today

DR Learner (Kennedy): a broadly applicable "Swiss army knife"

- Doubly-robust pseudo-outcome regression, extended to the longitudinal setting
- No structural model required; any ML method in the second stage

Extension of the R-Learner (Syrgkanis): closely related to g-estimation

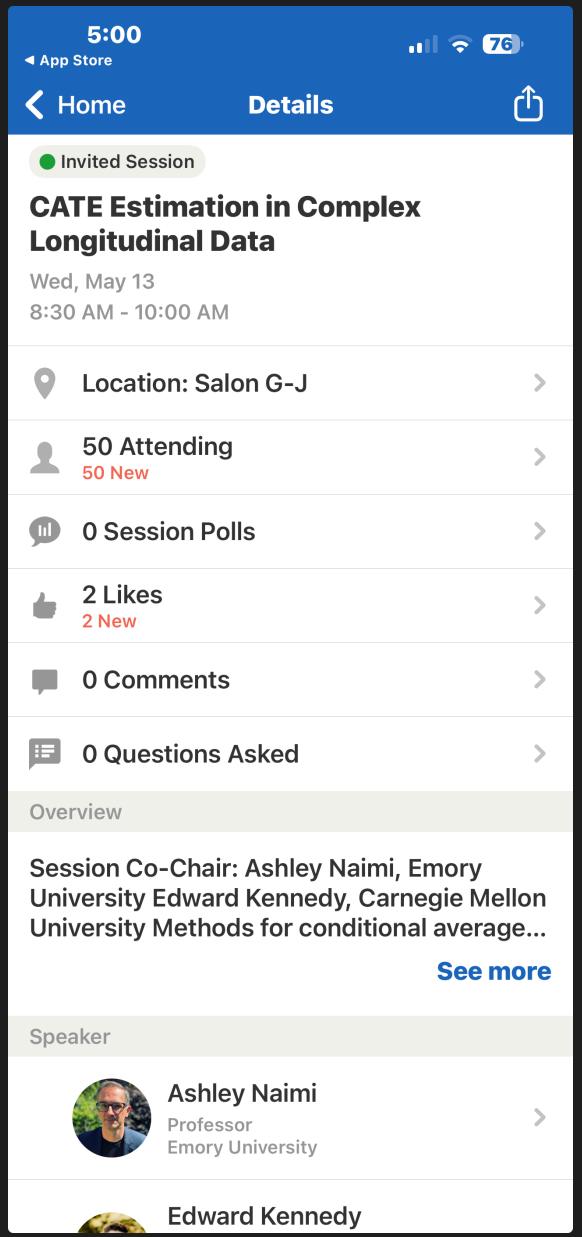
- Retains the Neyman-orthogonal estimating equation logic of SNMMs
- Incorporates ML flexibility directly into the structural model for the blip

Causal Excursion Effects (Chen): in micro-randomized trials

- Known randomization probabilities eliminate time-varying confounding
- Weighted estimating equation approach yields direct inference on day-specific CATEs

Panel Discussion (Scharfstein): tying it all together

- Lead the panel discussion on theoretical and practical challenges


5:00
App Store
76
Home
Details
Invited Session
CATE Estimation in Complex Longitudinal Data
Wed, May 13
8:30 AM - 10:00 AM
Location: Salon G-J
50 Attending
50 New
0 Session Polls
2 Likes
2 New
0 Comments
0 Questions Asked
Overview
Session Co-Chair: Ashley Naimi, Emory University Edward Kennedy, Carnegie Mellon University Methods for conditional average...
See more
Speaker
Ashley Naimi
Professor
Emory University
Edward Kennedy

Some Questions

How Hard Is This to Implement?

DR Learner: very accessible

- Two-stage regression; second stage uses any ML or parametric model
- We often implement in time-fixed settings with the Cross-Validated Super Learner
- Same ease in longitudinal extensions?

DML with G-Estimation: moderate complexity

- Very closely related to g Estimation proper
- Can it be implemented via a residual-on-residual type approach under a nonparametric structural model?

CEE / MRT methods: designed for accessibility

- Weighted GEE-type estimating equations
- R package `MRTAnalysis` available
- How do these translate to observational data?

Robustness: What Can Go Wrong?

"Double robustness" can mean different things in different settings. How "robust" is each approach?

What happens if:

- One or Both models misspecified?
- Near positivity violations?
- How "robust" is variance estimation
- Can we even estimate the variance of the CATE functional?

But also:

How amenable are these to:

- measurement error correction methods?
- sensitivity / bias analyses for unmeasured confounding?
- Nonparametric bounding of LCATE effects?

Machine Learning for Nuisance Functions?

When fitting nuisance functions, can ML/nonparametric methods be used?

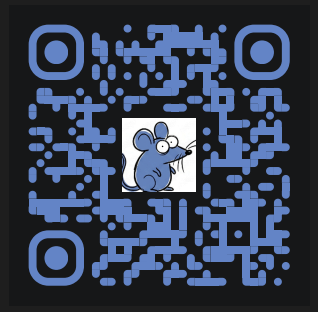
- What happens if we use very slowly convergent models (random forests?)
- Does longitudinal data affect how we might do cross-fitting?
- Does cross-fitting matter more, less, or the same for longitudinal CATEs?

SCI Annual Meeting, 2026
Symposium Session

CATE Estimation in Complex Longitudinal Data

An Introduction

Ashley I Naimi, PhD | Dept of Epidemiology
Professor | Emory University
ashley.naimi@emory.edu
https://ainaimi.github.io/



EMORY | ROLLINS SCHOOL OF PUBLIC HEALTH