Transporting Treatment Effects from Randomized Trials to Real-World Target Populations

Jennifer L. Lund, PhD
Associate Professor of Epidemiology
Gillings School of Global Public Health
University of North Carolina at Chapel Hill

Karolinska Institutet
Tuesday, November 16, 2021

Funding, Conflicts of Interest, and Acknowledgements

Funding support

- This work was supported through a Patient-Centered Outcomes Research Institute (PCORI) Program Award (ME-2017C3-9337)
- All statements in this presentation, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee

Conflicts of interest

- Sanofi provided access to the individual-level clinical trial data to the authors and reviewed findings before export from the online platform
- Jennifer Lund receives research support from the Center for Pharmacoepidemiology whose members include UCB, GlaxoSmithKline, Takeda, Boehringer Ingelheim, AbbVie, and (formerly) Merck

Acknowledgements

- Daniel Westreich developed selected content presented in this lecture; I thank and acknowledge him for his contributions.

Overview

- Randomized clinical trials (RCTs): The gold standard
- External validity: Why is this a potential problem for RCTs?
- Foundational concepts for addressing external validity

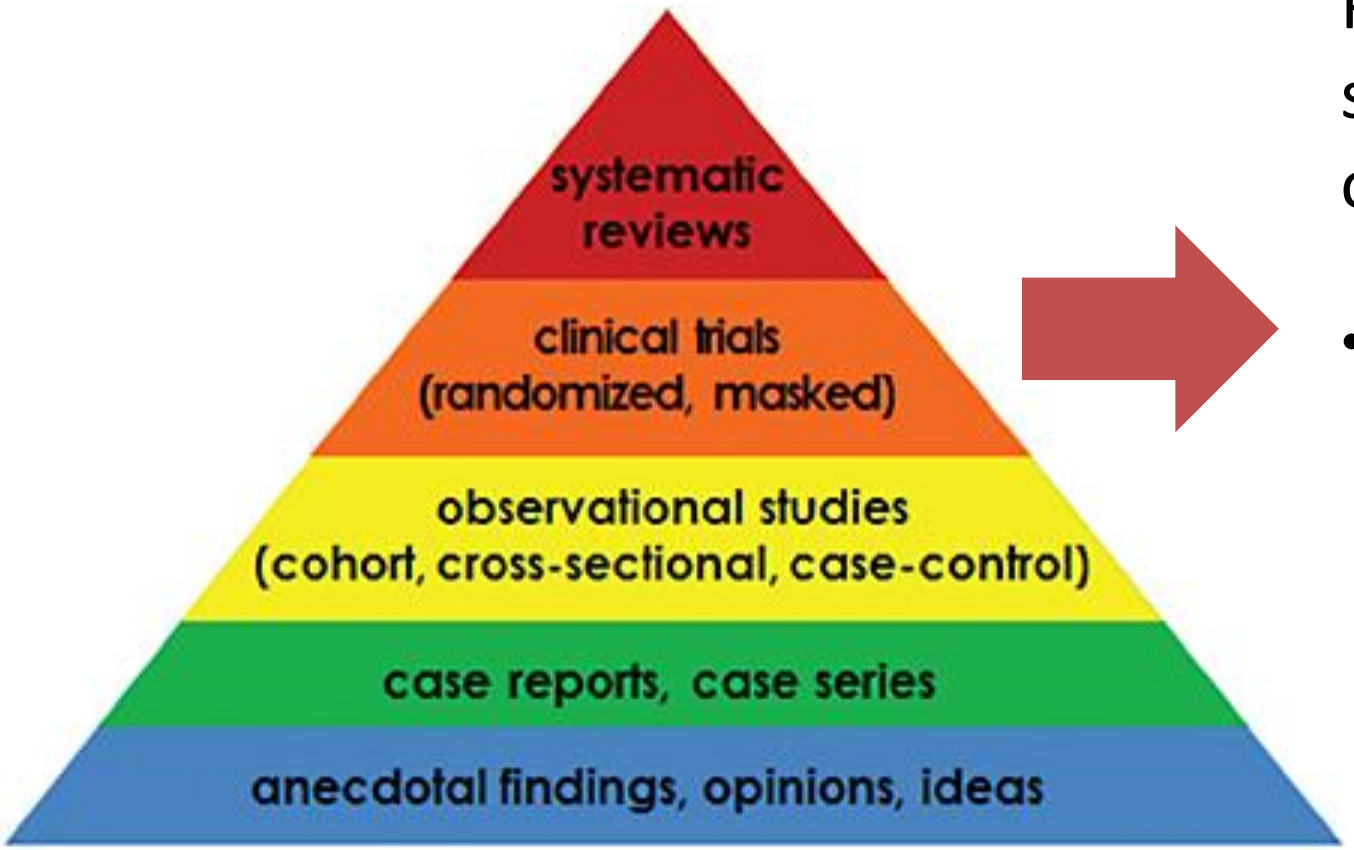
Randomized clinical trials (RCTs): The gold standard

Motivating example

- Adjuvant FOLFOX for the treatment of stage II and III colon cancer: how well does it work compared with the current standard of care?

- How do we address this question? In general, how do we learn about the efficacy of new treatments?

We do a randomized controlled trial…



Phase III clinical trials are the gold standard for determining the efficacy of interventions for drug approval

- Designed to reduce threats to ***internal validity*** through:
  - Restriction (inclusion/exclusion)
  - Randomization
  - Masking and standardized outcome assessment
  - Intensive treatment monitoring and follow-up

Randomized controlled trial design

This trial might look like the following:

- Screen N people and randomize them 1:1 to FOLFOX versus 5-FU
- Blind both participants and researchers
- Follow-up both groups for at least 5 years and count the number of deaths in each group

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Oxaliplatin, Fluorouracil, and Leucovorin as Adjuvant Treatment for Colon Cancer

Thierry André, M.D., Corrado Boni, M.D., Lamia Mounedji-Boudiaf, M.D., Matilde Navarro, M.D., Josep Tabernero, M.D., Tamas Hickish, M.D., Clare Topham, M.D., Marta Zaninelli, M.D., Philip Clingan, M.D., John Bridgewater, M.D., Isabelle Tabah-Fisch, M.D., and Aimery de Gramont, M.D., for the Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) Investigators

VOLUME 27 · NUMBER 19 · JULY 1 2009

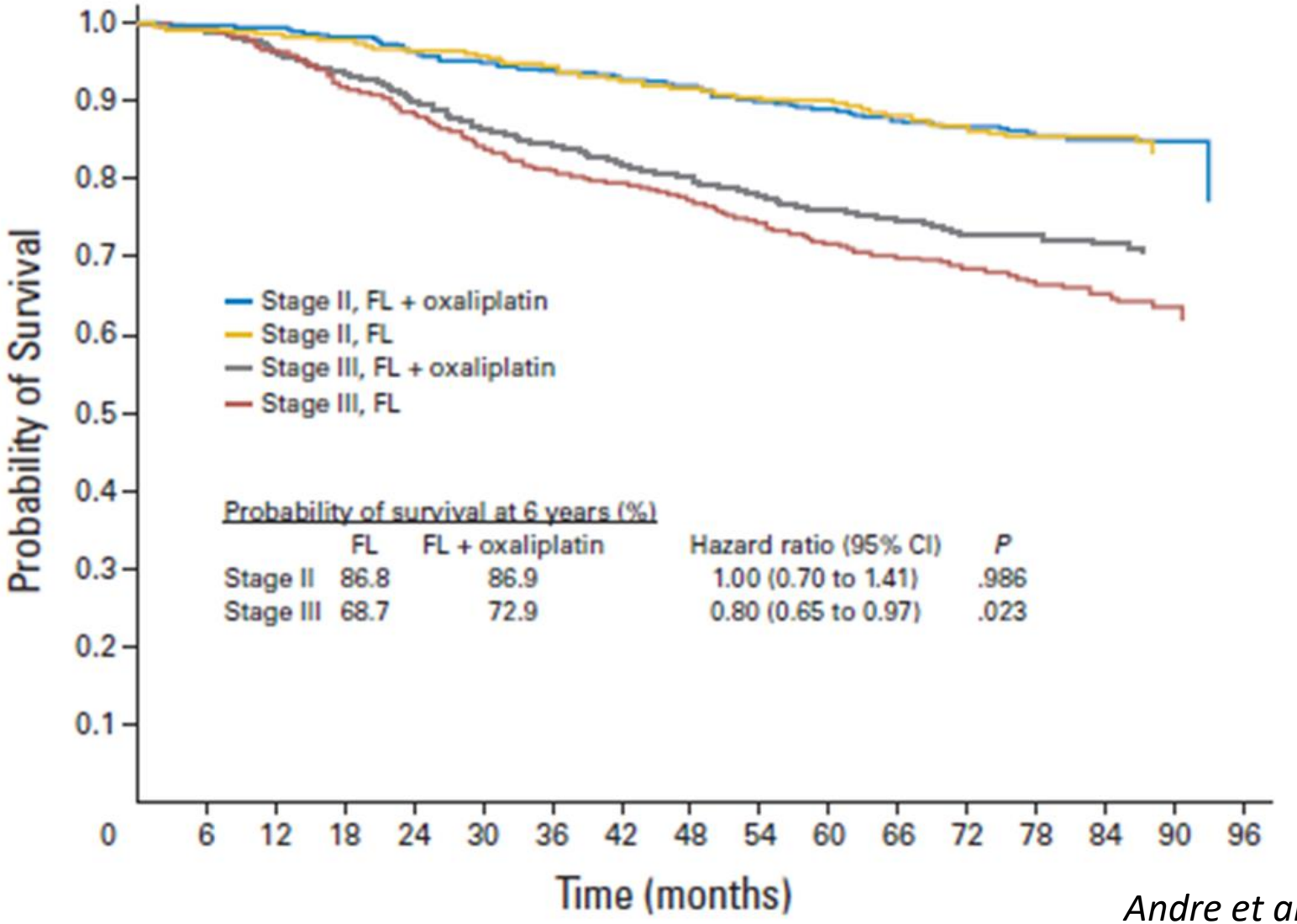
JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Improved Overall Survival With Oxaliplatin, Fluorouracil, and Leucovorin As Adjuvant Treatment in Stage II or III Colon Cancer in the MOSAIC Trial

Thierry André, Corrado Boni, Matilde Navarro, Josep Tabernero, Tamas Hickish, Clare Topham, Andrea Bonetti, Philip Clingan, John Bridgewater, Fernando Rivera, and Aimery de Gramont

What is reported from the trial



Andre et al, JCO, 2009

So, where do we go from here?

In the US, usually another confirmatory phase III trial is conducted and if successful, the treatment receives regulatory approval

Results from trials are incorporated into clinical treatment guidelines

Healthcare providers and patients make decisions about whether this new treatment is appropriate should be used *(note: oncology is sometimes unique in that off-label use is relatively common)*

External validity: Why is this a problem for RCTs?

Internal versus external validity

Internal validity

"The degree to which the results of a study are correct for the sample of individuals being studied."

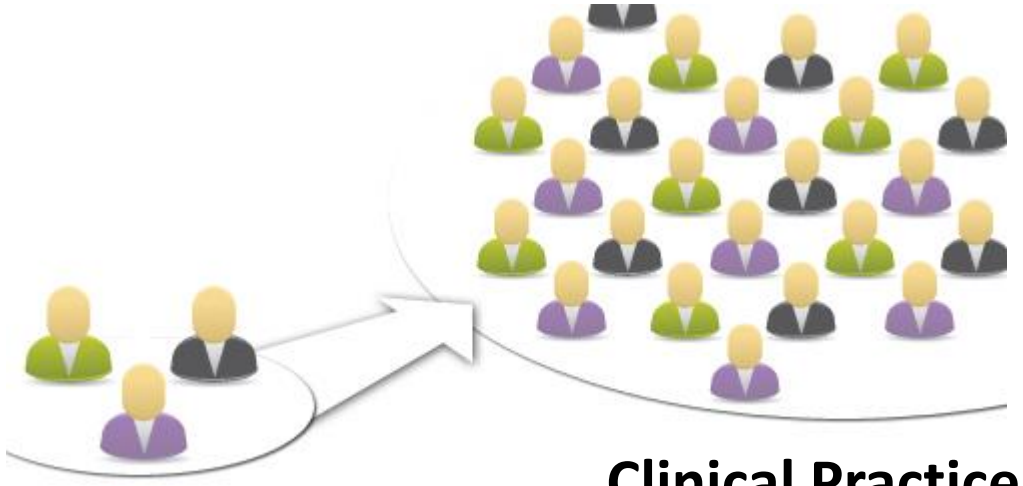
External validity

"The degree to which the results of an observation hold true in other settings."

Fletcher and Fletcher, *Clinical Epidemiology: The Essentials, 4th Ed*

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

External validity and comparative effectiveness research



One of the primary motivators of comparative effectiveness research (CER)

Understand the benefits and harms of alternative interventions in routine (or real-world) clinical practice settings.

Internal versus external validity

A trial will usually identify an internally valid sample **average causal effect (ACE)** – where the sample is the trial population.

But what is the **target population**? We are rarely interested in the trial population for its own sake, yet it is rare that we formally identify the target population. We almost never describe the target population in detail.

If the **target population differs systematically from the trial sample**, the average causal effect estimated from the study sample may not generalize (or "transport") unconditionally. Example:

Intervention is more effective in women than men.

Our study oversampled men compared to a target.

ACE from trial ≠ **ACE** in the target.

Why would treatment effects differ in trials and target populations?

Trial population

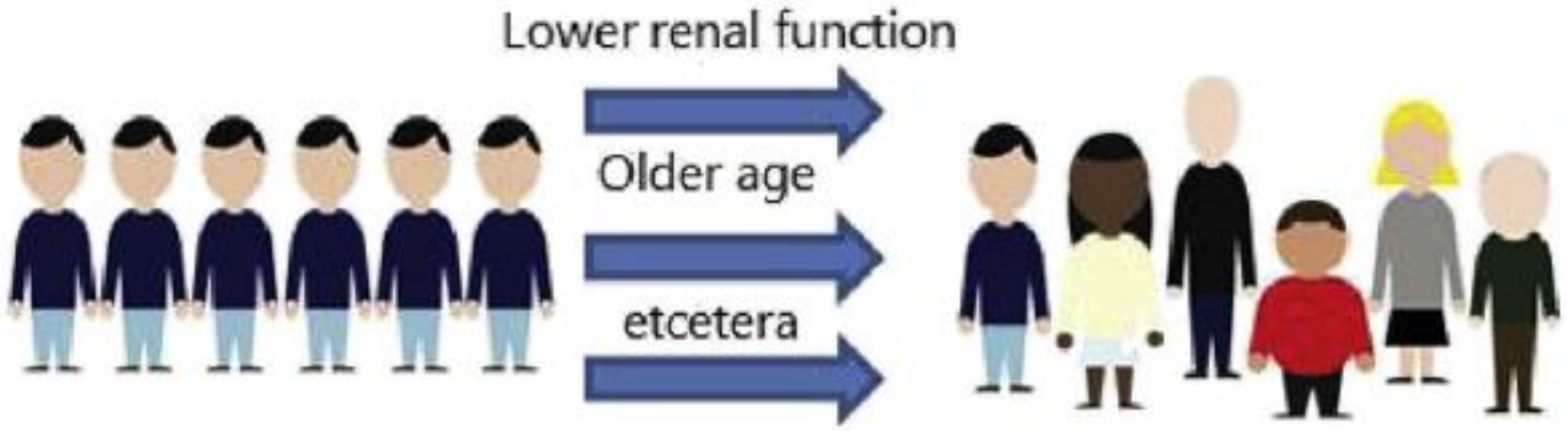
Clinical practice population

Zuidgeest et al, J Clin Epi, 2017

Why would treatment effects differ in trials and target populations?

Modifiers of drug effects

(i.e., treatment effect heterogeneity)



Trial population

Clinical practice population

Zuidgeest et al, J Clin Epi, 2017

Assessment of the external validity of trial results

Informal assessments

- Assume external validity (direct extrapolation)
- Discuss the potential impact of inclusion/exclusion criteria
- Note lack of generalizability outside of the trial

Quantitative approaches

- Explicitly evaluate assumptions of external validity
- Use trial and target population data to extend inferences from trial to target population of interest

Foundational concepts for addressing external validity

What is a target population?

From the survey sampling literature:

- The **target population** for a survey is the entire set of units for which the survey data are to be used to make inferences.

- A study (or trial) **sample** is the group of people who take part in the investigation. Importantly, these individuals may differ systematically from those in the target population.

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods* (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947

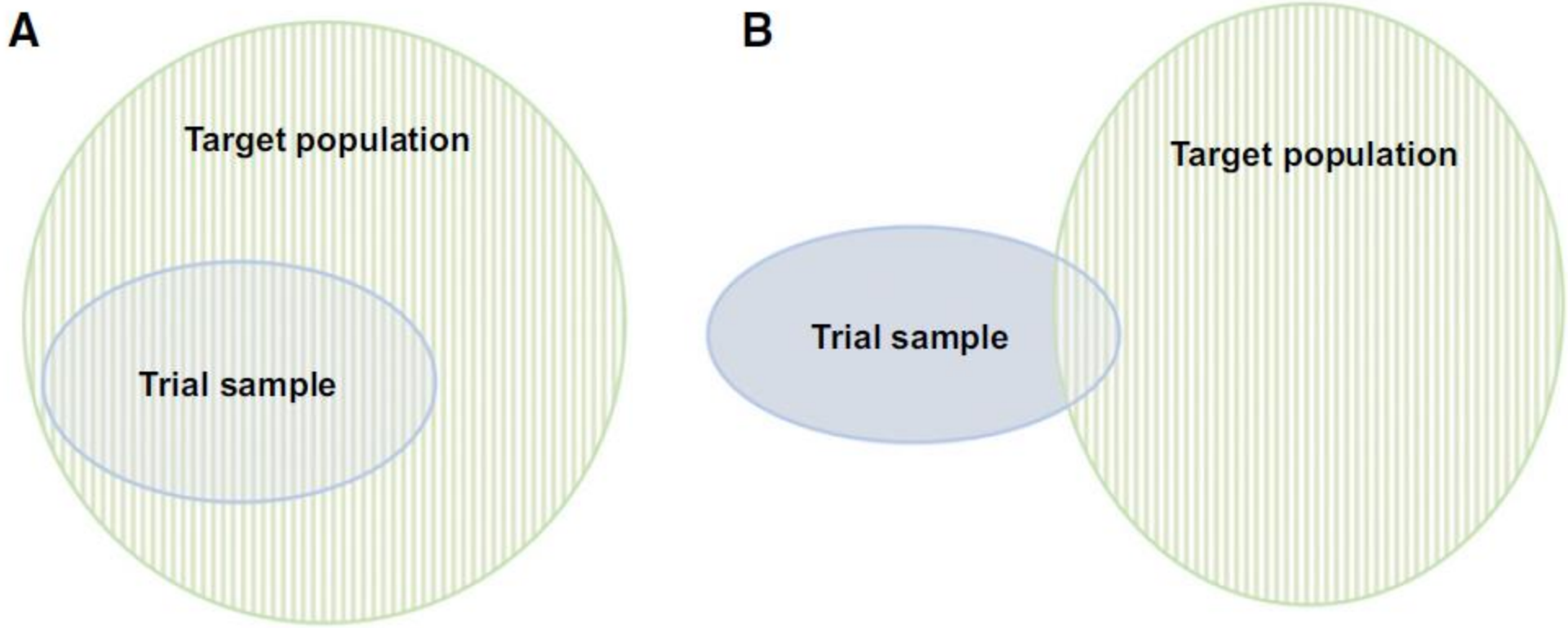
Two “flavors” of external validity – in words

- When study sample is a **subset of the target population** – we call this generalizability.

- When study sample is (at least partially) **external to target population** - call it transportability.

Note: *for purposes of this workshop, we can think of these largely as the same problem – extending inferences from trials to target populations of interest (whether a subset of the target or not)*

Two “flavors” of external validity – in pictures



Generalizability

Transportability

Inoue et al, CEBP, 2021

What is a target population in our example trial?

- No explicit mention of the target population (which is the norm)
- Defined by inclusion/exclusion criteria?
  - Enriched in ways to make trials efficient (e.g., high disease risk, low comorbidity risk)
- Defined by Table 1 – the super population implied by Table 1?
  - Those who agreed to participate and be randomized

METHODS

PATIENTS

Patients were eligible if they had undergone complete resection of histologically proven stage II (T3 or T4,N0,M0) or stage III (any T,N1 or N2,M0) colon cancer, as defined by the presence of the inferior pole of the tumor above the peritoneal reflection — that is, at least 15 cm from the anal margin. Treatment had to be started within seven weeks after surgery. Other eligibility criteria included an age of 18 to 75 years; a Karnofsky performance-status score of at least 60; a carcinoembryonic antigen level of less than 10 ng per milliliter; the absence of prior chemotherapy, immunotherapy, or radiotherapy; and adequate blood counts and liver and kidney function. Written informed consent was required from all patients, and the study was approved by the ethics committees of the participating centers.

What is a target population in our example trial?

- What about time and geography?

MOSAIC trial sites:

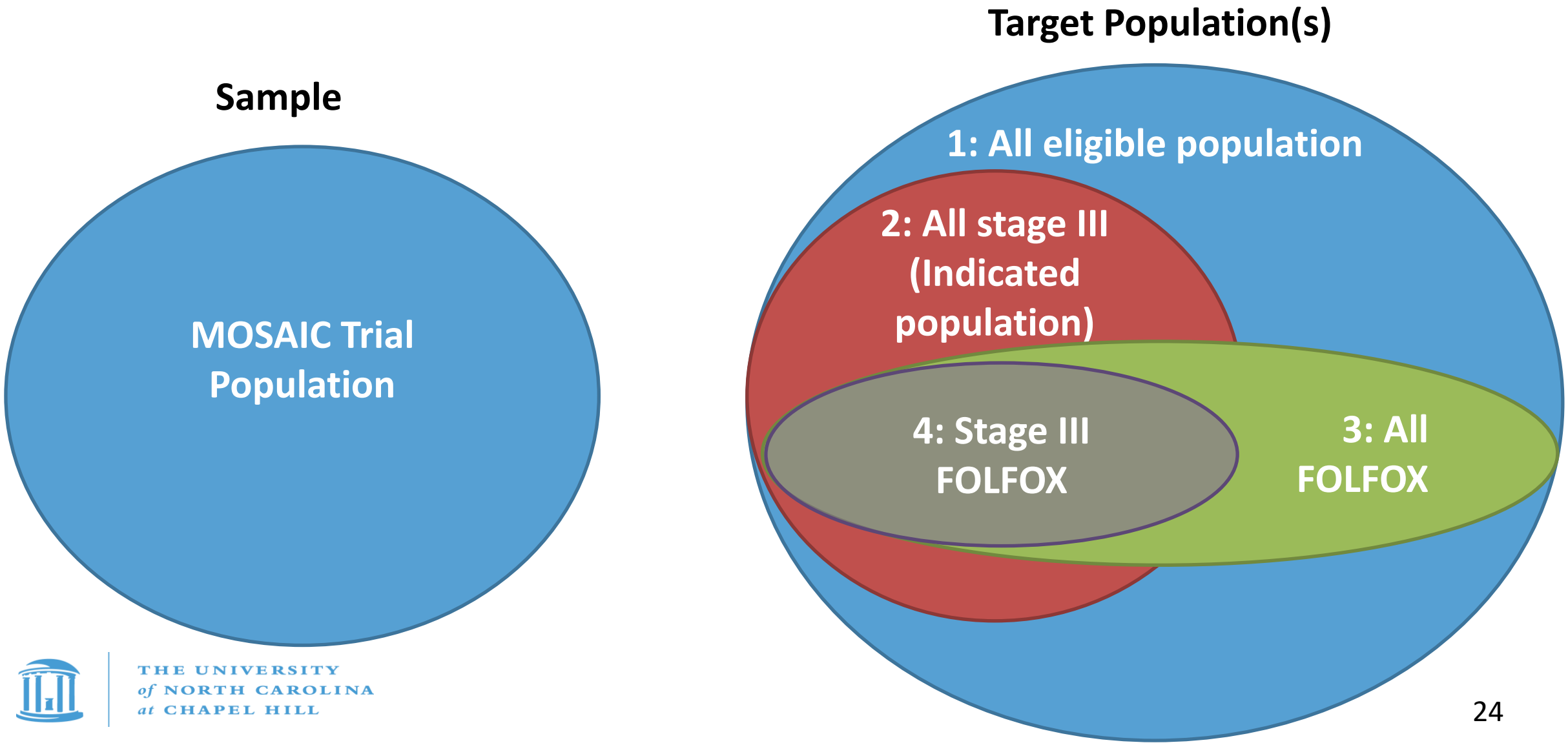
Australia, Austria, Belgium, Denmark, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Singapore, Spain, Sweden, Switzerland, UK

MOSAIC enrollment dates:

October 1998 – January 2001

What if our target population is geographically outside of the MOSAIC trial setting (e.g., the United States)? What if we want to make inference to a future population (e.g., patients in community setting) who will have access to this treatment after approval?

Example: Target populations for MOSAIC trial



How do we address the problem of external validity?

1) Trial design approaches

- Doubly randomized trials: Identify target population, randomly sample study sample, randomized treatment in study sample.
  - This won't work, because not everyone says yes.

- Pragmatic trials: Trials embedded within clinical practice settings
  - This may work but isn't always possible

2) Analytic approaches

- Purely observational analyses
- Hybrid or "data fusion" methods (using trial and observational data)

Target validity: Optimizing internal and external validity

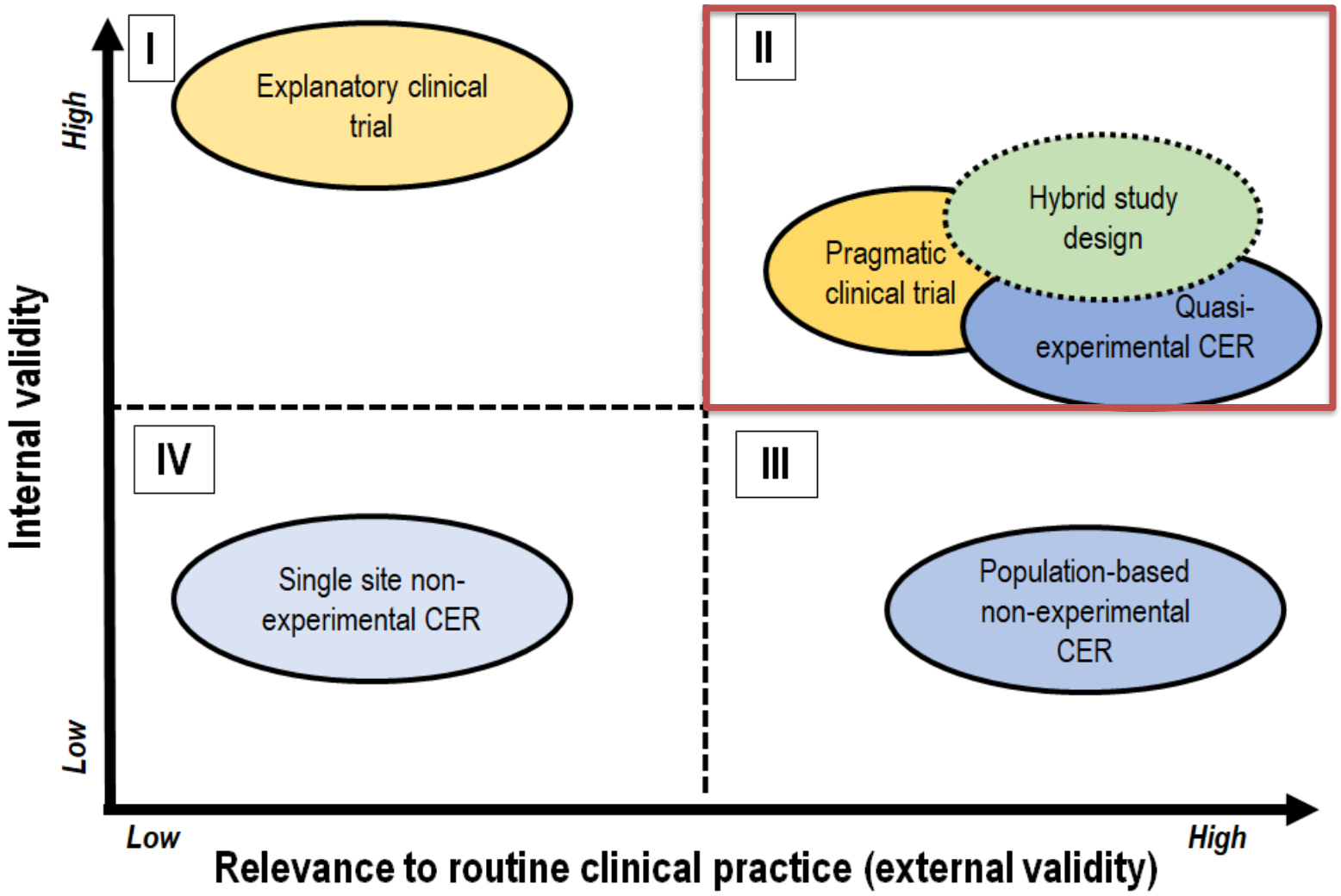


*Illustration demonstrates relative trade-offs among different approaches, but distance between each approach does not have quantitative meaning.

Growing attention to hybrid analytic approaches

Cole & Stuart Am J Epid 2010 presented a method using inverse probability weights to standardize clinical trial results to an external target population.

Pearl and Bareinboim (and others) have presented several papers presenting causal diagram-based methods for identification of externally valid causal effects.

Dahabreh & Hernan (2020) present a tutorial on identification criteria and methods that can be used to extend causal inferences from trials to a target population of non-participants

Identification conditions for generalizability (1)

Parallel to those for internal validity:

Instead of exchangeability (no uncontrolled confounding or selection bias) we have **external-exchangeability** (no uncontrolled effect measure modification).

Instead of positivity (non-zero probability that every subject under study could have any treatment – that is, no empty table cells) we have **transport-positivity** (no empty table cells in the target population vis a vis the study sample)

> *e.g., if you have a bunch of 70-year-old women in the target population, and no women over age 50 in your study sample, then you can't estimate effects in the 70-year-olds. That is, unless you extrapolate your models, which is exactly how you can overcome non-positivity for internal validity.*

Lesko C et al., Epidemiology, 2017
Dahabreh IJ et al, Statistics in Medicine, 2020

Identification conditions for generalizability (2)

- Per Hernán & VanderWeele (Epidemiology, 2011) we also need similar versions of treatment in the study sample and target population aka **transport-consistency**.

- Likewise, need similar patterns of interference between study sample and target population.

Implementation: How can we carry-out a hybrid analytic approach to address external validity?

- Hand-off to Michael…

Acknowledgements

This work was supported by a PCORI contract (ME-2017C3-9337)

Collaborators: Michael Webster-Clark, Alexander Keil, Hanna Sanoff, Til Stürmer, Daniel Westreich, Amy Lowman, Courtney Schlusser, Sharon Peacock Hinton, Allison Musty, Shahar Shmuel

Data partners: Sanofi and ClinicalStudyDataRequest.com for data access; colleagues at US Oncology (Marley Boyd, Jen Frytak, Nicholas Robert)

Questions?

Jennifer.Lund@unc.edu

Key References

1. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. *Am J Epidemiol.* 2010;172(1):107-115.

2. Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW, Grobbee DE; GetReal Work Package 3. Series: Pragmatic trials and real world evidence: Paper 1. Introduction. J Clin Epidemiol. 2017 Aug;88:7-13.

3. Andre T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med. 2004;350(23):2343-2351.

4. André T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C, Bonetti A, Clingan P, Bridgewater J, Rivera F, de Gramont A. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. J Clin Oncol. 2009 Jul 1;27(19):3109-16.

5. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR. Transportability of Trial Results Using Inverse Odds of Sampling Weights. Am J Epidemiol. 2017;186(8):1010-1014.

6. Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. J R Stat Soc Ser A Stat Soc. 2001 Apr 1;174(2):369-386.

7. Dahabreh IJ, Robertson SE, Steingrimsson JA, Stuart EA, Hernán MA. Extending inferences from a randomized trial to a new target population. Stat Med. (Epub 2020 Apr 6).