SAVANNAH CONVENTION CENTER // SAVANNAH, GEORGIA, USA

SIO 2026

Annual Scientific Meeting

4-8 FEBRUARY

Statistical Considerations for IO-Specific Trial Designs

Nicholas J. Seewald, PhD
Assistant Professor of Biostatistics
Department of Biostatistics, Epidemiology, and Informatics
University of Pennsylvania Perelman School of Medicine

2026 Annual Scientific Meeting

Disclosures

Scan for slides:

- Salary support from Merck (~5%), paid directly to Penn, 2024-2026

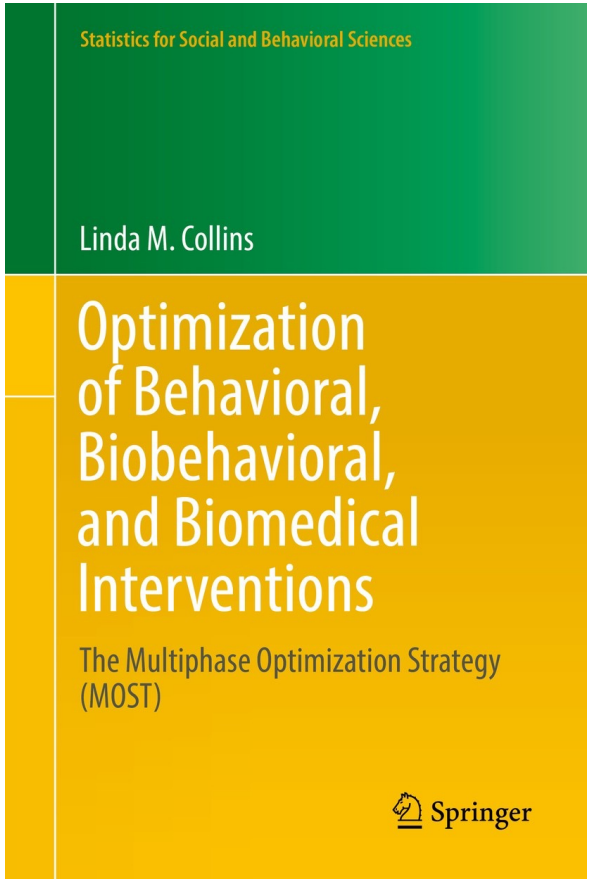
IO interventions are complex.

Scan for slides:

- Studying complex interventions requires **very** careful thought.
- Challenges can be mitigated with **high-quality design**.
  - Design of both the intervention *and* a trial.
- Well-designed interventions and studies make statistical analysis straightforward, *depending on the objective*.
  - Poor design makes it hard to achieve high-quality results ☹️

Plug: The Multiphase Optimization STrategy (MOST)

- **MOST** is a framework for developing and optimizing multicomponent or multiphase interventions
- Often used in behavioral or biobehavioral settings, but applicable in biomedical applications as well.
- *Preparation phase* develops a conceptual model for the intervention's mechanism.
- *Optimization phase* uses experimentation (RCTs) to identify effective components, delivery strategies, etc.
- *Evaluation phase* studies the effects of the optimized intervention.



Objectives, Endpoints, and Estimands

- A trial's **objective** is the scientific question of interest
  - "Is [intervention A] [superior to] [intervention B] in terms of [endpoint]?"
- An **endpoint** is the outcome measure for an *individual patient*.
  - PFS, OS, objective response (not response *rate*), etc.
  - Not estimated, *measured*.
- An **estimand** is a population-level quantity of interest that addresses the scientific question
  - Difference in ORR, median survival, etc.
  - Translates the objective to a statistical target quantity.
  - Estimated.

Some obvious points worth repeating:

- In a clinical trial,
- You can only study what you *do*.
- You can only compare what you *vary*.
- You can only *easily* compare what you *randomize*.

- ***Clearly establishing and justifying your primary scientific question is crucial.***

Establishing a Scientific Question

- Your trial is designed to answer specific scientific questions (objectives)
- Objectives must be **clear and explicit.**
- This is especially true for trials of complex interventions!

- Your scientific question should be about *one or more interventions* as they relate to some outcome/endpoint of interest.

- If investigating a multicomponent intervention, do you want to evaluate *each component* or the *entire package*?

Defining the Intervention

- Having **clearly defined, protocolized interventions** is crucial.
- What does this mean?
- Dose
  - For drugs *and* procedures: How much *stuff* are you delivering?
- Timing
  - When is the *stuff* being delivered?
    - Continuously? At pre-defined intervals? Trigger-based?

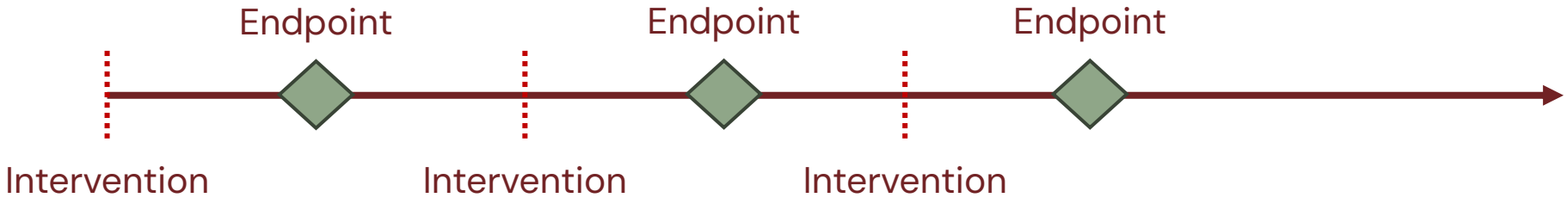
Defining the Intervention

- Having **clearly defined, protocolized interventions** is crucial.
- What does this mean?
- Duration
  - How long does the intervention last or take to be delivered?
  - E.g., how long between procedures, drug administrations, etc.?
- Minimum required components / compliance / adherence
  - What does a patient need to have received to have “complied” with the intervention?

Endpoint Timing and Measurement

- ”Time scale” of endpoints should match the scientific question and the intervention as defined.

- If the scientific question is about a single procedure, dose, etc., the endpoint should be **proximal** (near in time to treatment delivery).

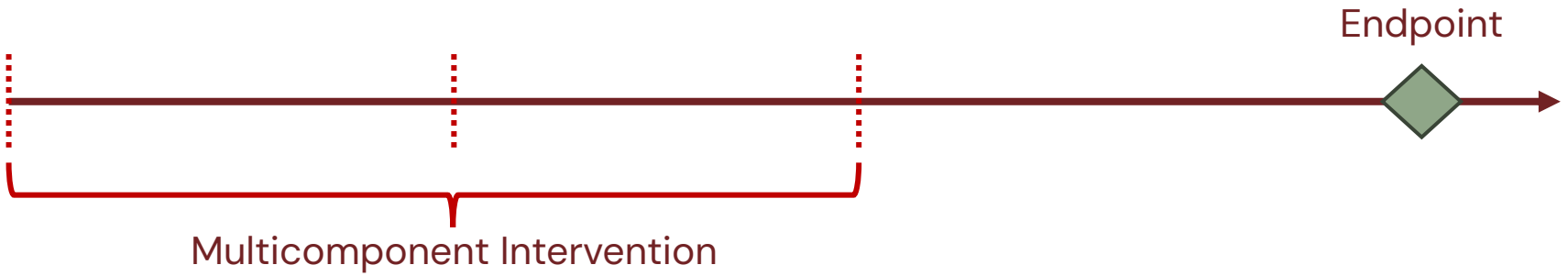


- This yields repeated measures on the same patient.

Endpoint Timing and Measurement

- ”Time scale” of endpoints should match the scientific question and the intervention as defined.

- If the endpoint is **distal** (far in time from treatment delivery), you may need to reconceptualize the intervention:



- We can only study the effects of all intervention components *jointly*.

Why does this matter?

- If delivering multiple treatments, and your question is about a single administration, the endpoint must be observable before the next administration.
- Otherwise, the endpoint could be affected by subsequent administrations, leading to a biased treatment effect estimate.
- Unpacking each administration's effect would complex statistical techniques.
- MOST is useful in this setting if interested in optimization.

Efficacy vs. Effectiveness

- **Efficacy** is how well something works in ideal settings (e.g., the lab)
  - Requires clear protocolization (of both treatment and control)
- **Effectiveness** is how well something works in the real world
  - Allows flexibility in implementation, clinical judgment, etc.

- Be very explicit about what in your intervention needs to be protocolized and what's allowed to vary (and *how*).

Time-to-event endpoints with staged or multicomponent interventions

- It's important to define endpoints to avoid **immortal time bias**
- Clocks should start with the intervention, *not* when the intervention is "complete"
  - If randomizing, clocks should start at randomization.
- Example: trial of continuous drug vs. staged embolization
  - If PFS starts after last embolization, the continuous drug arm has already had time for progressions to occur (BIAS!)

Get a biostatistician involved ASAP!

Scan for slides:

- **Involve us as early as possible!**
- We can help with:
  - Clarifying objectives, hypotheses, and endpoints
  - Trial design
  - Analytic strategies to improve efficiency
  - Improving trial data collection and quality
  - And much more!
- We’re more than a checkbox or a power calculation.
- You know more about IO than we do, but we learn fast.
  - Outside knowledge can be helpful: we’re good at asking “why?”