CER in Oncology:
Emerging Methods and Application

Jennifer L. Lund, PhD

Associate Professor of Epidemiology

February 11, 2020

UNC
GILLINGS SCHOOL OF
GLOBAL PUBLIC HEALTH

Funding, Collaborators, and Acknowledgements

This work was supported by a Lineberger Developmental Award and PCORI contract ME-2017C3-9337

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

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

Overview

RCTs and CER in oncology

Motivating example: Adjuvant chemotherapy for stage II-III colon cancer

- MOSAIC trial results
- Observational evidence

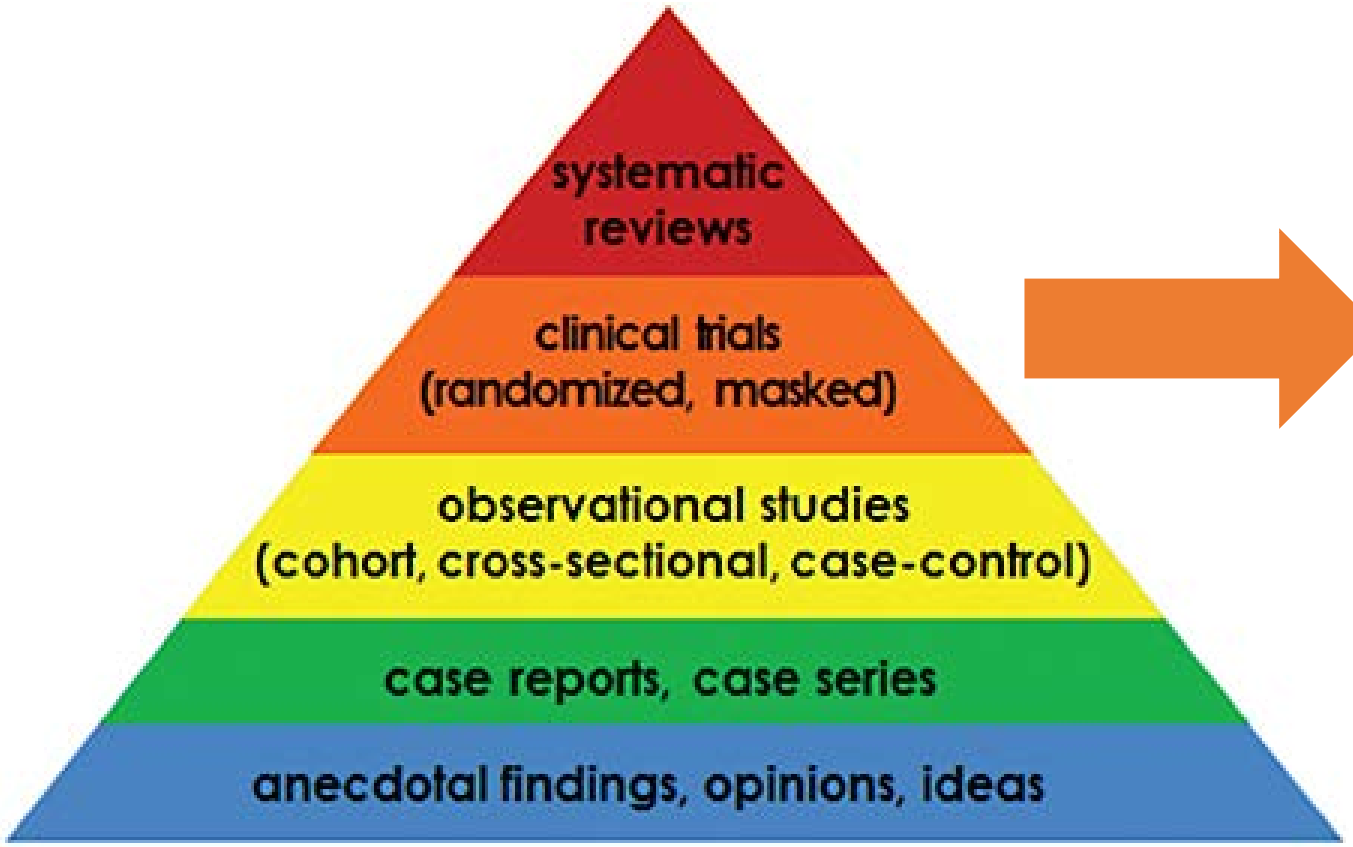
An emerging hybrid approach for oncology CER

- Study 1: FOLFOX versus 5FU in older adults
- Study 2: FOLFOX versus 5FU in community oncology practice

Clinical and research implications and next steps

RCTS and CER in oncology

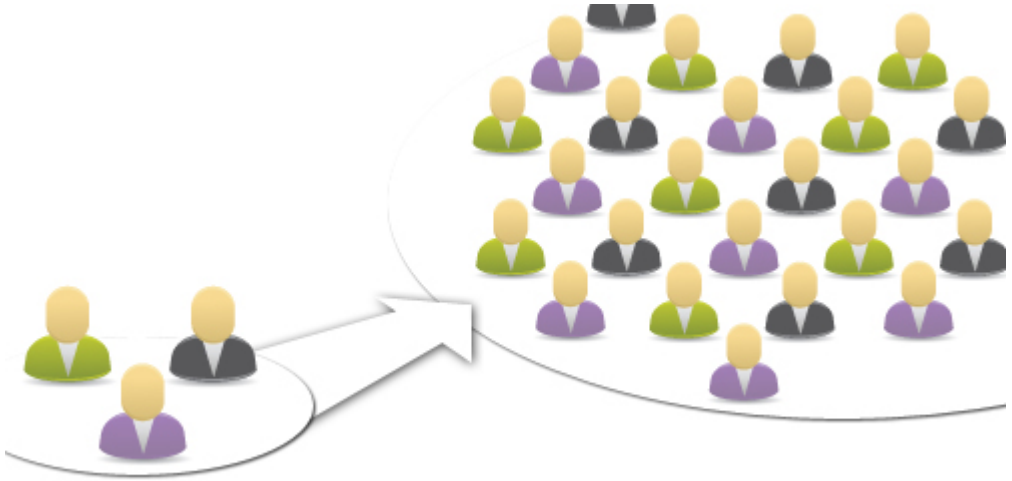
RCTs in clinical medicine



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

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

But what about external validity of RCTS?



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

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

One of the primary motivators of comparative <u>effectiveness</u> research (CER)

Understand the benefits and harms of alternative interventions in routine clinical practice settings.

Efficacy versus effectiveness

Efficacy: Can it work?

The extent to which a specific intervention, procedure, regimen, or service produces a beneficial effect under ideal conditions.

Effectiveness: Does it work?

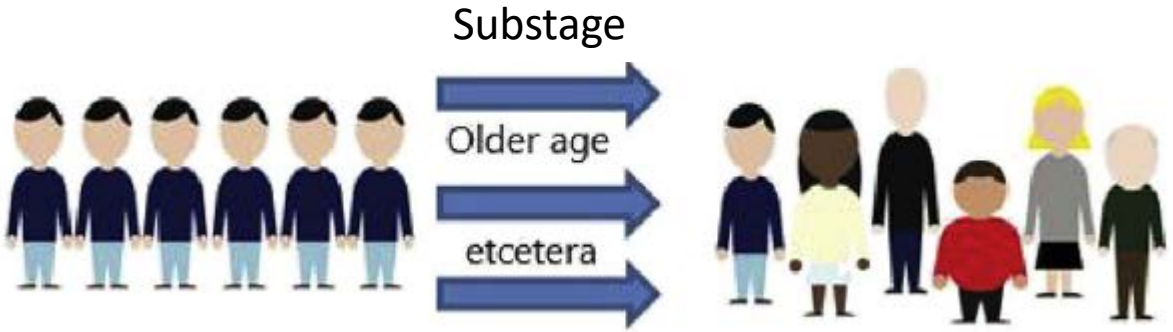
The extent to which a specific intervention, procedure, regimen, or service when applied in the real-world (i.e., clinical practice) Does what it is intended to do.

Last et al, Dictionary of Epidemiology, 5th Edition

Why would the effects of interventions differ between clinical trial and clinical practice settings?

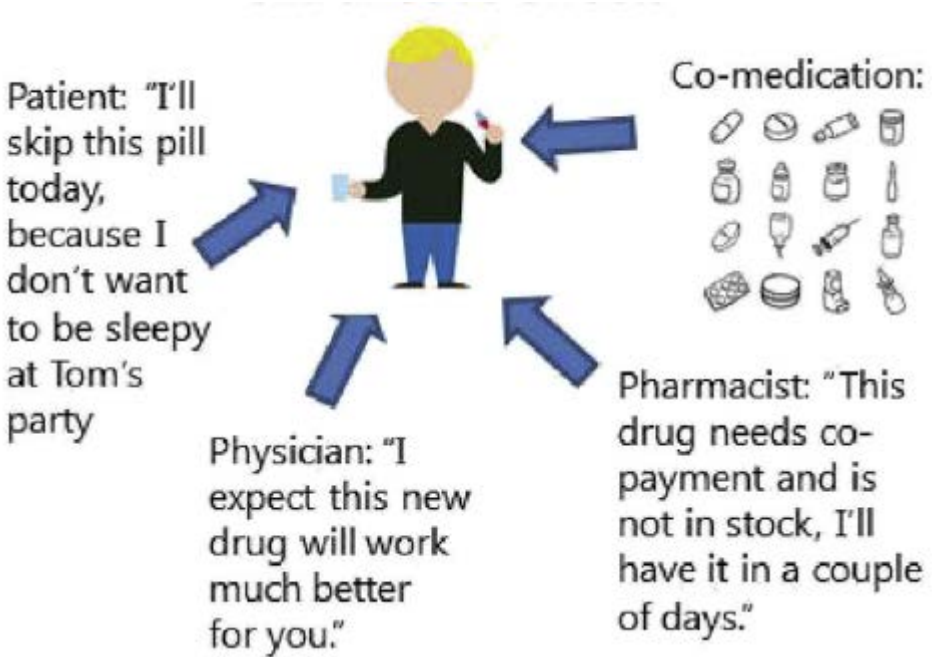
Modifiers of drug effects

(i.e., treatment effect heterogeneity)



Differential treatment delivery

(e.g., dose delays and reductions)



Zuidgeesta et al, J Clin Epi, 2019

Who cares about the effectiveness of cancer therapies?

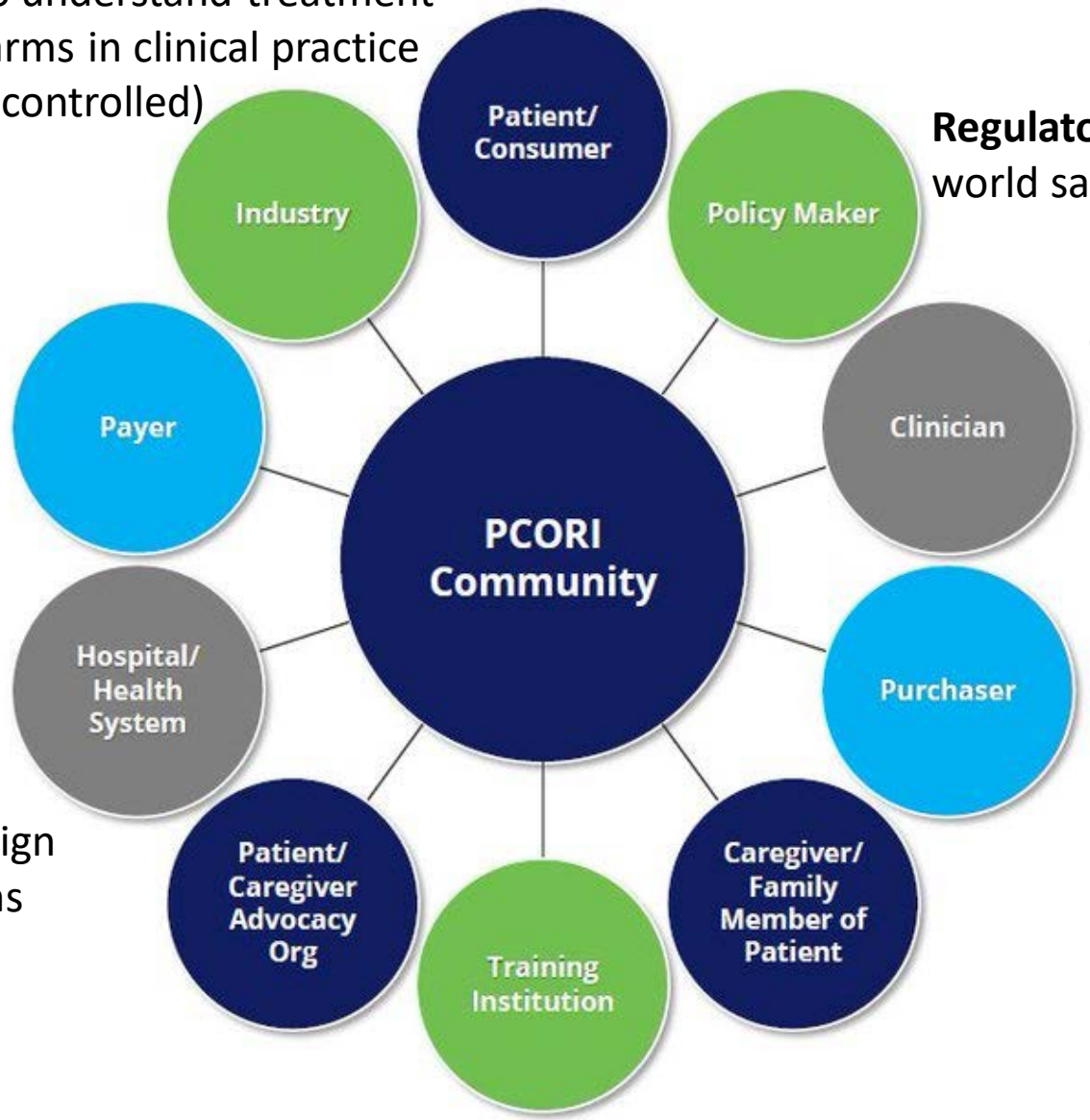
Patients: To understand treatment benefits/harms in clinical practice (not highly controlled)

Regulators: To monitor real-world safety and effectiveness

Payers: To understand value; to make coverage and formulary decisions

Oncologists: To understand the benefits and harms to expect in real-world patient populations

Healthcare systems: To understand value; to design clinical pathway programs



How can we generate **robust knowledge**

about the effectiveness and safety of cancer therapies in **clinical practice**

to better inform **decisions**

made by **patients and healthcare providers, payers, and regulators**?

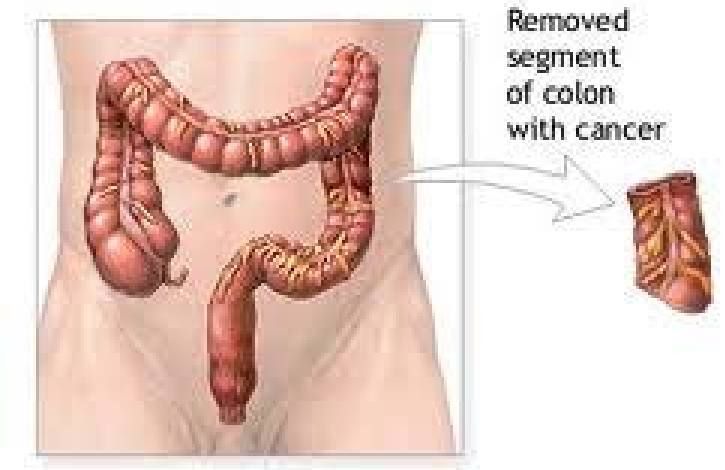
Motivating example

Adjuvant chemotherapy for stage II-III colon cancer

Comparative effectiveness of two adjuvant chemotherapy regimens for treating stage II and III colon cancer

FOLFOX **versus** **5-FU**

Treatment for stage II and III colon cancer



Surgical resection

Adjuvant chemotherapy

1980–2004: 5-fluorouracil (5-FU) was the mainstay for stage III disease*

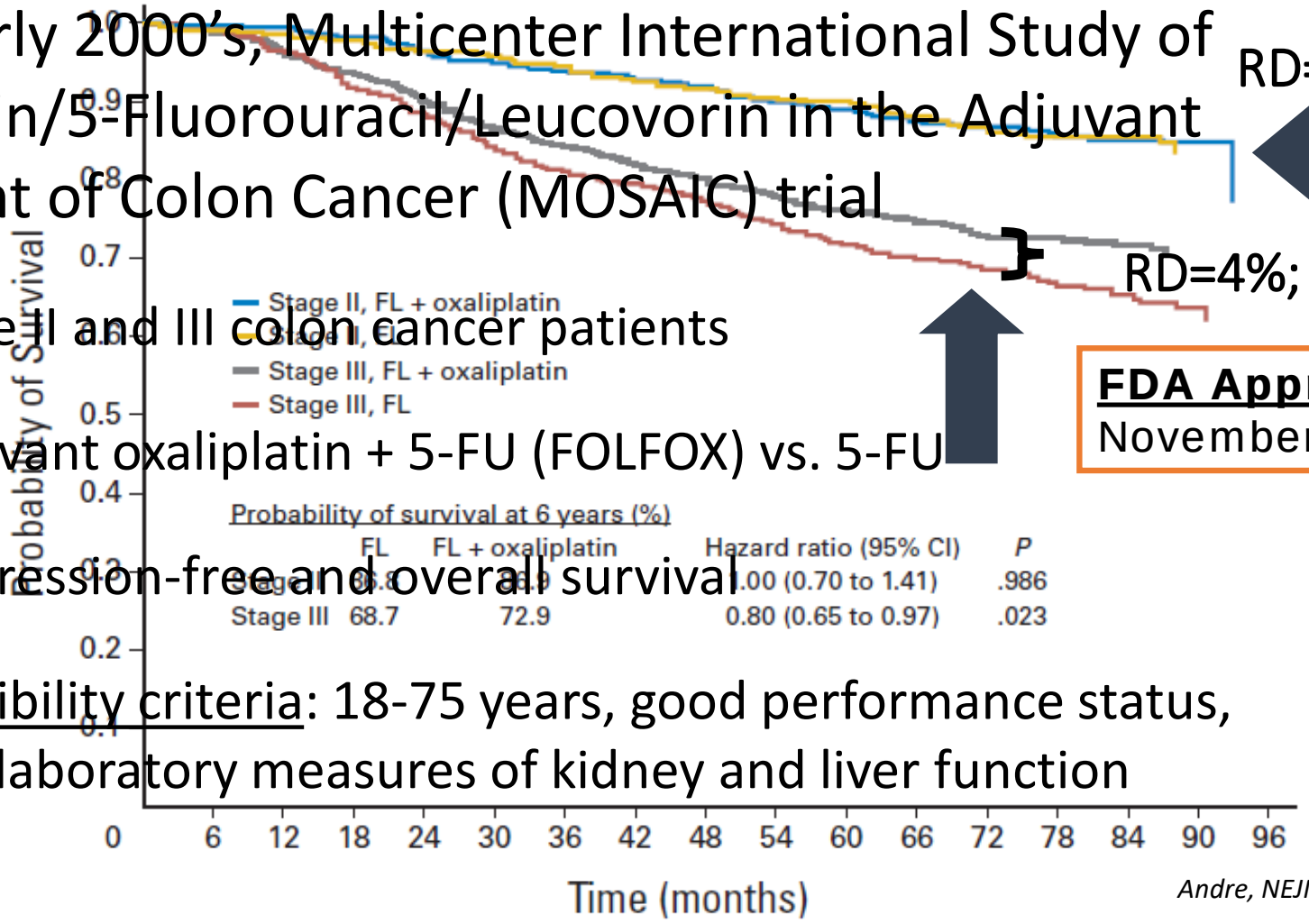
- 16% absolute reduction in mortality compared with surgery alone

Moertel, NEJM (1990)

Adjuvant oxaliplatin for stage II/III colon cancer

In the early 2000's, Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) trial

- Stage II and III colon cancer patients
- Adjuvant oxaliplatin + 5-FU (FOLFOX) vs. 5-FU
- Progression-free and overall survival



RD=0%; HR=1.0

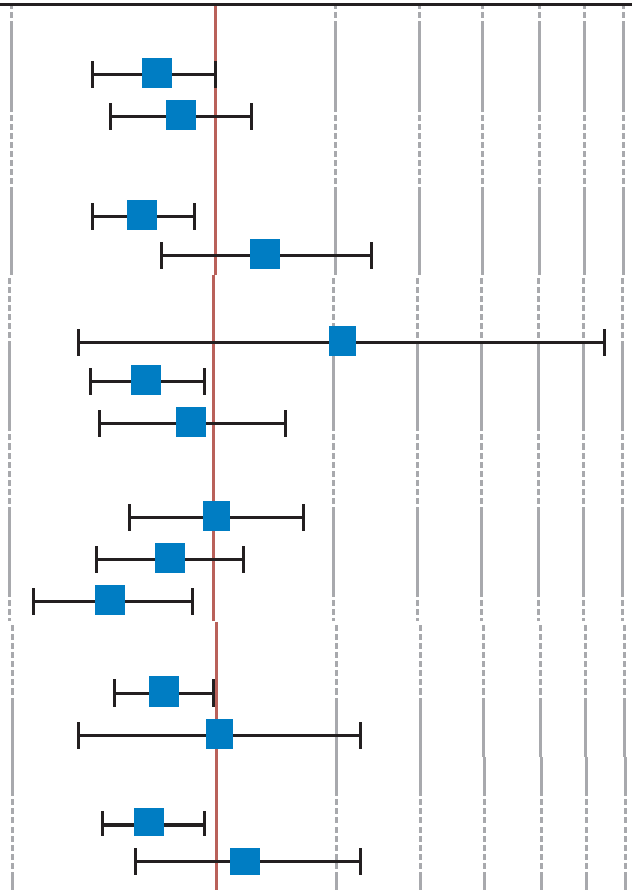
RD=4%; HR=0.80

FDA Approval: November, 2004

Major eligibility criteria: 18-75 years, good performance status, adequate laboratory measures of kidney and liver function

Andre, NEJM (2004,2009)

Treatment effect heterogeneity in MOSAIC, Overall survival

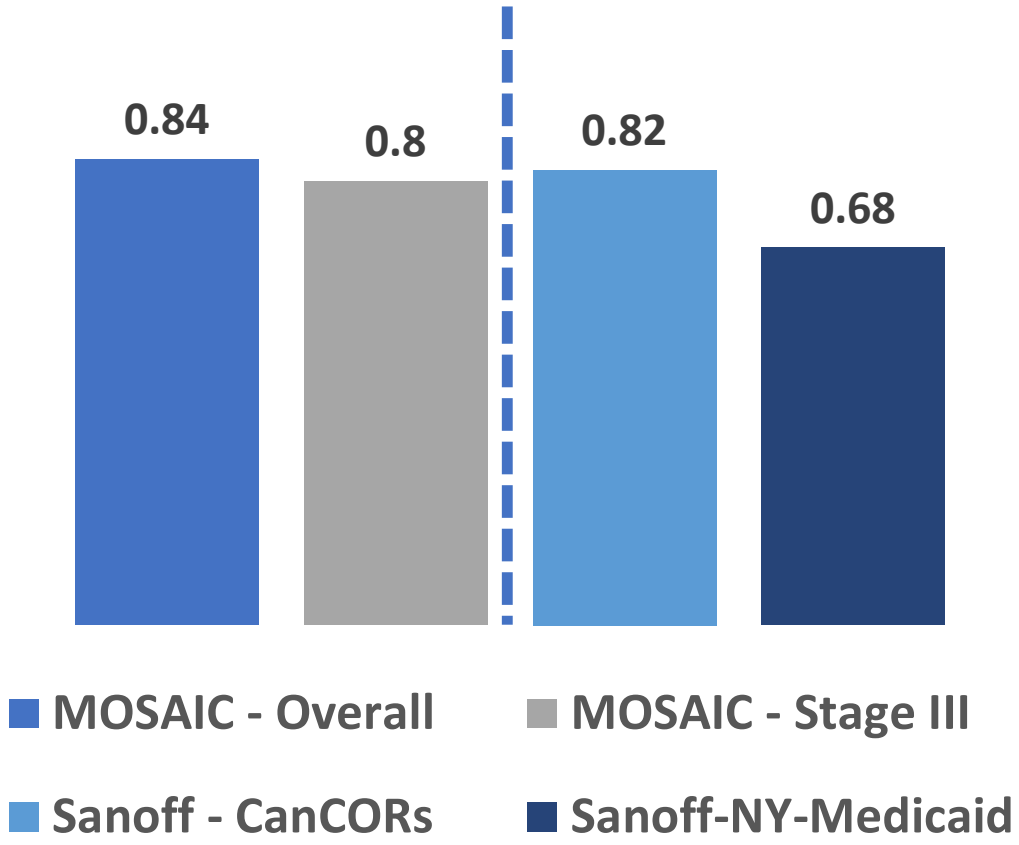


Subgroup	LV5FU2, n	FOLFOX4, n	HR	LCL	UCL
Sex					
Male	588	630	0.82	0.66	1
Female	535	493	0.89	0.7	1.13
Age, years					
≤ 70	963	968	0.78	0.66	0.93
> 70	160	155	1.19	0.83	1.7
T					
1–2	62	57	1.55	0.63	3.8
3	852	853	0.8	0.66	0.96
4	208	213	0.93	0.68	1.27
N					
0	447	452	1.01	0.75	1.36
1	442	440	0.86	0.67	1.11
2	231	229	0.71	0.54	0.93
Perforation					
No	1,045	1,045	0.84	0.71	0.99
Yes	78	78	1.01	0.62	1.64
ECOG PS					
0,1	984	968	0.8	0.68	0.95
2,3	139	155	1.11	0.76	1.64

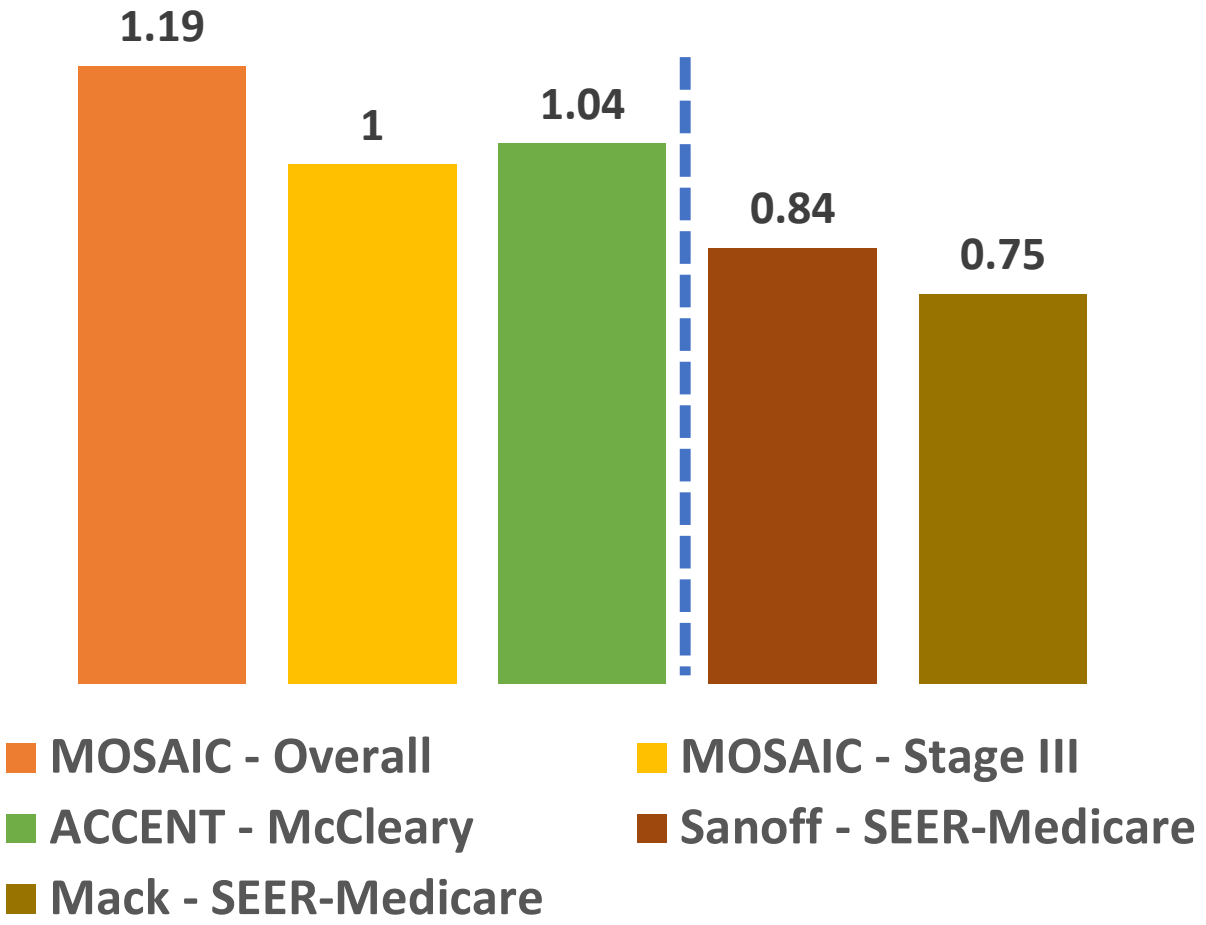
Andre, JCO (2015)

Evidence from trials and observational studies

Hazard Ratios for Mortality – All Ages



Hazard Ratios for Mortality – Older Adults



UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

“Balancing” internal and external validity in CER

Internal validity

Confounding

Selection bias

Misclassification

External validity

Diverse patients

Real-world monitoring and adherence

Clinical endpoints of interest (not surrogates)

An emerging hybrid approach for oncology CER

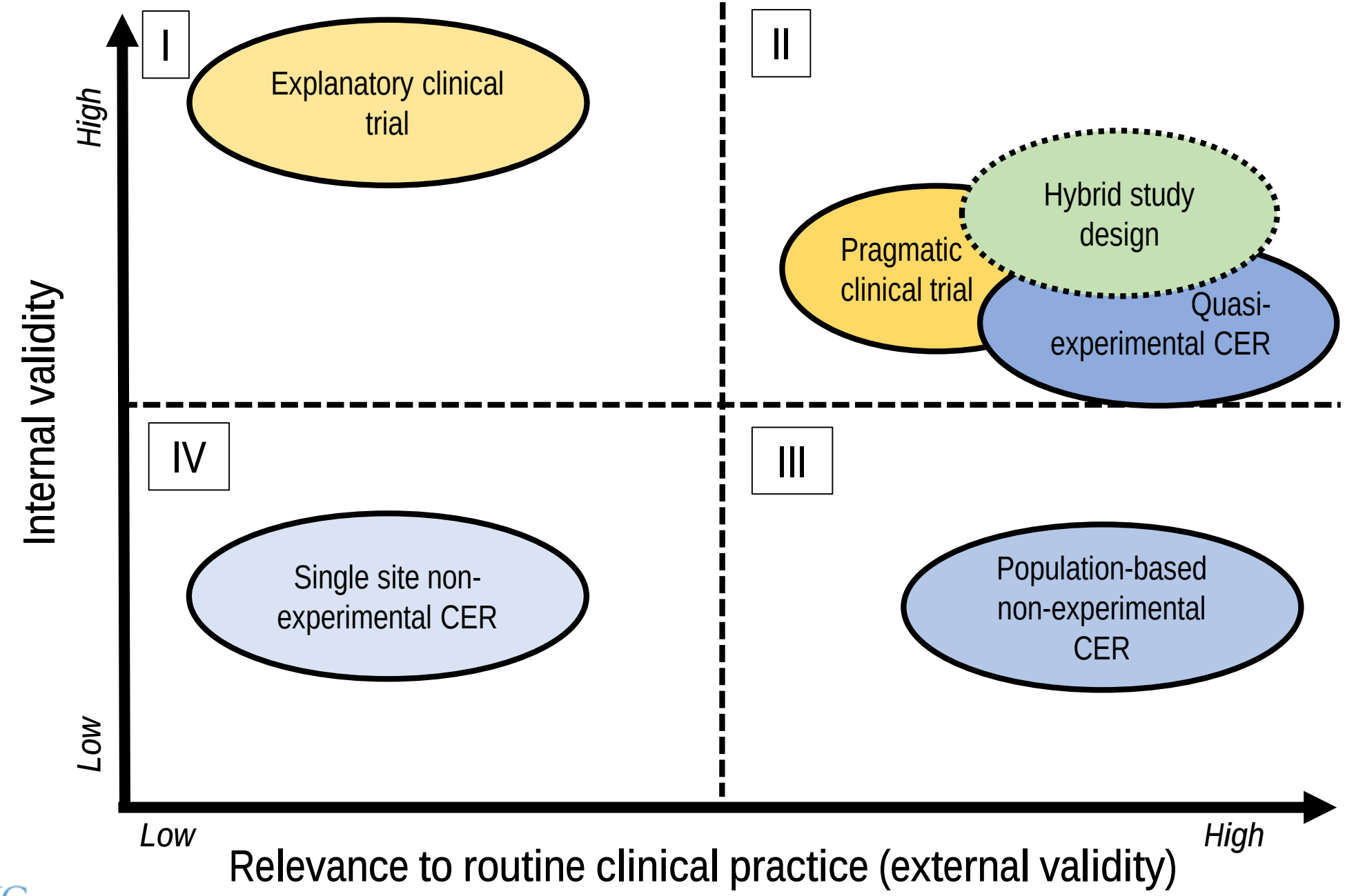
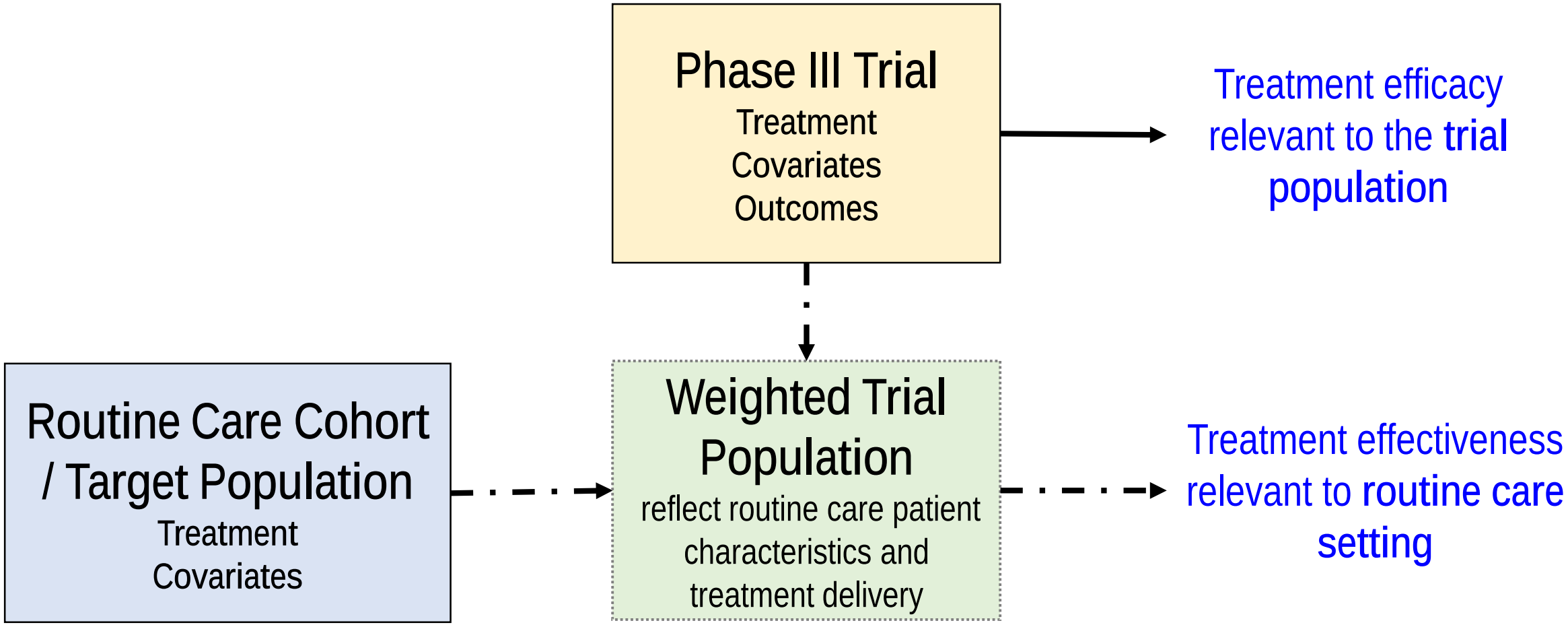
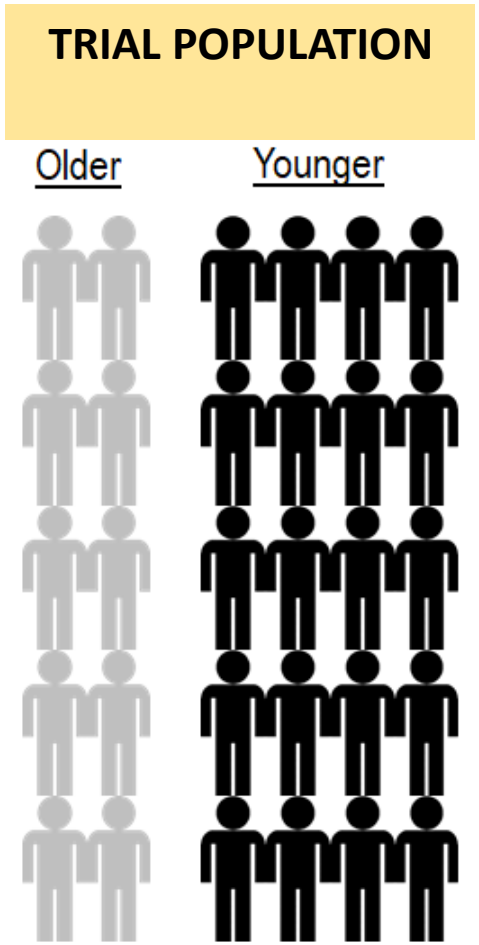


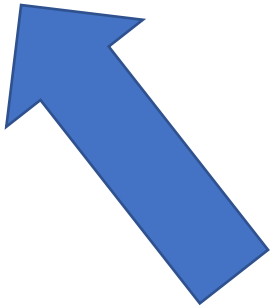
**Illustration demonstrates relative trade-offs among different approaches, but distance between each approach does not have quantitative meaning.*

Hybrid approach: blending of trial and observational data



A simple example of this hybrid approach





Each older person gets a **weight of 2** and each younger person gets a **weight of 0.5.**

How do you create these weights with multiple variables?

1. Similar to a propensity score, you can model the probability of trial enrollment,

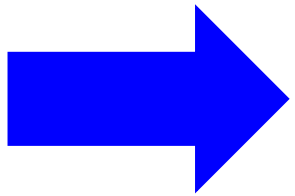
$$ln\left[\frac{Pr(trial{=}1 \mid z)}{Pr(trial{=}0 \mid z)}\right] = \beta_0 + \beta_1 z_1 + \cdots + \beta_k z_k$$

, as a function of covariates, z,

that are potential effect measure modifiers.

2. Create weights, as follows:

$$W_i = \begin{cases} \frac{Pr(trial_i{=}0 \mid Z_i)}{Pr(trial_i{=}1 \mid Z_i)} \times \frac{p(trial{=}1)}{p(trial{=}0)}, & trial{=}1 \\ 0, & trial{=}0 \end{cases}$$



3. Analyze the treatment effects of interest in the weighted trial population

Study 1: FOLFOX versus 5FU in older adults

- More than **50% of all colon cancers** are diagnosed in adults age 65 years+.
- This age group tends to be underrepresented in cancer clinical trials; thus evidence base for treating these patients is limited.
- Trial subgroup analyses and observational studies have produced conflicting findings regarding the potential benefits of FOLFOX vs 5FU in patients 65 and older.

What would the results of the MOSAIC trial have been if the trial population was randomly drawn from the Medicare beneficiaries aged 66-75 with stage II-III colon cancer initiating adjuvant chemotherapy with FOLFOX or 5FU?

Data sources and study population

MOSAIC Trial

Enrolled: 1998-2001 (146 centers, 20 countries)

Key eligibility: age 18-75, stage II-III colon cancer, KPS 60+, adequate organ function

Study population: age 66-75 years

Surveillance, Epidemiology and End Results (SEER)-Medicare

Diagnosed: 2004-2011 (US SEER regions)

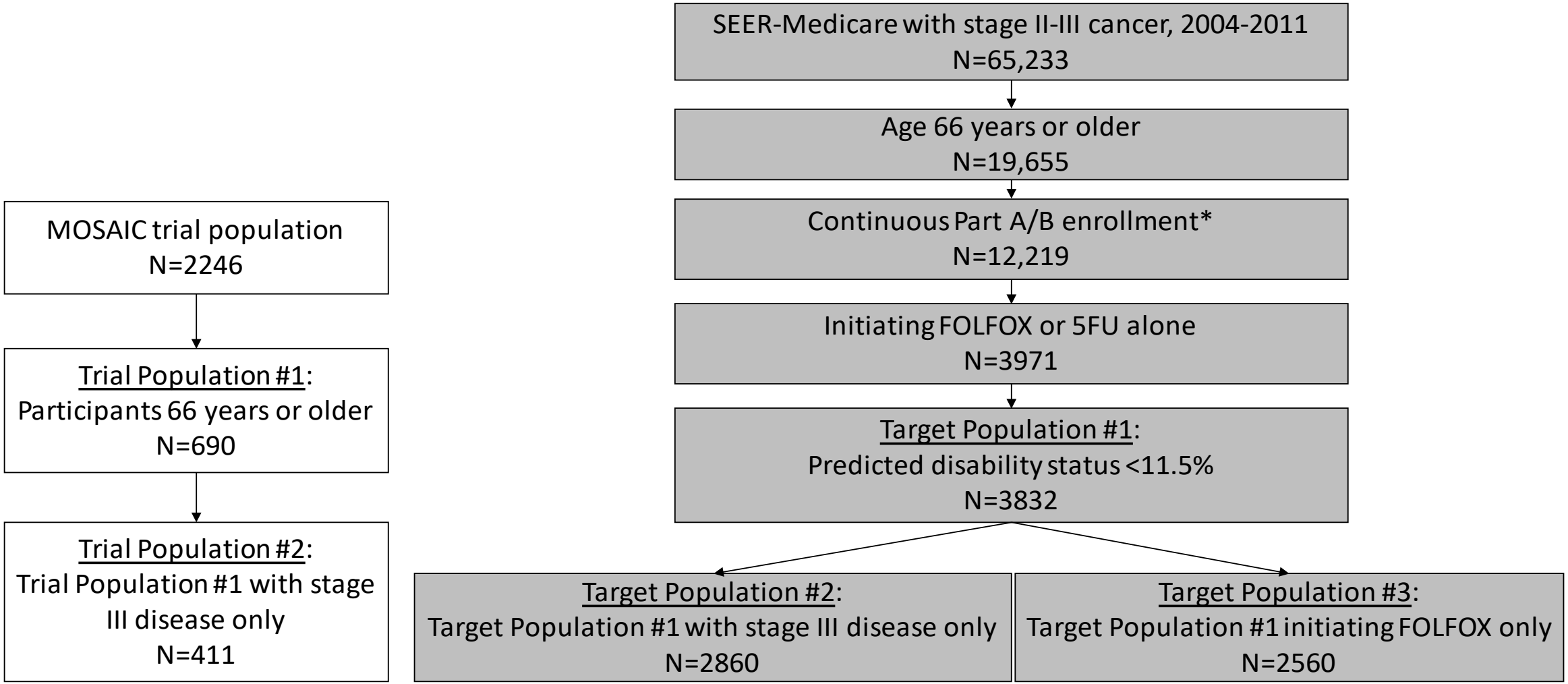
Study population: age 66-75, stage II-III colon cancer, claims-based proxy for good (ECOG 0, 1, 2) performance status, initiating FOLFOX or 5FU

Secondary analyses: (1) stage III patients and (2) patients initiating FOLFOX)

Methods and analysis

- Compared the two study populations with respect to the distribution of age, sex, and substage.
- Calculated the inverse odds of sampling weights
- Cox proportional hazards regression and Kaplan Meier methods to estimate the mortality hazard ratios and risk differences comparing FOLFOX vs. 5FU alone.
- Secondary analyses: (1) stage III patients only and (2) all patients initiating FOLFOX.

Study population flow diagram

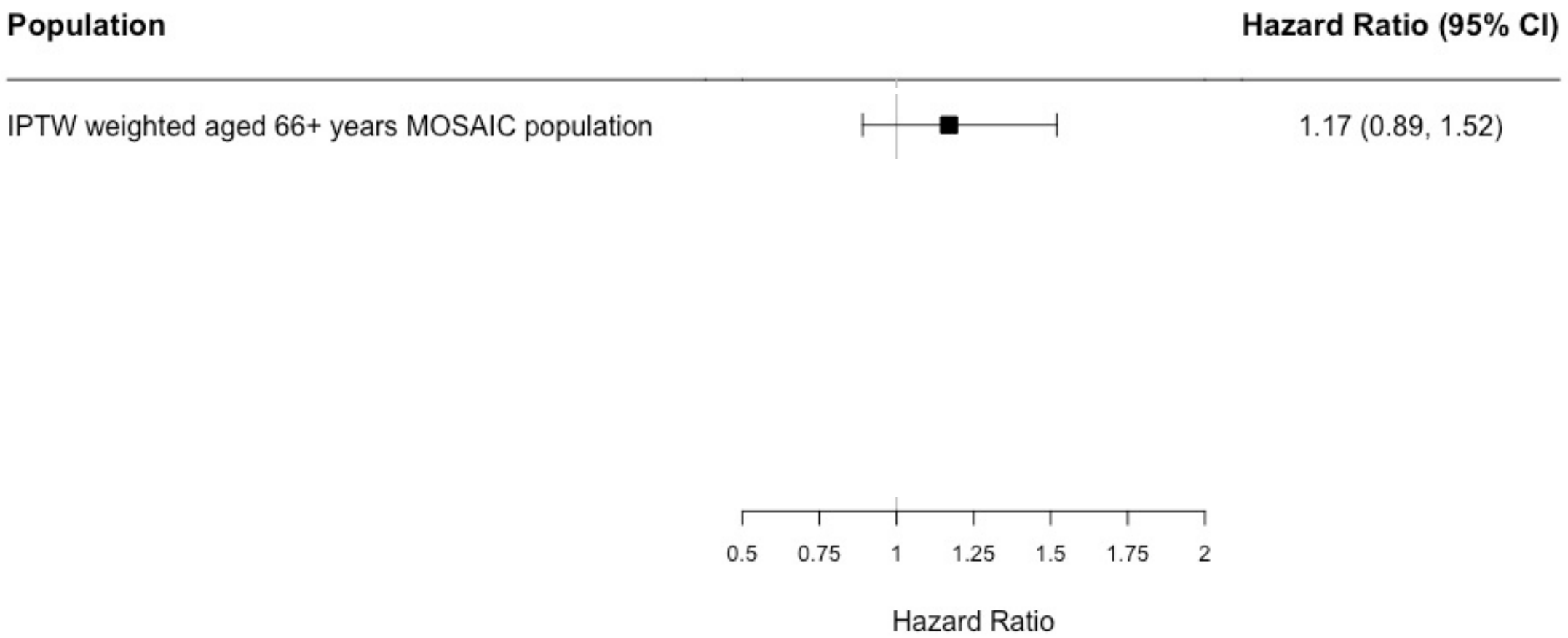


Results: Comparison of study populations

	MOSAIC Trial Population				SEER-Medicare Population		
	FOLFOX	5FU	Total		FOLFOX	5FU	Total
			n=690	%			
Age group							
66-69			375	54			
70-75			315	46			
Sex							
Male			373	54			
Female			317	46			
AJCC Substage							
IIA			232	34			
IIB			47	7			
IIIA			22	3			
IIIB			259	38			
IIIC			130	19			

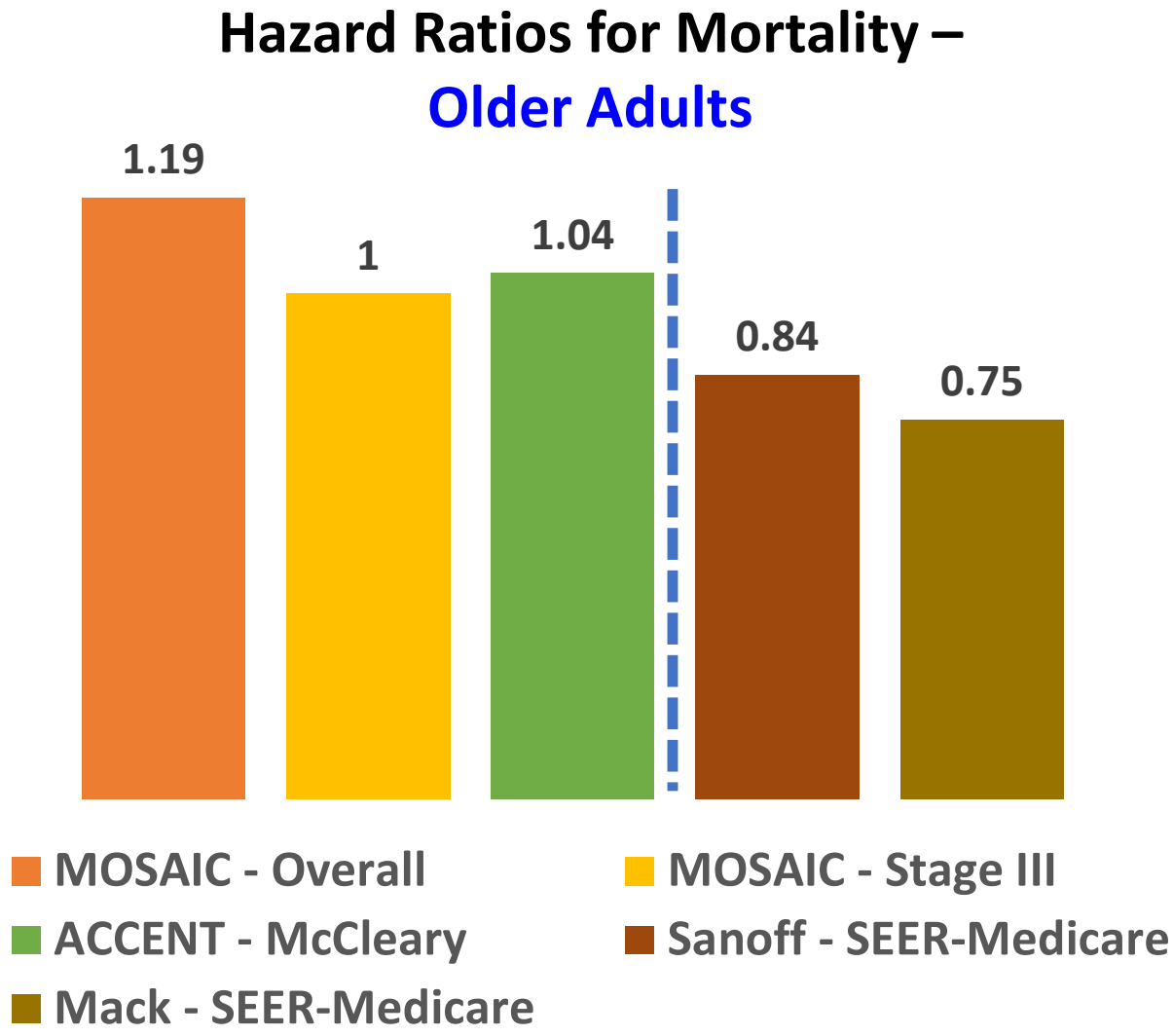
Abbreviations: MOSAIC=Multicenter International Study of Oxaliplatin/5FU-LV in the Adjuvant Treatment of Colon Cancer, SEER=Surveillance, Epidemiology and End Results program; FOLFOX=oxaliplatin +5FU; 5FU-5-fluorouracil; AJCC=American Joint Commission on Cancer.

Mortality hazard ratios comparing FOLFOX vs 5FU



Study 1: Conclusions and implications

- Results of this hybrid study are in line with the results of the MOSAIC trial and ACCENT pooled findings.
- Depending upon the target, our HR estimates ranged from 1.19-1.21.
- Hybrid study approaches reduce the risk of confounding by leveraging trial randomization.
- Limited in identifying the target due to lack of clinical data



Study 2: FOLFOX versus 5FU in community oncology practice

- Most patients diagnosed with cancer in the United States received their treatment in community oncology practices.

- As clinical trial enrollment is low overall and largely dominated by patients treated in academic medical centers, there is little evidence about the effectiveness of cancer therapies in the community setting.

- Extend our work in the older adult population by evaluating a wider population of adult colon cancer patients with better capture of clinical data through data from an oncology-specific electronic health records system.

What would the results of the MOSAIC trial have been if the trial population was randomly drawn from the US Oncology Network with stage II-III colon cancer?

Data sources and study population

MOSAIC Phase III Trial

Trial sponsor: Sanofi

Enrolled: 1998-2001 (146 centers, 20 countries)

Key eligibility: age 18-75, stage II-III colon cancer, KPS 60+, adequate organ function

US Oncology EHR: iKnowMed

US Oncology: Network of community oncology practices

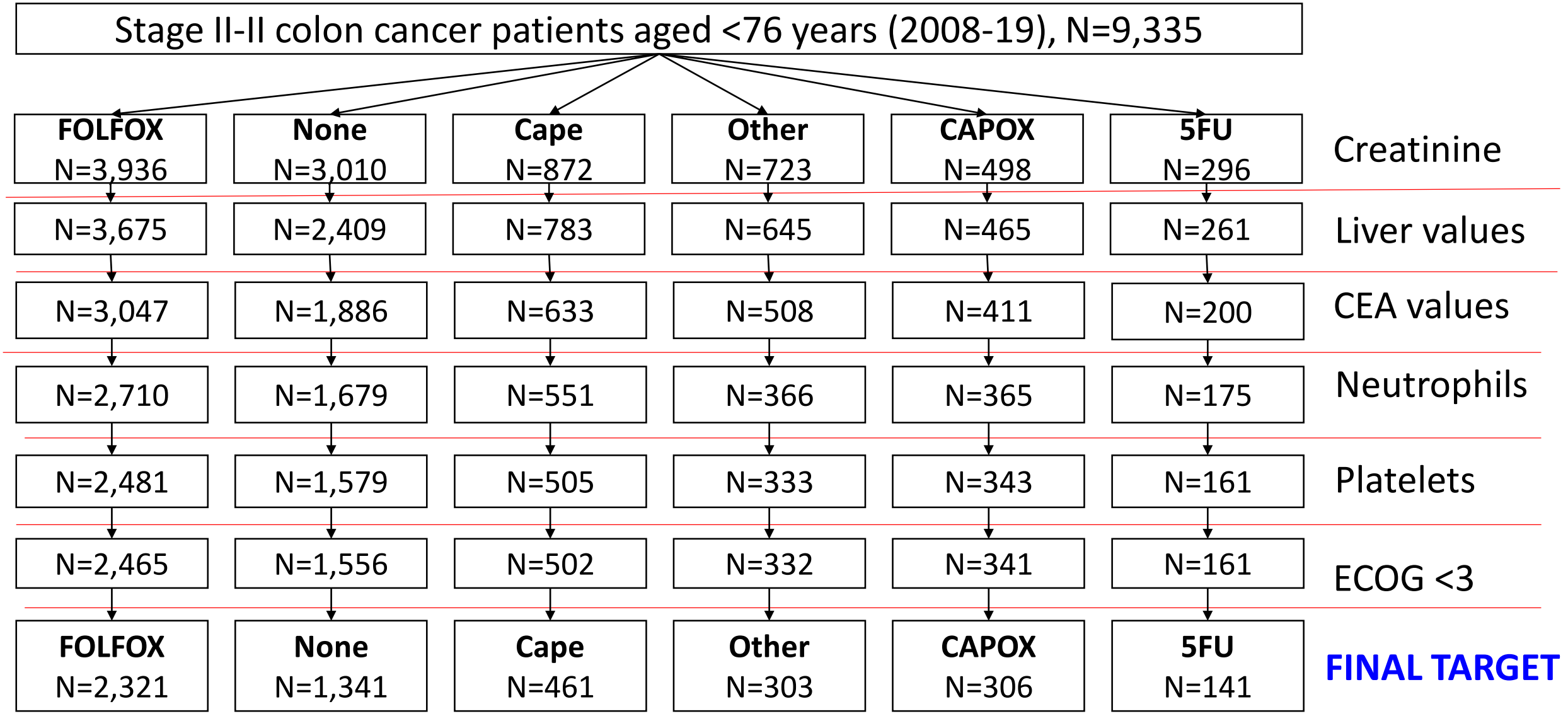
Data system: Oncology-specific, integrated, web-based EHR system, iKnowMed®

Study population: age 18-75, stage II-III colon cancer, KPS 60+, adequate organ function

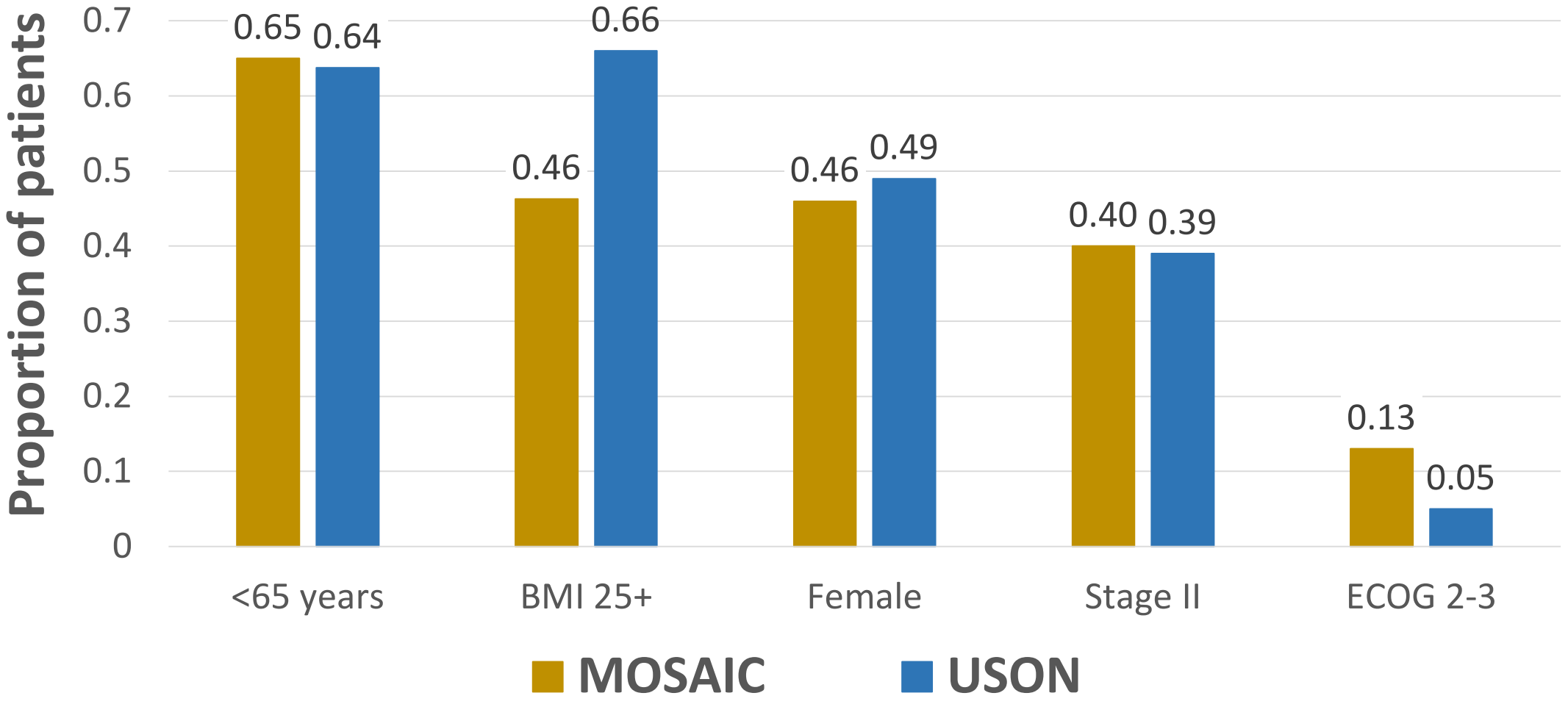
Methods and analysis

- Compared the two study populations with respect to the distribution of age, sex, and substage.
- Calculated the inverse odds of sampling weights and described their distribution.
- Cox proportional hazards regression to estimate hazard ratios for all cause mortality comparing FOLFOX versus 5FU.
- Secondary analyses: (1) stage III patients only and (2) stage III patients initiating FOLFOX.

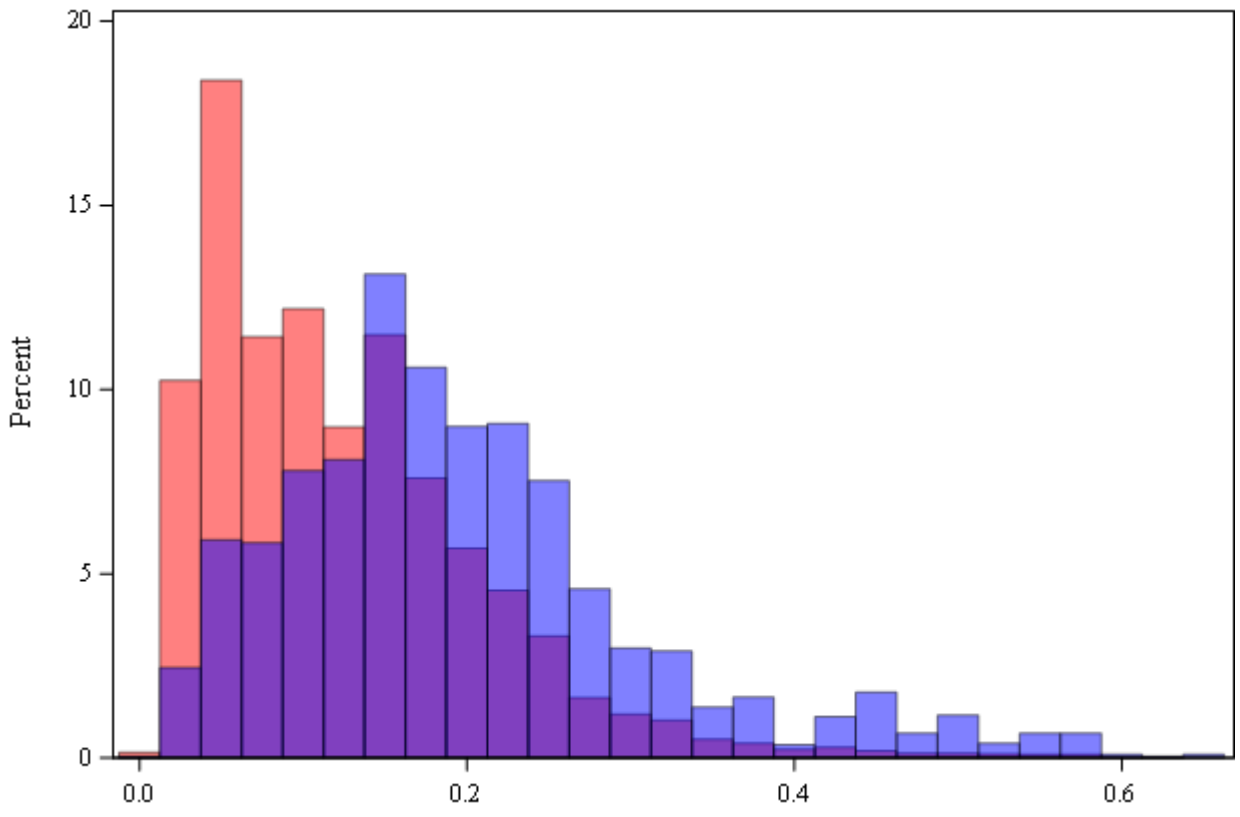
Study population flow diagram in US Oncology



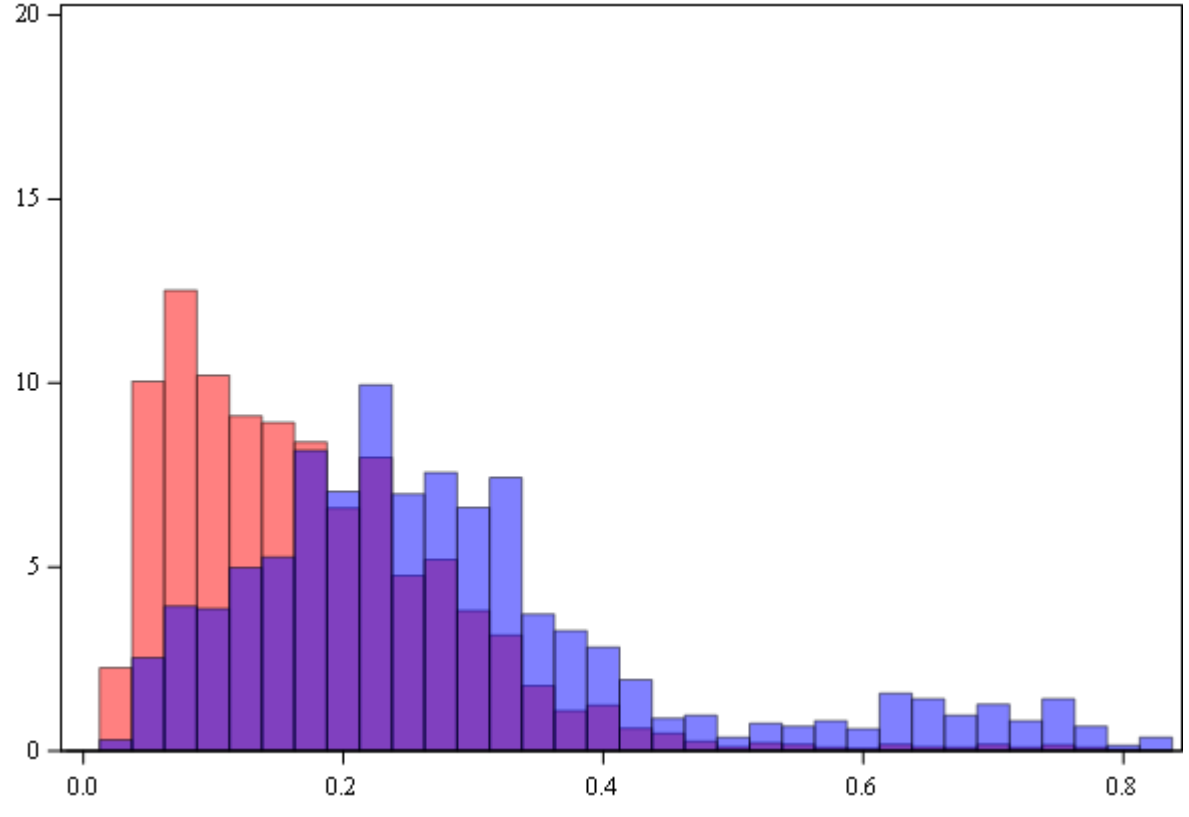
Results: Comparison of study populations



Probability of trial enrollment in the trial (blue) and USON (red)



Target Population:
All Stage II-III Colon Cancer Patients



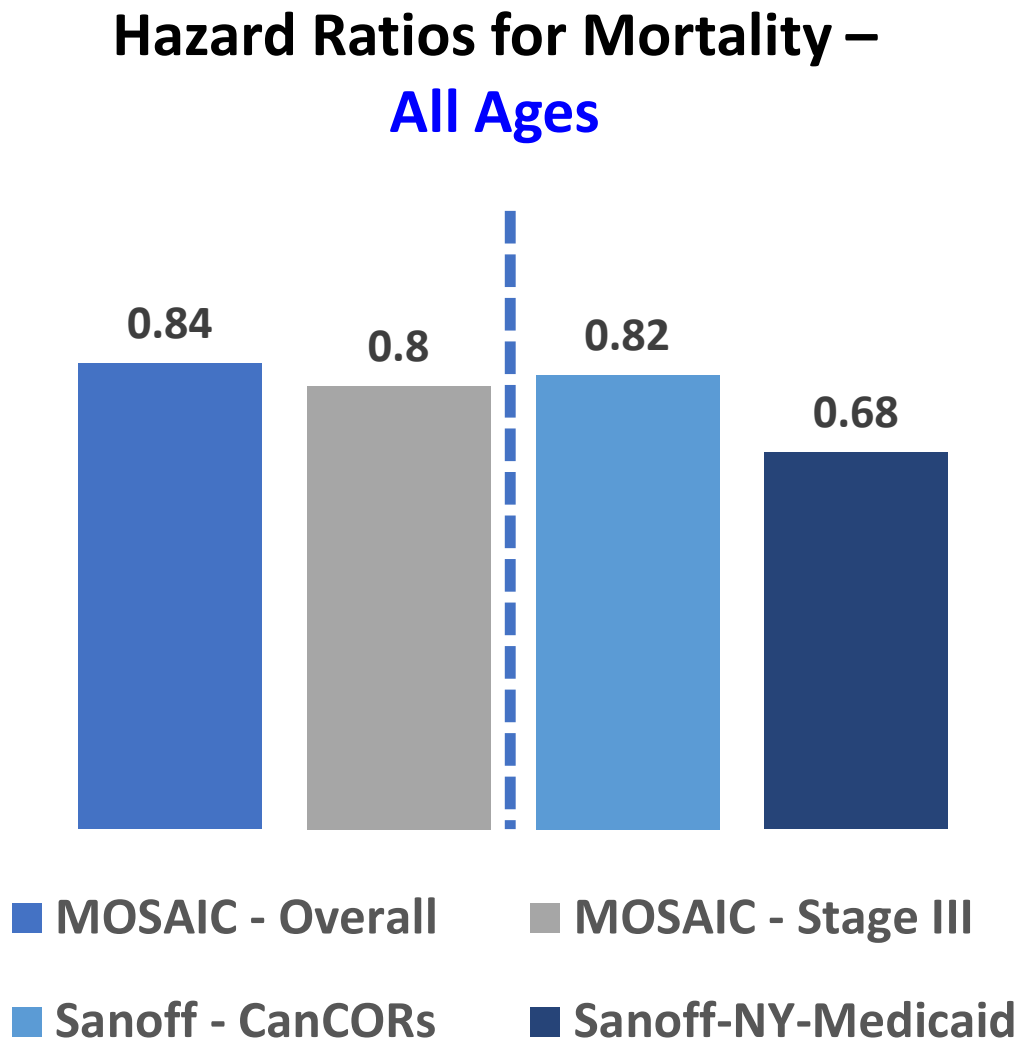
Target Population:
All FOLFOX-Treated Patients

Mortality hazard ratios comparing FOLFOX vs 5FU

Population	Trial Arm	6-Year Mortality	HR (95% CI)
MOSAIC population – all patients	5FU	24.0%	1.0
	FOLFOX	21.5%	0.84 (0.71, 1.00)
MOSAIC population – stage III patients	5FU	31.3%	1.0
	FOLFOX	27.1%	0.80 (0.65, 0.97)

Study 2: Conclusions and implications

- Results of this hybrid study indicate that the incremental effectiveness of FOLFOX versus 5FU are attenuated compared with what was observed in the trial.
- Depending upon the target population, our mortality HR estimates ranged from 0.93-1.05.
- Of note, confidence intervals are wide and include the overall MOSAIC trial result.



Clinical and research implications and next steps

Clinical implications

In general, our results show that the incremental benefits of FOLFOX versus 5FU are diminished in real-world, clinical practice settings

- FOLFOX does not appear to provide incremental reductions in mortality compared with 5FU in older adult populations
- Even in a broader population (age 18-75 years), benefits were largely diminished

Informative to a variety of stakeholders: regulators, payers, clinicians and patients – who want to know about the population-level benefits and harms of therapies in clinical practice

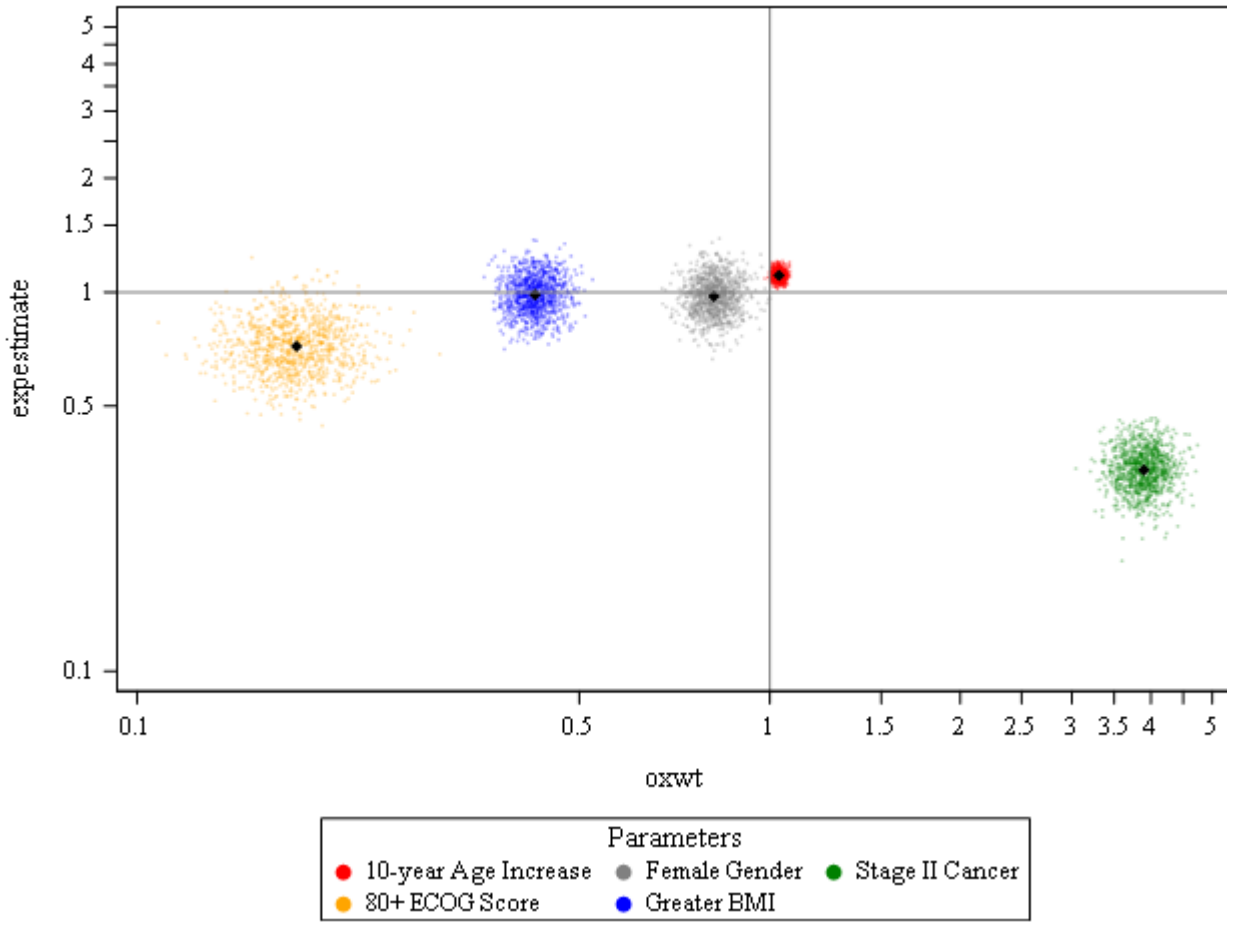
Research implications

- Proliferation of real-world data sources (e.g., EHRs) and increasing access to clinical trial data make hybrid designs increasingly feasible
  - Project Data Sphere, ClinicalStudyDataRequest, Vivli, YODA, etc.
- Application of these methods to the CER context is limited
- Diagnostic tools to guide model selection and identification of relevant and appropriate target populations for hybrid studies are needed.
- Consideration of differential therapy adherence between trial and clinical practice populations should also be considered in the hybrid study framework

Next steps (1): Diagnostic tools to guide model selection and identification of target populations

- How do we know what variables need to be included in our sampling models?

- Several target populations exist, how can we characterize differences between populations and identify targets of interest?



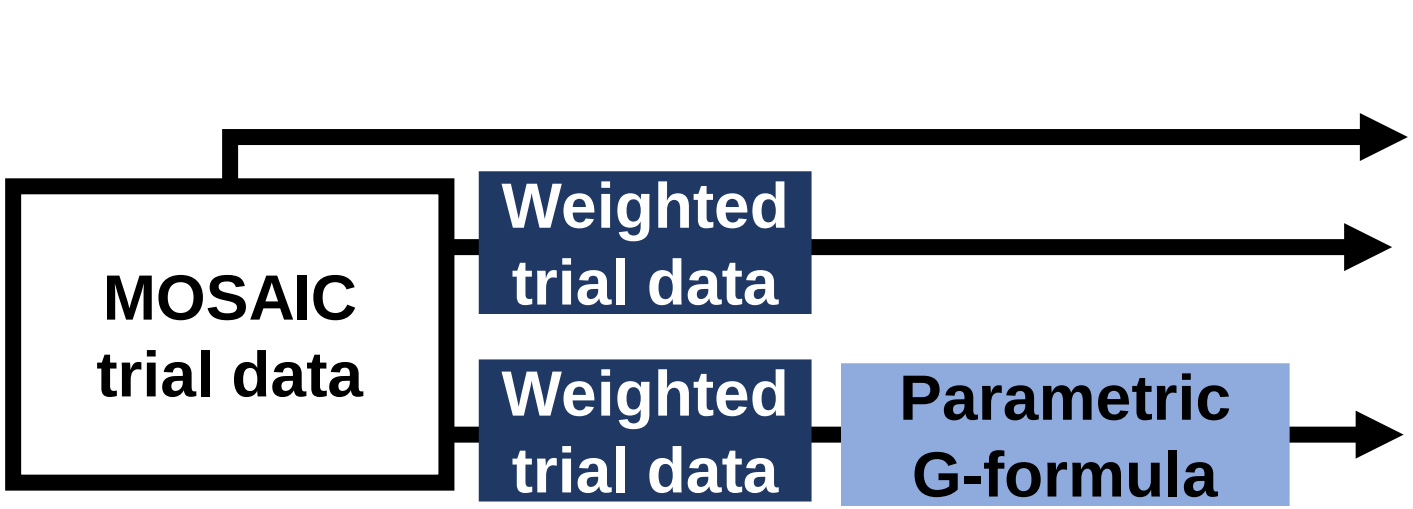
Next steps (2): Differential adherence

Clinical trials are generally designed to optimize treatment adherence

- Inclusion/exclusion criteria (e.g., run-in periods)
- Close monitoring via regular study visits

Examine patterns of FOLFOX and 5FU adherence in both MOSAIC and US Oncology populations.

Next steps (3): A flexible analytic approach for CER



Treatment effect of interest

(0) Trial ITT effect

(1) Weighted ITT effect

(2) Weighted per protocol (full adherence) effect

(3) Weighted treatment duration protocol effect

Key take-aways

- RCTs are the gold standard for evaluating efficacy of cancer therapies.
- Most stakeholders really want to know how effective cancer therapies are in patients like the ones they treat in clinical practice.
- Hybrid study approaches offer a novel way to leverage the internal validity of RCTs while also enhancing the generalizability of study results to relevant patient populations.
- Future work needs extend this framework and assess its relevance (e.g., RWE) for decision-making for patients, healthcare providers, payers, and regulators.

Questions and discussion

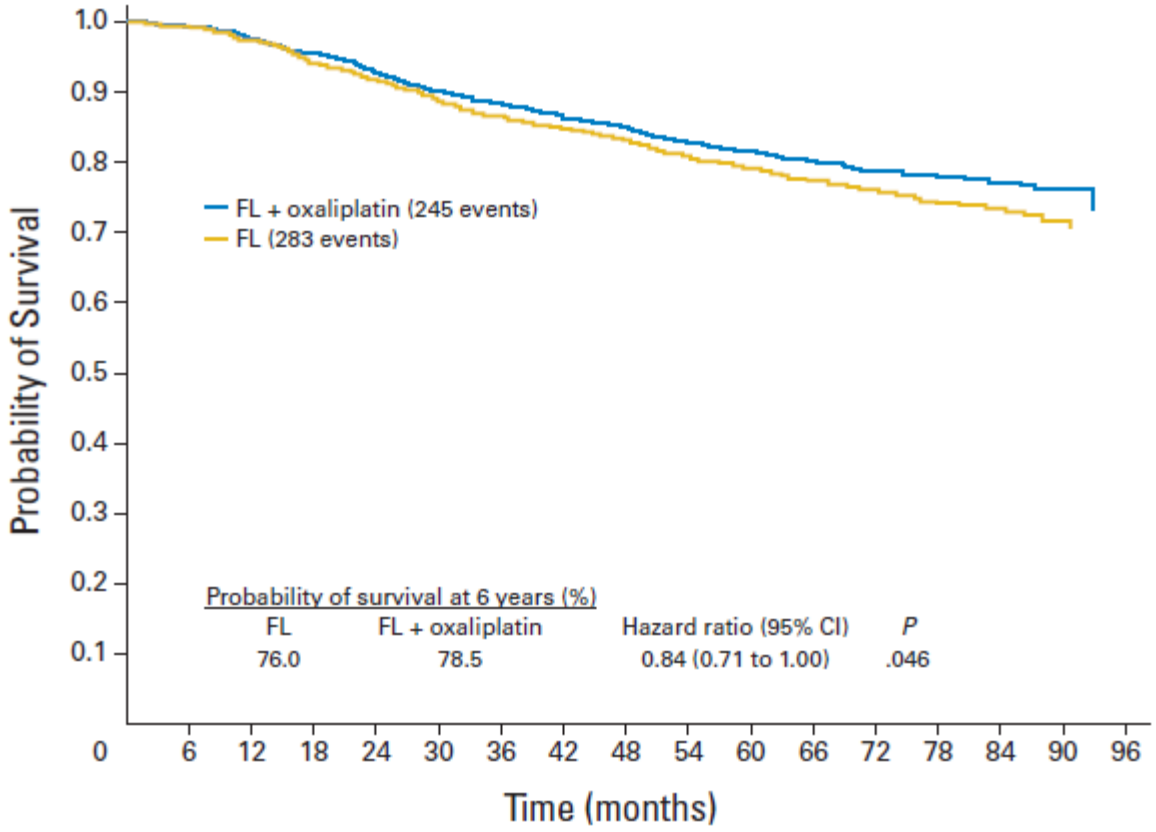
Contact information

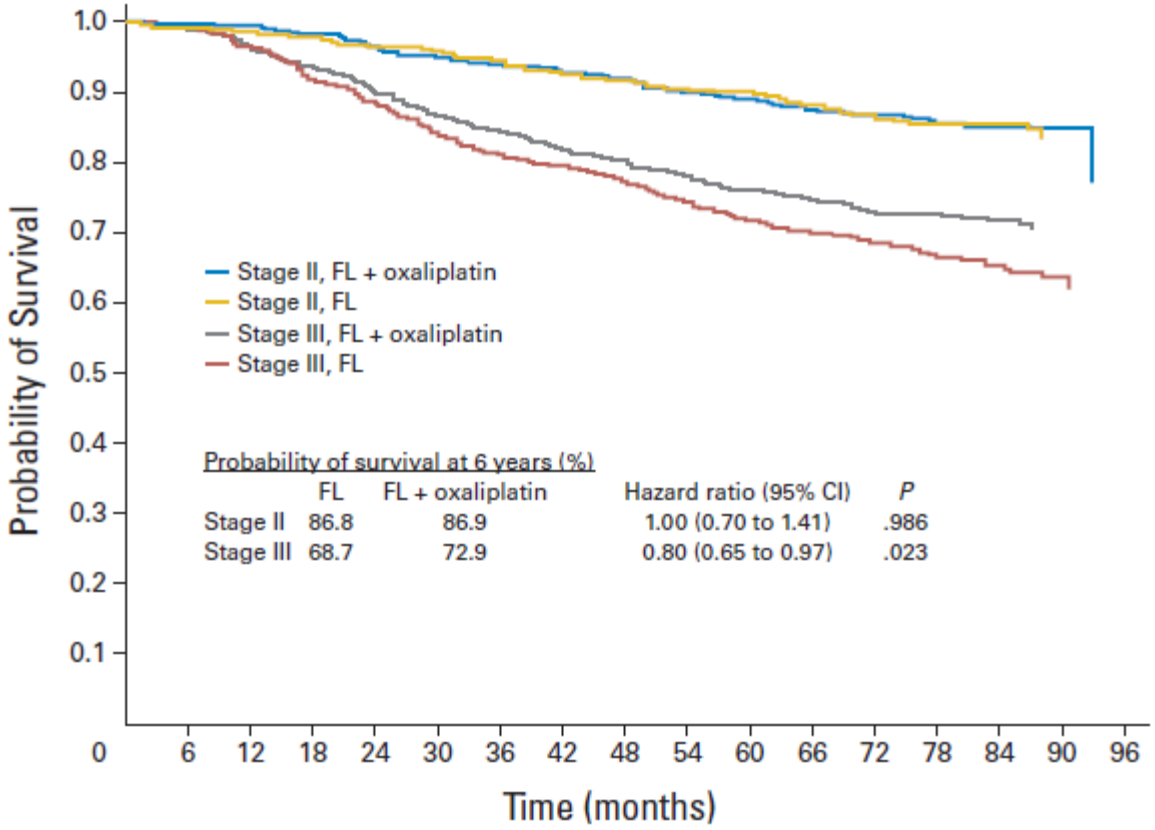
Jennifer L. Lund

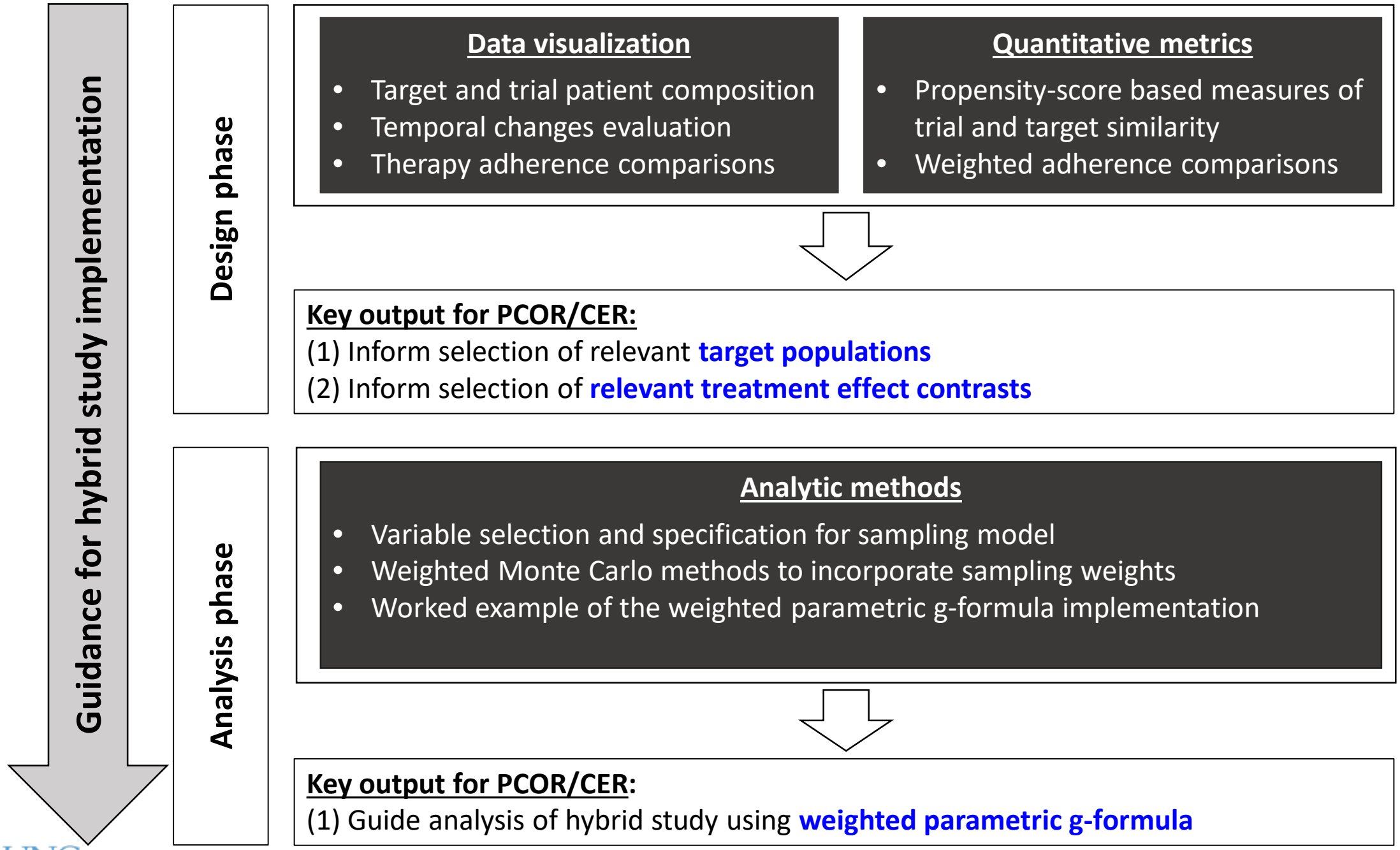
Email: Jennifer.Lund@unc.edu

Twitter: @LundJenny

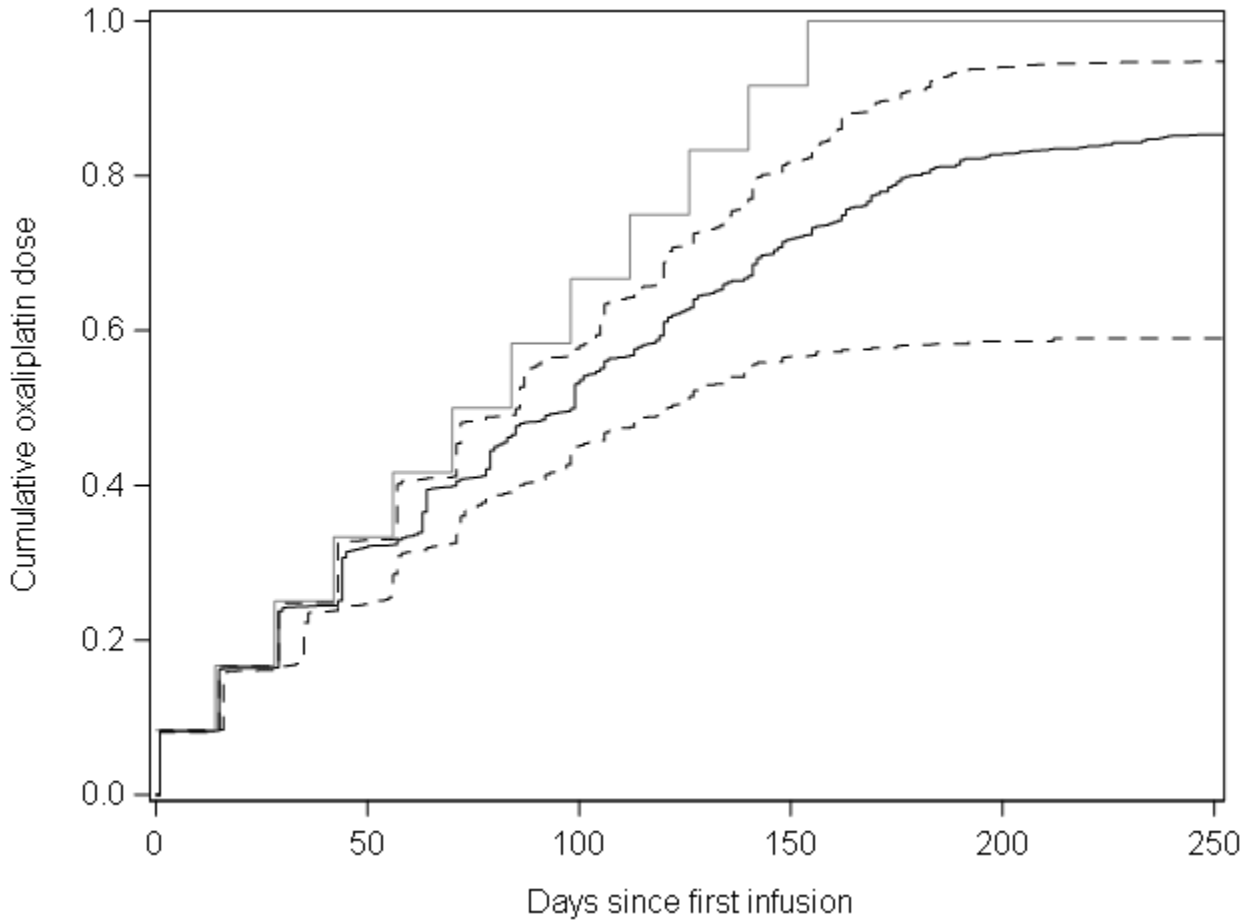
MOSAIC Trial Results






Guidance for hybrid study implementation
Design phase
Data visualization
• Target and trial patient composition
• Temporal changes evaluation
• Therapy adherence comparisons
Quantitative metrics
• Propensity-score based measures of trial and target similarity
• Weighted adherence comparisons
Key output for PCOR/CER:
(1) Inform selection of relevant target populations
(2) Inform selection of relevant treatment effect contrasts
Analysis phase
Analytic methods
• Variable selection and specification for sampling model
• Weighted Monte Carlo methods to incorporate sampling weights
• Worked example of the weighted parametric g-formula implementation
Key output for PCOR/CER:
(1) Guide analysis of hybrid study using weighted parametric g-formula

Oxaliplatin cumulative dose

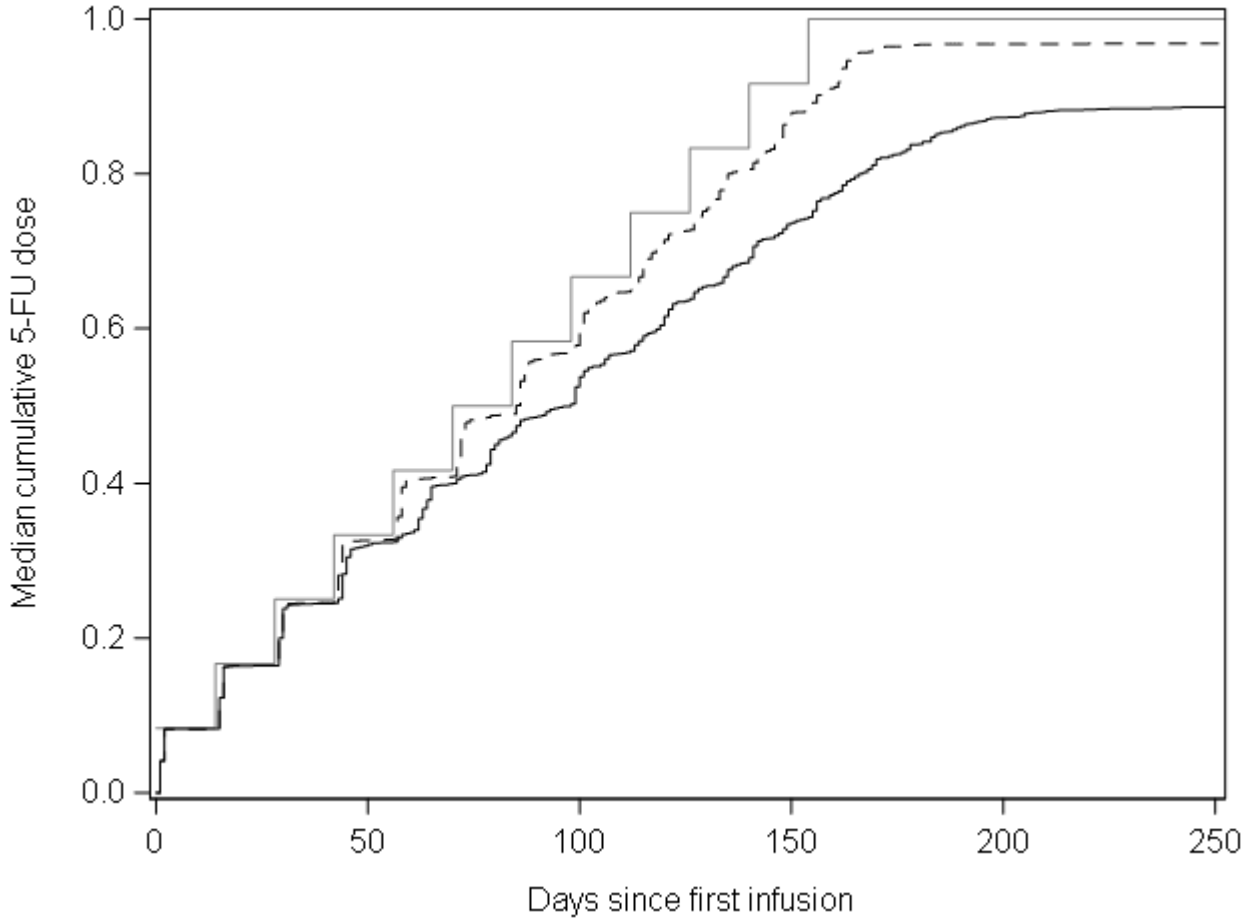


75th percentile

Median

25th percentile

5-FU cumulative dose (by arm)



5-FU arm

FOLFOX arm

Cumulative dose (by arm and age group)