Key Issues in CER for Design and Analysis

(A statistician's view of intervention studies)

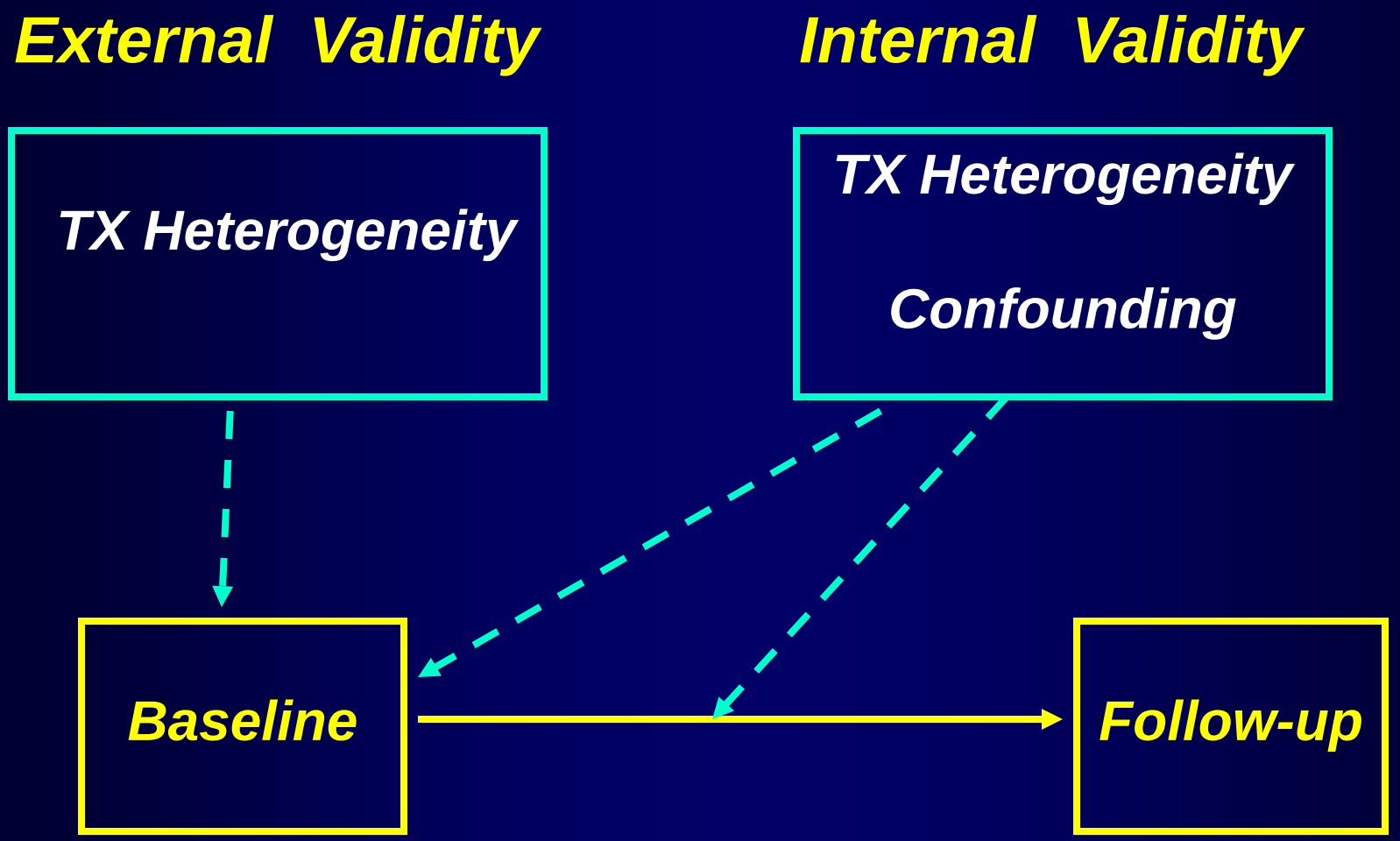
Thomas Ten Have
University of Pennsylvania

Feb 23, 2010

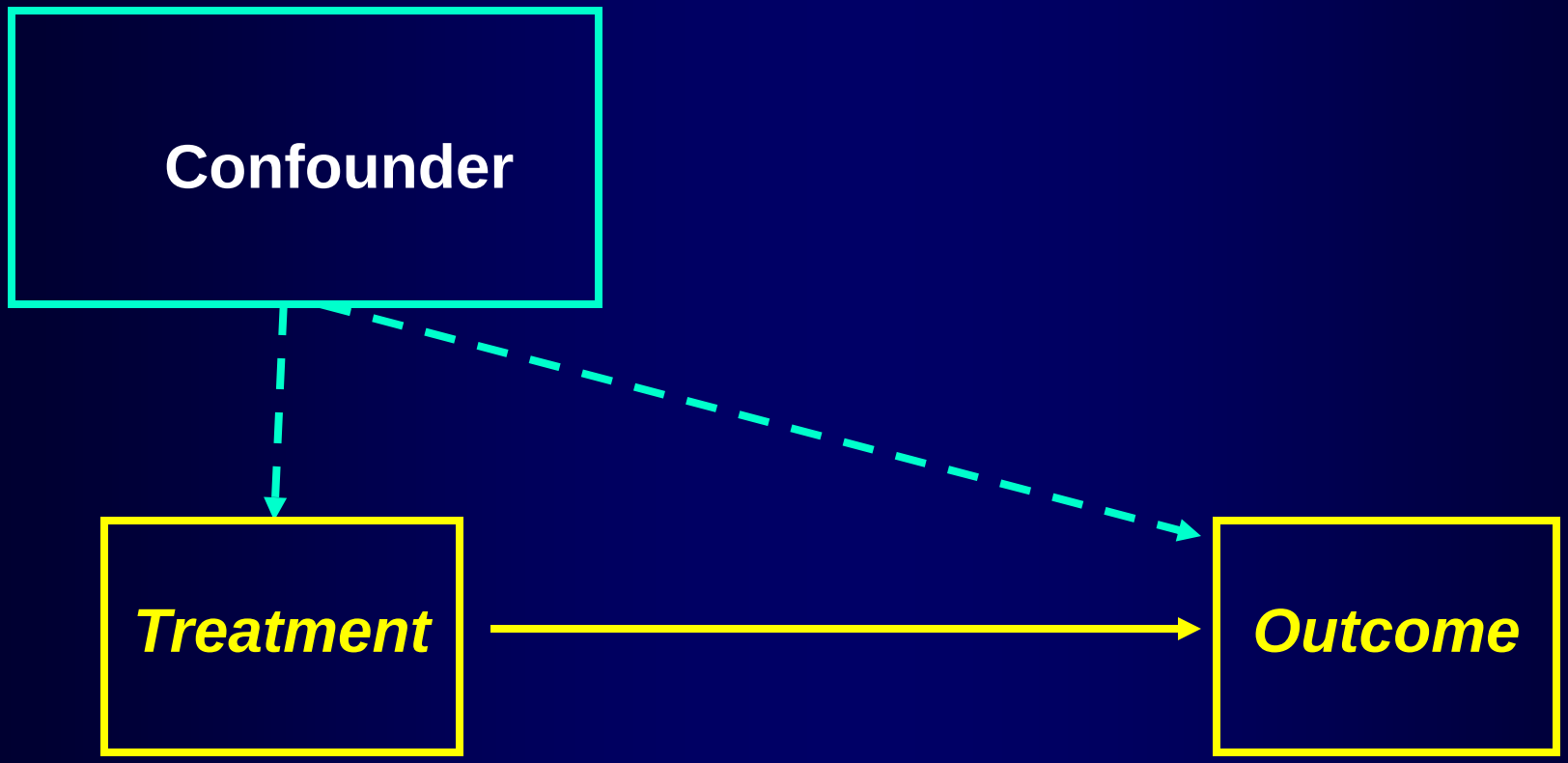
External vs. Internal Validity

- **External Validity: Inference for a specific study sample generalize to other populations and conditions**

- **Internal Validity: A causal relation between intervention and outcome is "appropriately" demonstrated for a given study sample.**


External Validity
Internal Validity
TX Heterogeneity
TX Heterogeneity
Confounding
Baseline
Follow-up

Confounding

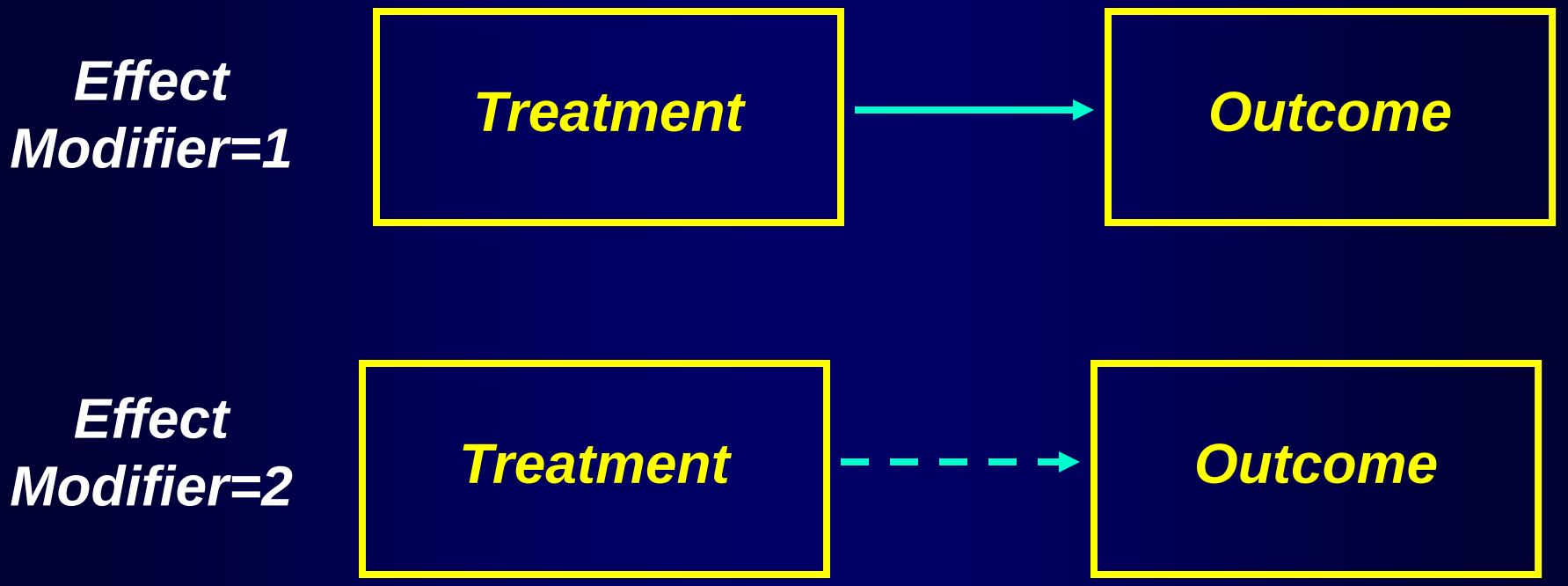


Confounding

Selection bias into treatment

- Disease severity -- Indication bias
- Patient/provider preference

Treatment Heterogeneity



Baseline Treatment Heterogeneity

- Gene-treatment interactions
- Optimizing interventions in sub-groups
- Personalizing treatment
- Cultural, social, and environmental, provider heterogeneity

Post Baseline Treatment Heterogeneity

- Early response/non-response
- Non-adherence
- Adverse events
- Changing co-morbidities
- Changing social and familial relationships

Design	Time 1	Time 2
Stratified Equipoise (External Validity)	patient preference disease severity	randomization
Randomized Encouragement/ Consent (Internal Validity)	Randomization	patient preference disease severity
Within-Subject Randomized Adaptive (Internal Validity)	randomization/ history, patient preference	randomization/ early response, patient preference, outcome
Step-Wedge (Extended Wait LIsting) (Interna+External Validity)	staggered randomization	staggered randomization

Stratified Equipoise Designs

Generalizability Benefit (External Validity):

Comparison of treatment effects between randomized and preference samples (treatment heterogeneity).

Confounding (Internal Validity):

Propensity score for **observed confounding**

Instrumental variable for **unobserved confounding**

Improving generalizability of fully randomized designs

Baseline data collection of confounding factors that relate to treatment

Survey sampling methods to extrapolate results to general population

Post-stratification/ raking

Assumes no treatment heterogeneity

Assessing post-baseline treatment heterogeneity:

Sequential randomization **design protects against confounding**

Without sequential randomization Instrumental variable methods **protect against such bias but…**

Post-baseline heterogeneity in baseline randomized trials

Under MacArthur criteria, post-baseline treatment heterogeneity is "mediation" (Kraemer et al. 2008).

Standard methods are biased and causal methods are required for unbiased effect size estimates.

The challenge is to assess assumptions of standard and causal methods.

Achieving both external and internal validity

More complex randomization designs

Affordable?

More complex analytic methods

Good baseline and post-baseline data on potential confounders?

Close examination of assumptions?