An Introduction to Causal Inference for Pharmacometricians

Jim Rogers, PhD
May 18, 2023

METRUM
RESEARCH GROUP

CPT: Pharmacometrics & Systems Pharmacology

TUTORIAL | Open Access

An introduction to causal inference for pharmacometricians

James A. Rogers, Hugo Maas, Alejandro Pérez Pitarch

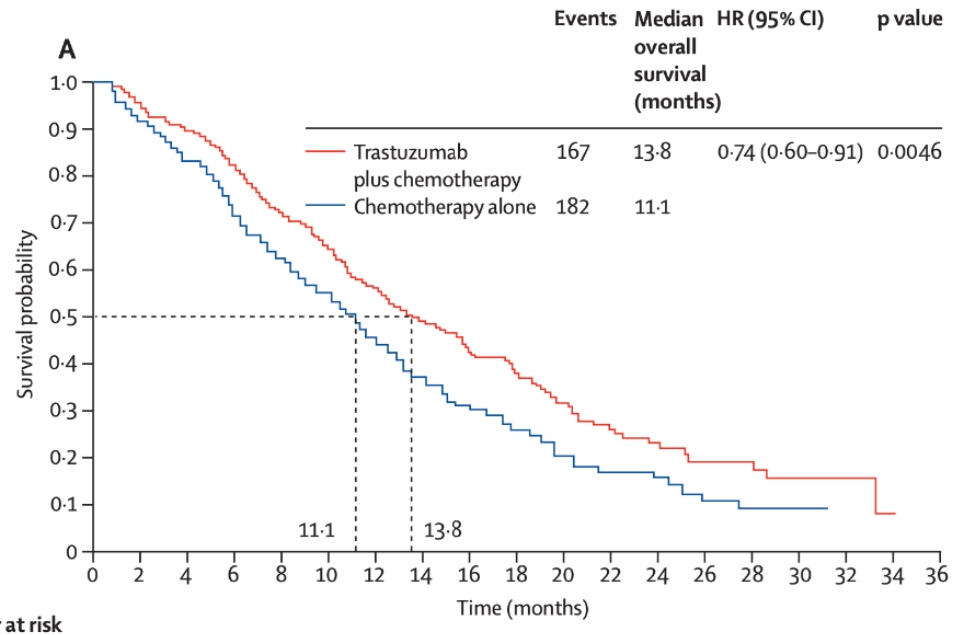
First published: 16 November 2022 | https://doi.org/10.1002/psp4.12894

Quantitative Data Integration to Enhance Clinical Trials
QUIC
A Global BDS-TMCP Initiative

Boehringer Ingelheim

METRUM RESEARCH GROUP

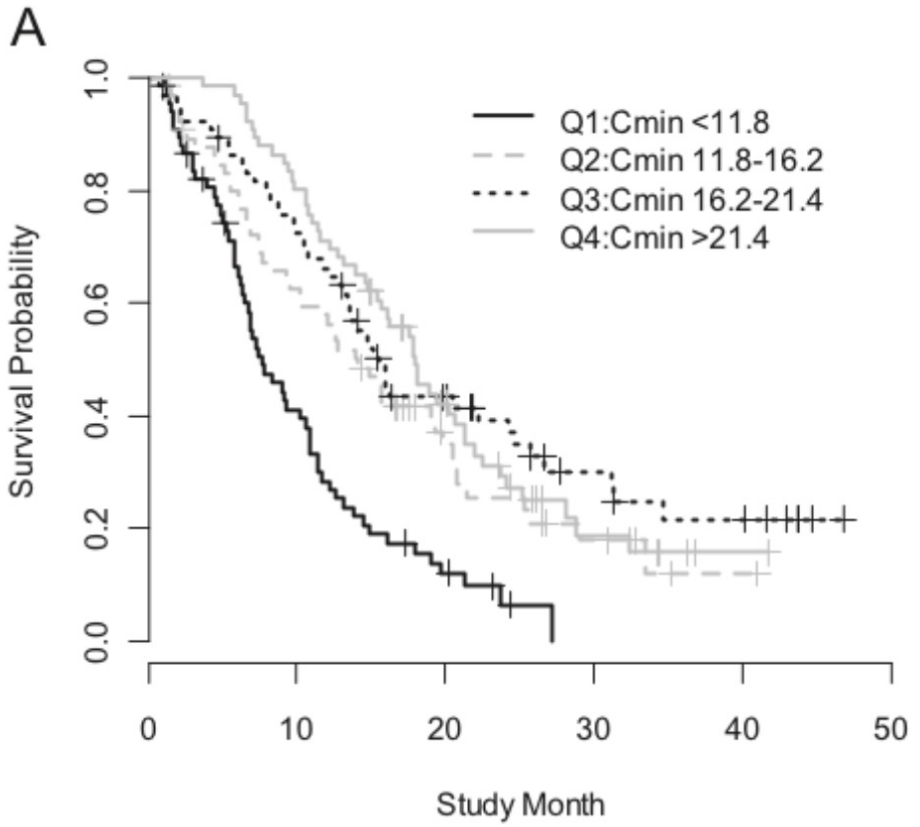
Trastuzumab in Gastric Cancer: ToGa



Clinical Trial > Lancet. 2010 Aug 28;376(9742):687-97. doi: 10.1016/S0140-6736(10)61121-X. Epub 2010 Aug 19.

Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial

Yung-Jue Bang [1], Eric Van Cutsem, Andrea Feyereislova, Hyun C Chung, Lin Shen, Akira Sawaki, Florian Lordick, Atsushi Ohtsu, Yasushi Omuro, Taroh Satoh, Giuseppe Aprile, Evgeny Kulikov, Julie Hill, Michaela Lehle, Josef Rüschoff, Yoon-Koo Kang; ToGA Trial Investigators



Pharmacokinetics and Pharmacodynamics

ACCP

The Journal of Clinical Pharmacology
53(2) 160–166
© The Author(s) 2012
DOI: 10.1177/0091270012445206

The Combination of Exposure-Response and Case-Control Analyses in Regulatory Decision Making

Jun Yang, PhD[1], Hong Zhao, PhD[1], Christine Garnett, PharmD[1], Atiqur Rahman, PhD[1], Jogarao V. Gobburu, PhD[1], William Pierce, PharmD[2], Genevieve Schechter, MD[2], Jeffery Summers, MD[2], Patricia Keegan, MD[2], Brian Booth, PhD[1], and Yaning Wang, PhD[1]

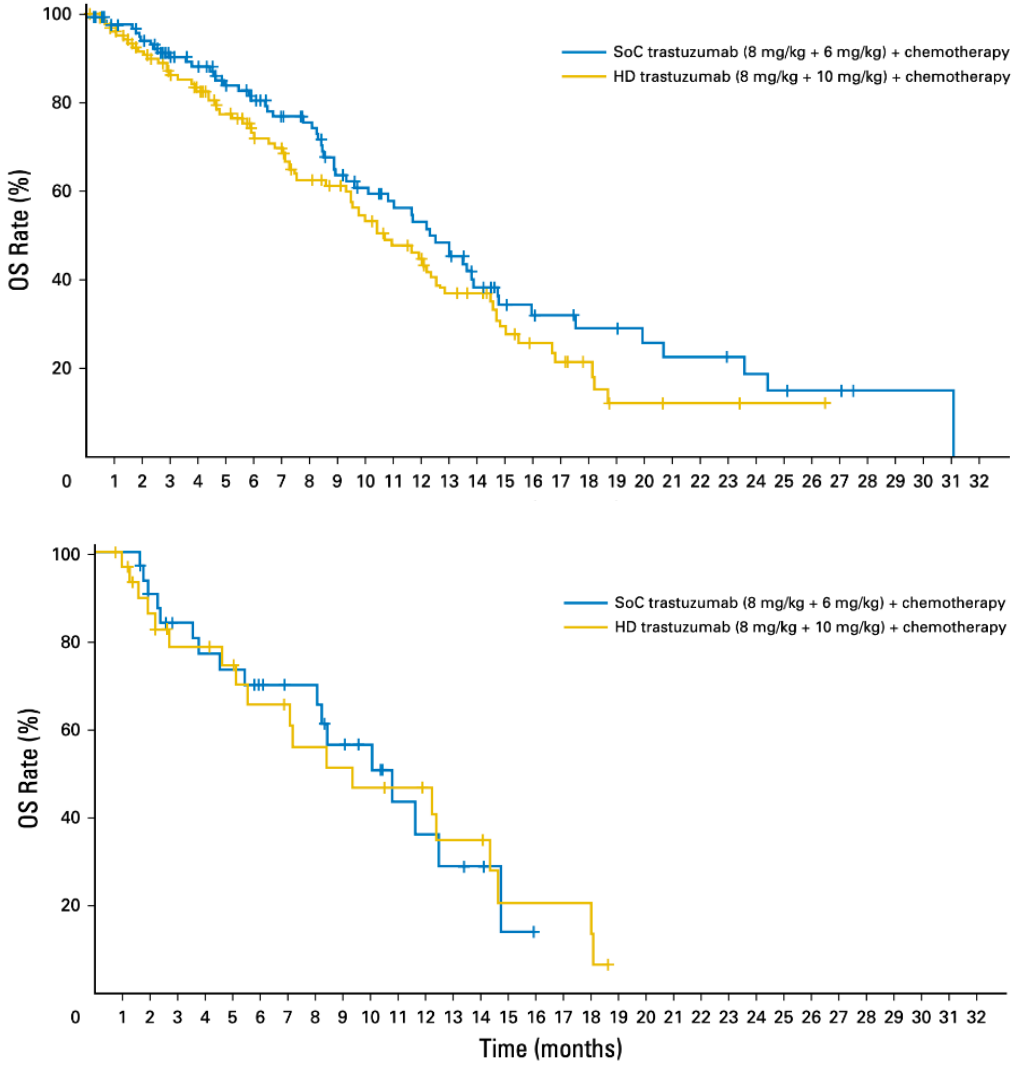
METRUM RESEARCH GROUP

Trastuzumab in Gastric Cancer: HELOISE

Clinical Trial > J Clin Oncol. 2017 Aug 1;35(22):2558-2567. doi: 10.1200/JCO.2016.71.6852. Epub 2017 Jun 2.

HELOISE: Phase IIIb Randomized Multicenter Study Comparing Standard-of-Care and Higher-Dose Trastuzumab Regimens Combined With Chemotherapy as First-Line Therapy in Patients With Human Epidermal Growth Factor Receptor 2-Positive Metastatic Gastric or Gastroesophageal Junction Adenocarcinoma

Manish A Shah [1], Rui-Hua Xu [1], Yung-Jue Bang [1], Paulo M Hoff [1], Tianshu Liu [1], Luis A Herráez-Baranda [1], Fan Xia [1], Amit Garg [1], Mona Shing [1], Josep Tabernero [1]



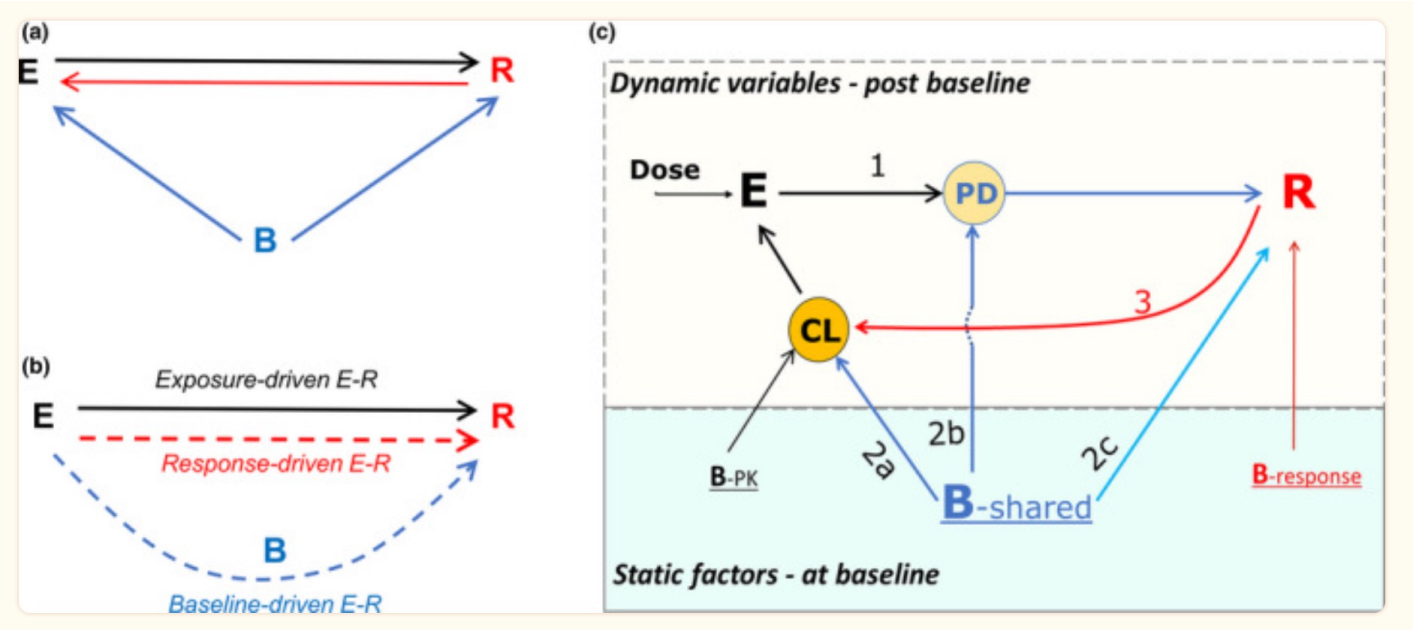
METRUM
RESEARCH GROUP

Trastuzumab: What Happened?

Clin Pharmacol Ther. 2020 Dec; 108(6): 1156–1170. PMCID: PMC7689749
Published online 2020 Aug 2. doi: 10.1002/cpt.1953 PMID: 32557643

Characterizing Exposure–Response Relationship for Therapeutic Monoclonal Antibodies in Immuno-Oncology and Beyond: Challenges, Perspectives, and Prospects

Haiqing Isaac Dai,[1,†] Yulia Vugmeyster,[1,†] and Naveen Mangal[1]



~~Low exposures caused larger tumors~~
Larger tumors caused low exposures

for efficacy and a confounding factor for the E–R relationship. A strong correlation was observed between baseline CL and patient survival in oncology studies.[4,5,7,8,23,26] Once CL was included as a covariate, the correlation between exposure and survival diminished. It appears that CL is closely related to the baseline disease condition, such as cancer-related cachexia.[8,27] It is conjectured that the elevated inflammation and proteolytic activity in cancer-related cachexia is responsible for the high CL.[27,28] Under

METRUM
RESEARCH GROUP

Review > J Pharmacokinet Pharmacodyn. 2023 Mar 4. doi: 10.1007/s10928-023-09850-2.
Online ahead of print.

A comprehensive regulatory and industry review of modeling and simulation practices in oncology clinical drug development

Ana Ruiz-Garcia [1] [2], Paul Baverel [3] [4], Dean Bottino [5], Michael Dolton [6], Yan Feng [7], Ignacio González-García [8], Jaeyeon Kim [9], Seth Robey [10], Indrajeet Singh [11], David Turner [6], Shu-Pei Wu [12], Donghua Yin [13], Di Zhou [14], Hao Zhu [15], Peter Bonate [16]

Affiliations + expand

PMID: 36870005 DOI: 10.1007/s10928-023-09850-2 Paperpile

Typically in E–R analyses, drug exposure is assumed to be the cause, and response to be the outcome. However, if disease progression or remission influences pharmacokinetic (PK) parameters over time, this interaction between treatment response and PK parameters could result in artificial E–R relationships. Anti-programmed death-1 (anti-PD1) immunotherapies nivolumab and pembrolizumab exhibited time-dependent pharmacokinetics and a correlation between drug clearance changes over time and survival rates [17,18,19]. In these situations, directly linking drug exposure at steady state to clinical outcomes in a single-dose trial may yield an over-steep E–R relationship, deviating from the true underlying relationship.

METRUM
RESEARCH GROUP

Unifying Conceptual Frameworks

What are some general framework to encourage the type of thinking shown in, e.g. Dai, Vugmeyster, Mangal (2020) , Ruiz-Garcia et al (2023) ?

Estimands (in the ICH E9 addendum sense)

Directed Acyclic Graphs (DAGs)

Target Trial Emulation

Pearl's "do-Calculus"

Neyman-Rubin Potential Outcomes Framework (*)

- dating to Jerzy Neyman in 1923. Happy 100th Birthday PO Notation!

METRUM RESEARCH GROUP

Unifying Conceptual Frameworks

What are some general framework to encourage the type of thinking shown in, e.g. Dai, Vugmeyster, Mangal (2020) , Ruiz-Garcia et al (2023) ?

Estimands (in the ICH E9 addendum sense)

Directed Acyclic Graphs (DAGs)

Target Trial Emulation

Pearl's "do-Calculus"

Neyman-Rubin Potential Outcomes Framework (*)

- dating to Jerzy Neyman in 1923. Happy 100th Birthday PO Notation!

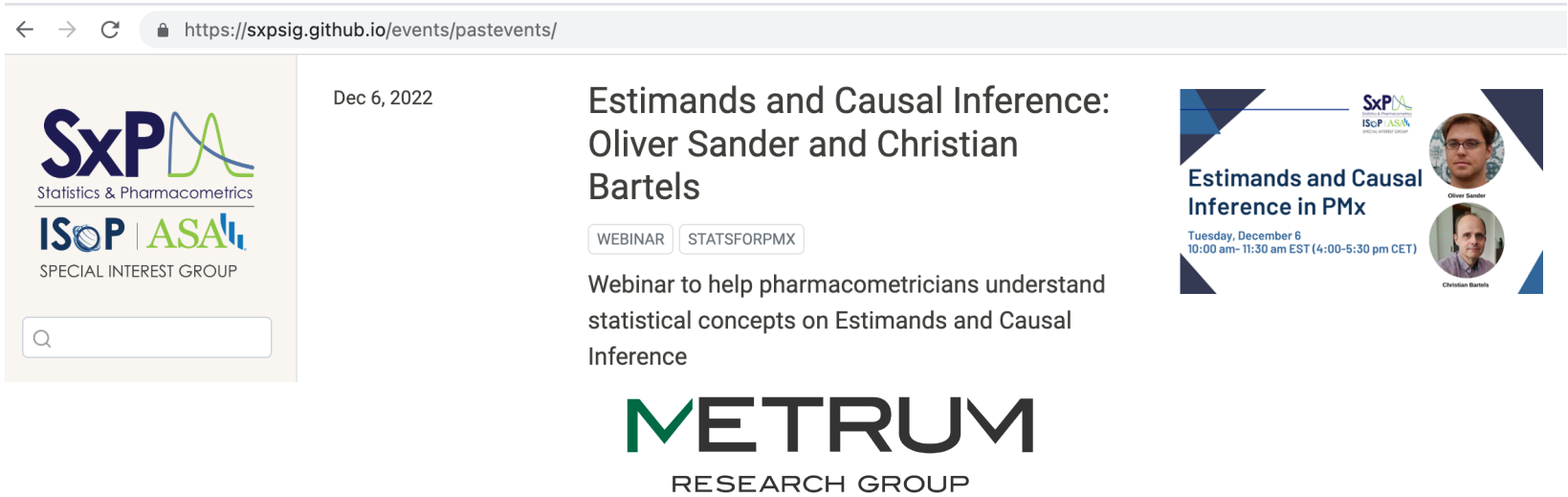
METRUM RESEARCH GROUP

Estimands as a Unifying Framework

Estimands—What they are and why they are important for pharmacometricians

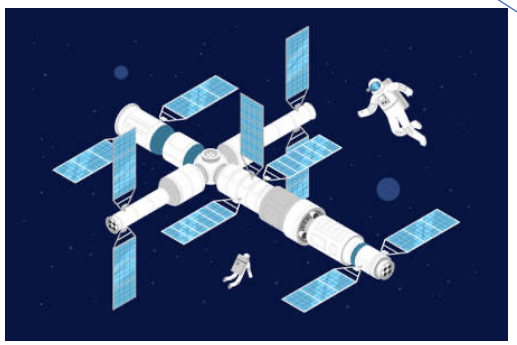
Mouna Akacha, [1] Christian Bartels, [1] Björn Bornkamp, [1] Frank Bretz, [1] Neva Coello, [1] Thomas Dumortier, [1] Michael Looby, [1] Oliver Sander, [1] Heinz Schmidli, [1] Jean-Louis Steimer, [1] and Camille Vong [1]

and assumptions. Although the ICH E9(R1) focuses on the causal effects of treatments, the same framework may also be used to evaluate estimands not related to treatment comparisons. The estimand framework can be understood as an attempt to restore "the intellectual primacy to the questions we ask, not the methods by which we answer them." [1] It is our firm belief that "opening the box" on the *what* questions in drug development through the estimand framework will ultimately lead to wider use of model-informed drug development methodologies [10] because they are often the most effective or the only way to address key questions such as method effectiveness.



Causal Inference as Part of "Pharmacometrics 101"

Analysis Plan



Objectives, Formalization of Questions, Scope of Work

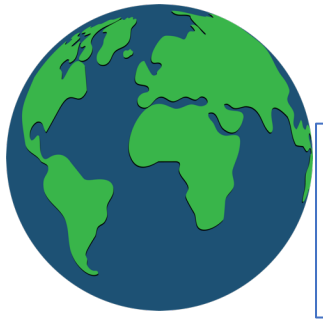
Code, Results (TFL)

Elicitation

Causal Inference especially valuable for "take off and landing"

Even for "simple" problems

Interpretation

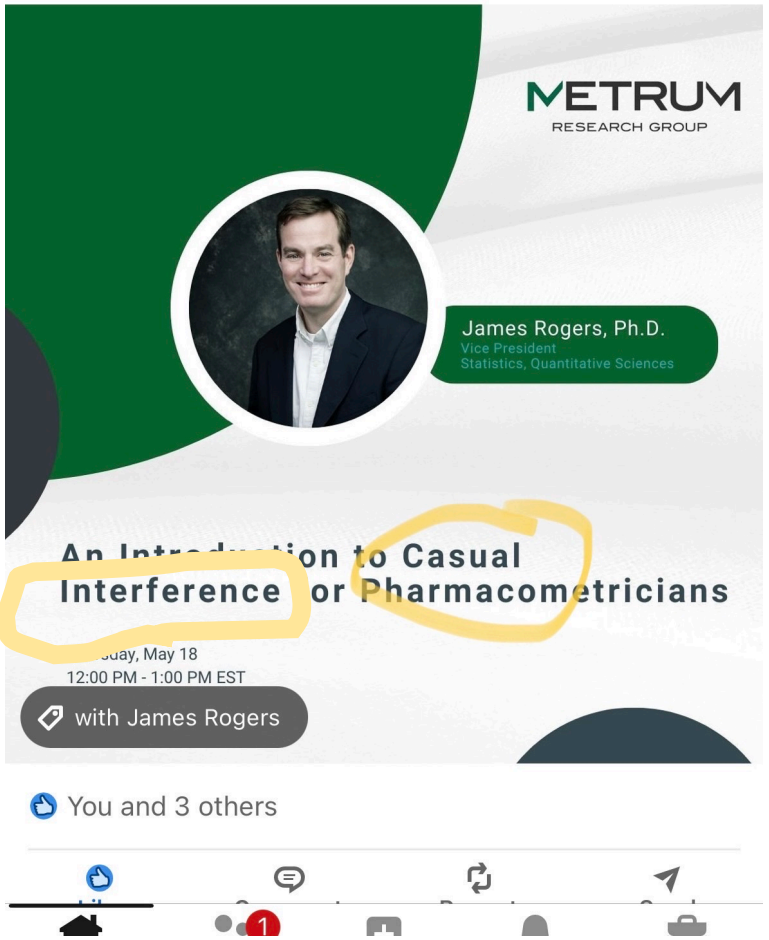


Vague inchoate questions and evidential needs + some data that might be useful

Discussion, Conclusions, Decisions

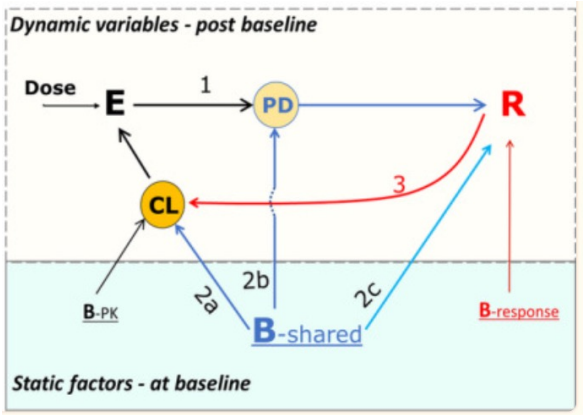
METRUM RESEARCH GROUP

The Casual Interference of Causal Inference

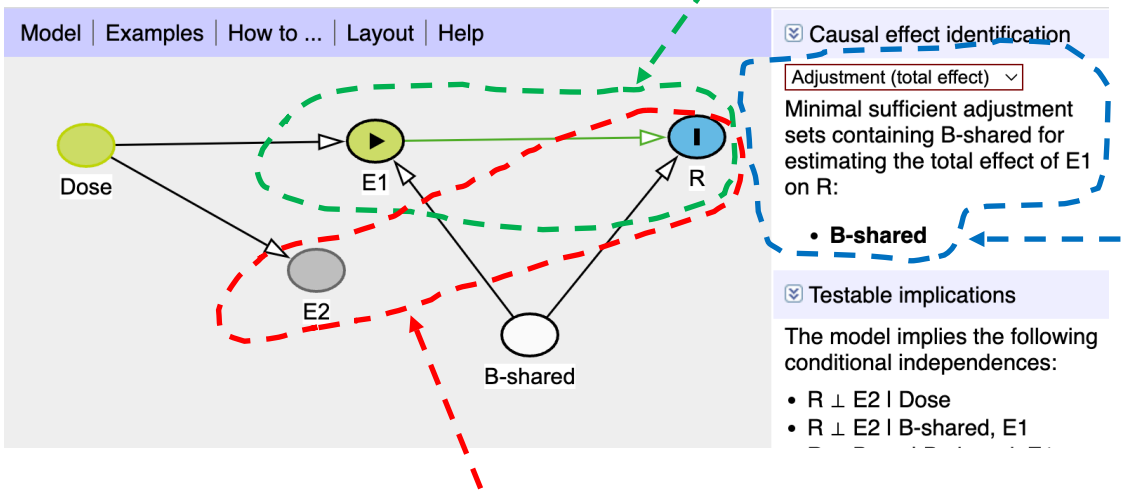


METRUM
RESEARCH GROUP

Using DAGitty to Clarify Questions, Assumptions, And Analysis Options



One of our questions: what is the effect of exposure feature E1 on response when other factors are held constant



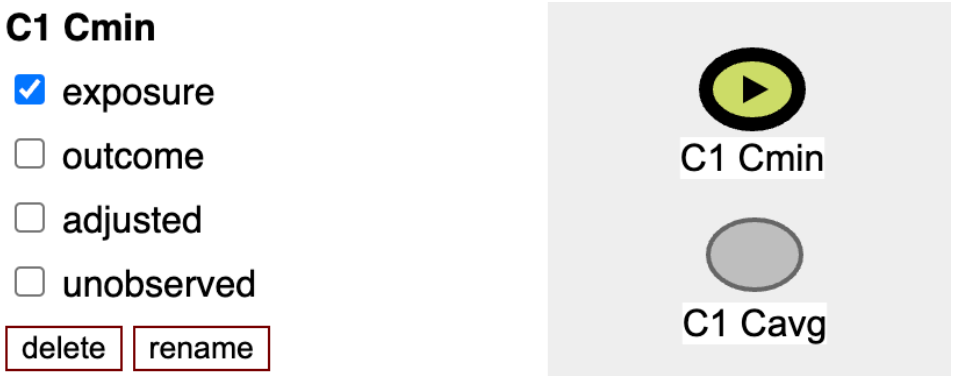
One of our analysis options, as determined by automated analysis of "backdoor paths"

One of our assumptions (indicated by the absence of an arrow): other exposure features such as E2 do not have an effect on the response

Johannes Textor, Benito van der Zander, Mark K. Gilthorpe, Maciej Liskiewicz, George T.H. Ellison. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *International Journal of Epidemiology* 45(6):1887-1894, 2016.

METRUM RESEARCH GROUP

The Discipline of DAG Building, Step 1: Identify the "Treatment" Variable



Thought exercise involving hypothetical randomization may seem unusual, but some of the best thinking in the pharmacometrics tradition has employed this conceptual framework:

> Because of variability in the relationship between dose and concentration (Dose/PK) among patients, the spread of concentrations within a dose group may be large, and concentration distributions may even overlap between dose groups. Patients who achieve a given concentration in a dose-controlled trial may have certain characteristics that lead to a different expected response from the expected response of patients from the same population randomly assigned to that concentration. We then say that the 'assignment' of concentrations to patients is confounded. The conventional PK/PD relationship will then differ from the true PK/PD relationship. Such a difference is due to the confounded assignment of concentrations, and we also say that the conventional PK/PD relationship is a confounded estimate of the true relationship. This paper concerns when such confounding can reasonably be assumed to be absent.

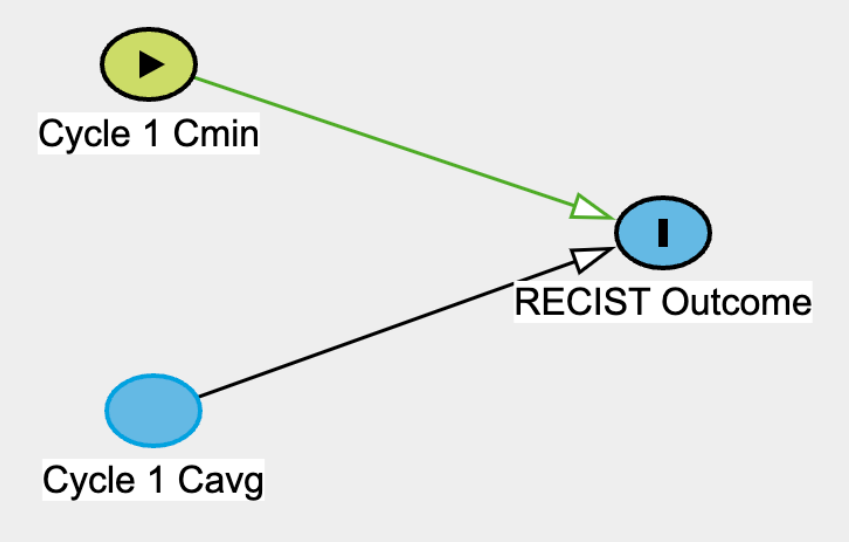
Nedelman, Rubin, Sheiner (2007)

Target Trial Emulation thought exercise:

- In principle, could we imagine randomizing patients to different levels of Cycle 1 Cmin?
- Is the treatment defined with sufficient specificity that we could describe it in a trial protocol?
- How would we make these assignments while holding other relevant factors constant?

METRUM RESEARCH GROUP

The Discipline of DAG Building, Step 2: Identify the "Outcome" Variable



Target Trial Emulation thought exercise:

Is the outcome defined with sufficient specificity that we could describe it in a trial protocol?

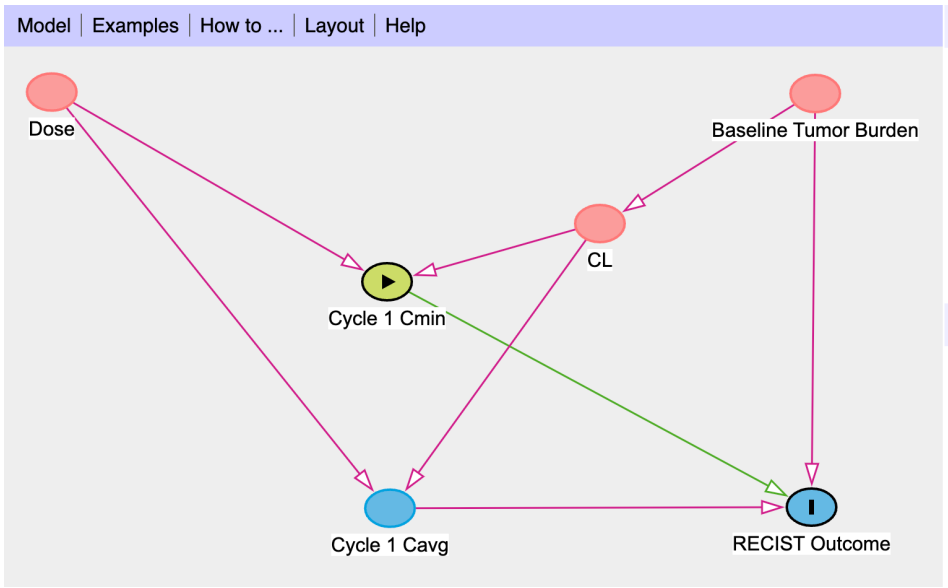
METRUM
RESEARCH GROUP

The Discipline of DAG Building, Step 3: Identify Selection Variables / Processes

Let’s gloss over this for now. More to come …

METRUM
RESEARCH GROUP

The Discipline of DAG Building Step > 3 : Identify Likely Common Causes; Repeat



Causal effect identification

Adjustment (total effect)

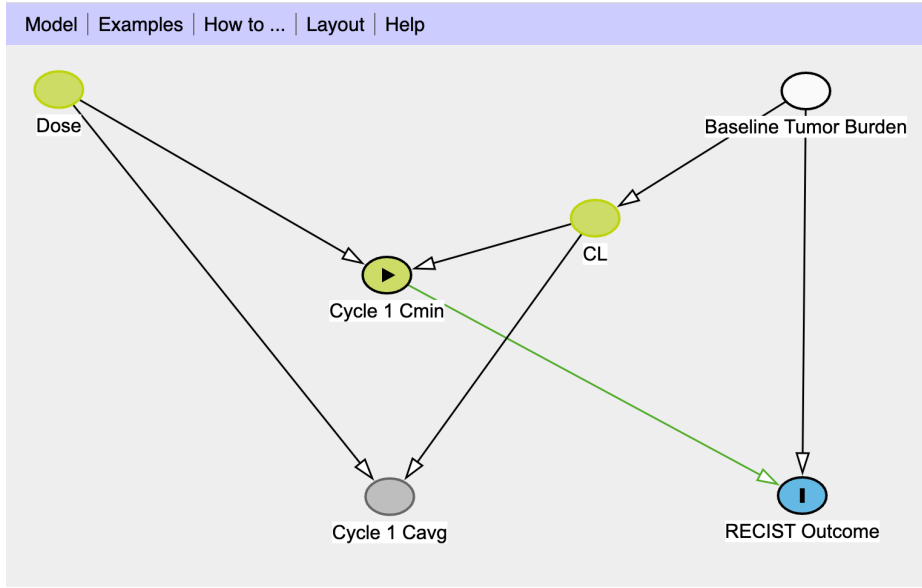
Minimal sufficient adjustment sets for estimating the total effect of Cycle 1 Cmin on RECIST Outcome:

- Baseline Tumor Burden, Cycle 1 Cavg
- CL, Cycle 1 Cavg
- CL, Dose

Testable implications

The model implies the following conditional independences:

- RECIST Outcome ⊥ Dose | CL, Cycle 1 Cavg, Cycle 1 Cmin
- RECIST Outcome ⊥ Dose | Baseline Tumor Burden, Cycle 1 Cavg, Cycle 1 Cmin
- RECIST Outcome ⊥ CL | Baseline Tumor Burden



Causal effect identification

Adjustment (total effect)

Minimal sufficient adjustment sets containing Baseline Tumor Burden for estimating the total effect of Cycle 1 Cmin on RECIST Outcome:

- **Baseline Tumor Burden**

Testable implications

The model implies the following conditional independences:

- RECIST Outcome ⊥ Dose | CL, Cycle 1 Cmin
- RECIST Outcome ⊥ Dose | Baseline Tumor Burden, Cycle 1 Cmin
- RECIST Outcome ⊥ Cycle 1 Cavg | CL, Dose
- RECIST Outcome ⊥ Cycle 1 Cavg | CL, Cycle 1 Cmin
- RECIST Outcome ⊥ Cycle 1

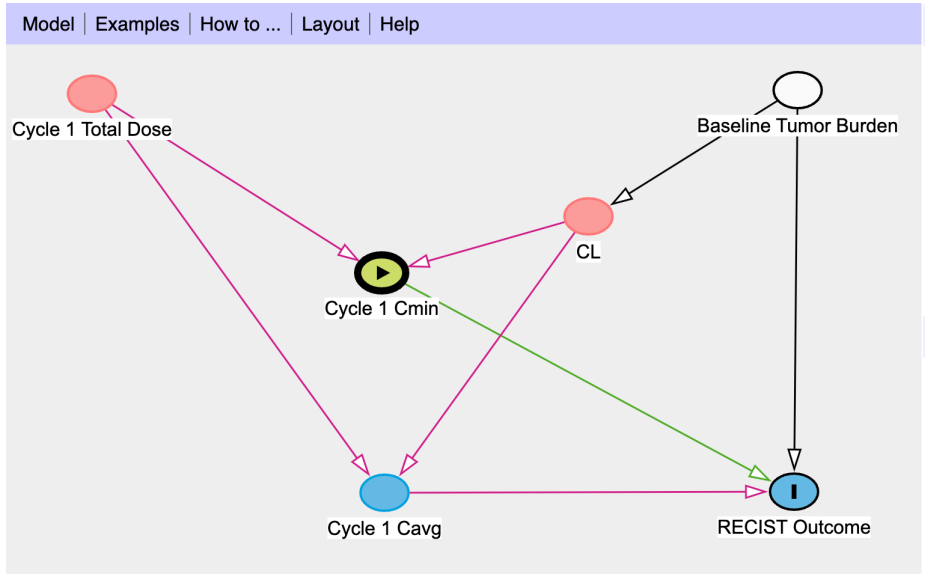
Questions that this prompts at both the planning and interpretation stages:

- Are we comfortable assuming that (almost) all of the effect of dose is mediated through Cycle 1 Cmin?
- Are we collecting the data and planning the analyses to adequately adjust for baseline tumor burden?
- Do we really care about isolating the effect of Cycle 1 Cmin? (As opposed to total effect of dose?)

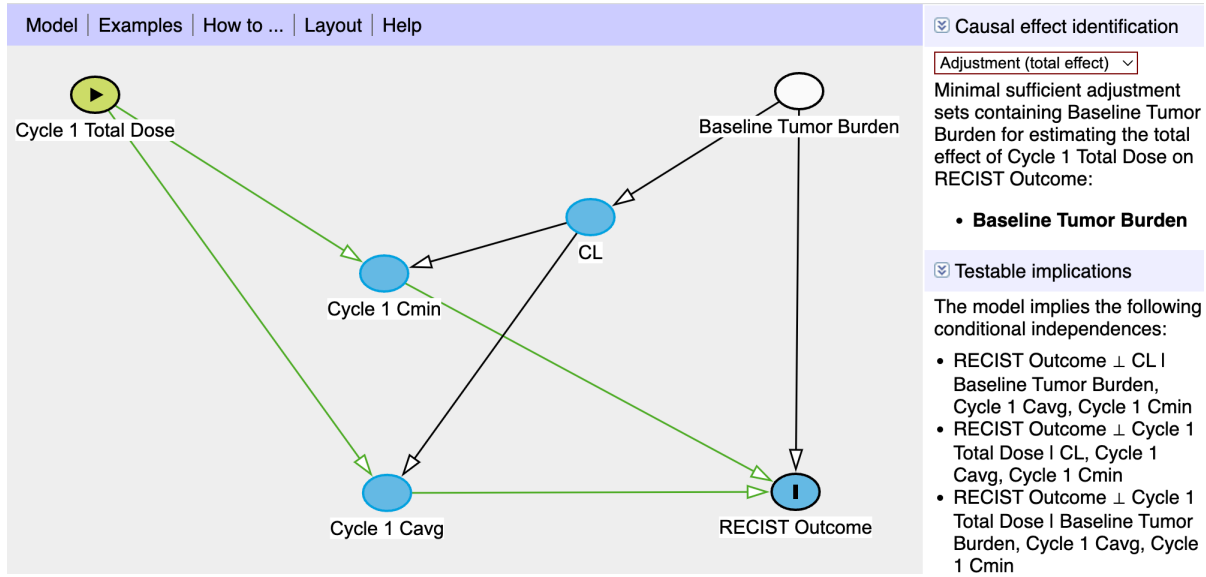
These questions **<u>are</u>** the (proximate) goal. Better questions → better planning and interpretation

METRUM
RESEARCH GROUP

Maybe it Suffices to Answer a Different (Simpler) Question?



Cycle 1 Cmin as the cause of primary interest



Cycle 1 Total Dose as the cause of primary interest

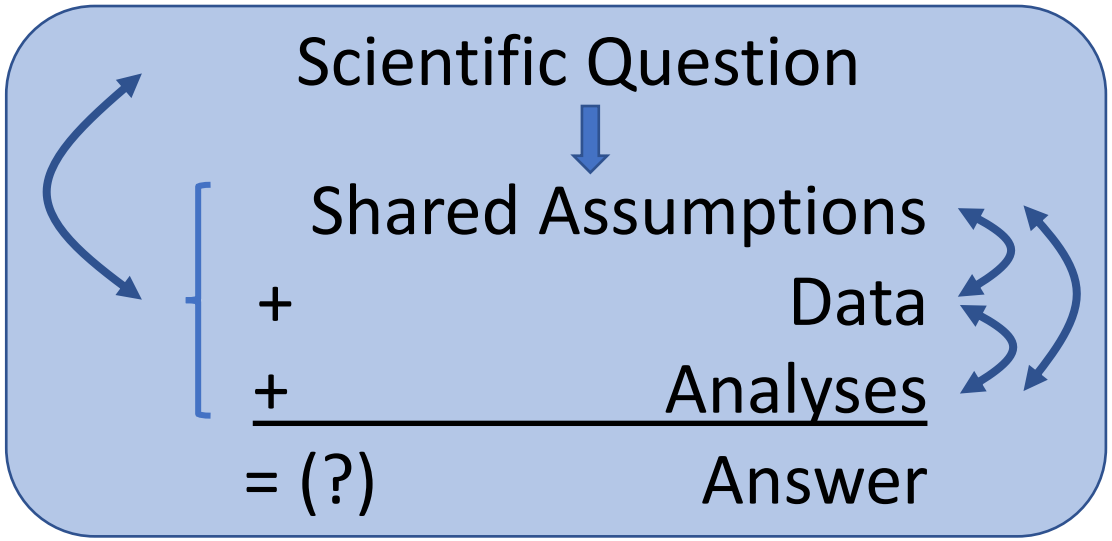
NB: It might be sufficient to answer the dose-response question in this particular case, but there are many situations where you would want to isolate the effects Cmin, Cavg, etc!

METRUM RESEARCH GROUP

Value Prop: Consistent Framework to Ensure Good Plans & Valid Interpretations

Would it suffice to answer a more limited question?

Can we make simplifying (but still defensible) assumptions?



Can we generate or obtain more data?

Can we adjust for certain biases in our analysis?

At every stage of iteration in our analysis planning, causal DAGs help to answer the question: **"Will this give us the type of answer we need?"**

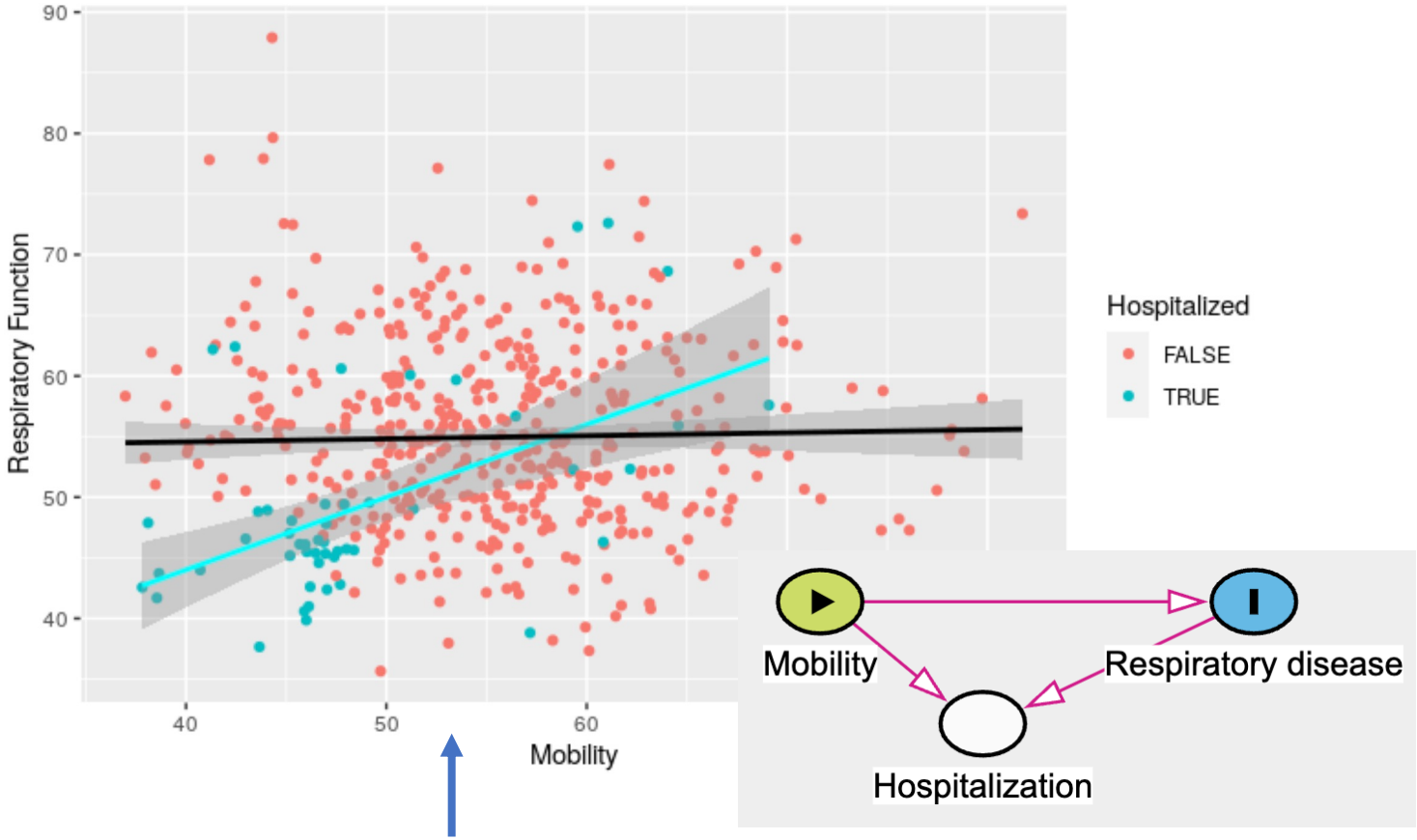
The next few slides provide a sampling of other applications where DAGs can quickly help to clarify our thinking:

- Retrospective study design (e.g. with real-world data)
- Covariate screening / covariate interpretation in pop PK
- Choice of exposure metric in exposure-response
- … the list goes on!

METRUM
RESEARCH GROUP

Selection and Collider Bias

Sackett DL. Bias in analytic research. J Chron Dis 1979; 32: 51-63



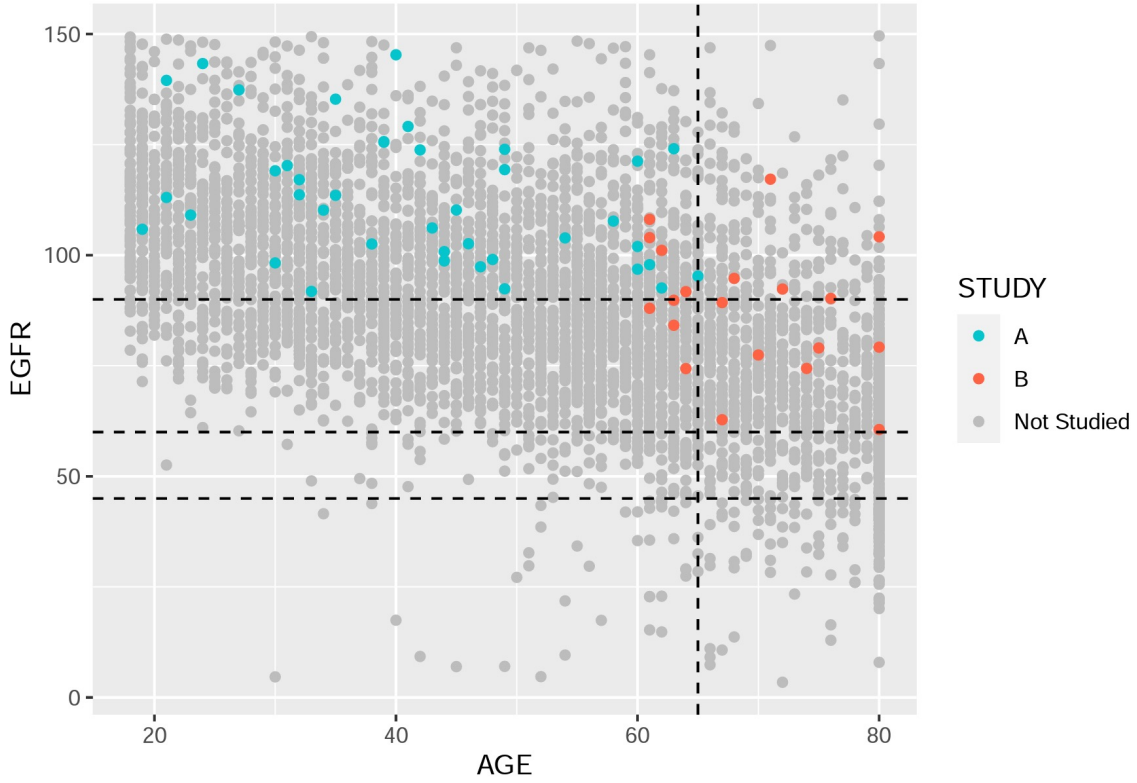
(this is not the real data analyzed by Sackett, but the real data followed a similar pattern)

- Hospital patients with locomotor disease appear more likely to have have respiratory disease.
- It is physiologically plausible that lack of mobility could cause respiratory impairment …
- **… but, not so fast …**
- Analyzing only hospitalized individuals creates false association (cyan regression line).
- Association is created by hospital admission patterns, not by causal relationship
- **The risk of this type of bias and the associated caveat to interpretation can be recognized prior to any data analysis**
- **Think through this issue when designing queries of RWD!**

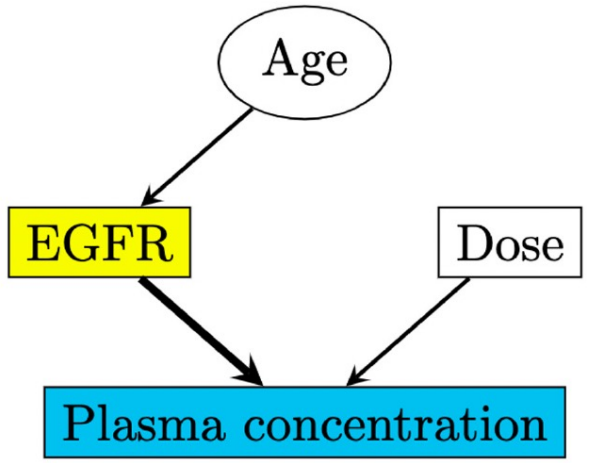
METRUM RESEARCH GROUP

Correlated Covariates in a Pop PK Model: What to Do?

Age and EGFR (mL/min/1.73m^2) are generally (negatively) correlated. Does this mean we shouldn't include both covariates in our model for clearance?



We might have good a priori evidence that clearance is (almost) entirely renal, in which case a reasonable DAG for estimating the effect of EGFR might be:



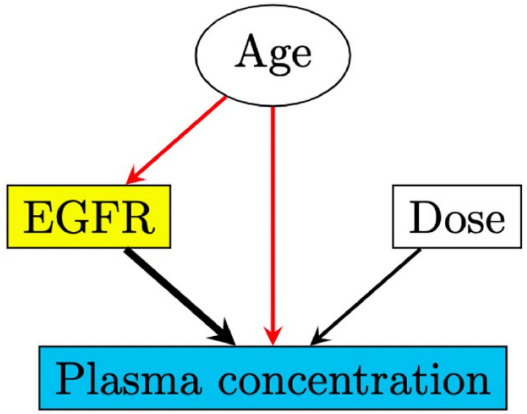
In this case, DAG tells us there are no backdoor path, so just state renal-mediation assumption clearly and remove age as a covariate in our model

METRUM RESEARCH GROUP

However …

Sometimes We Absolutely Should Include Correlated Covariates

We might not have sufficient prior evidence to rule out other age-related effects on clearance (e.g. via concomitant mediations). In this case we should err on the side of drawing the arrow:

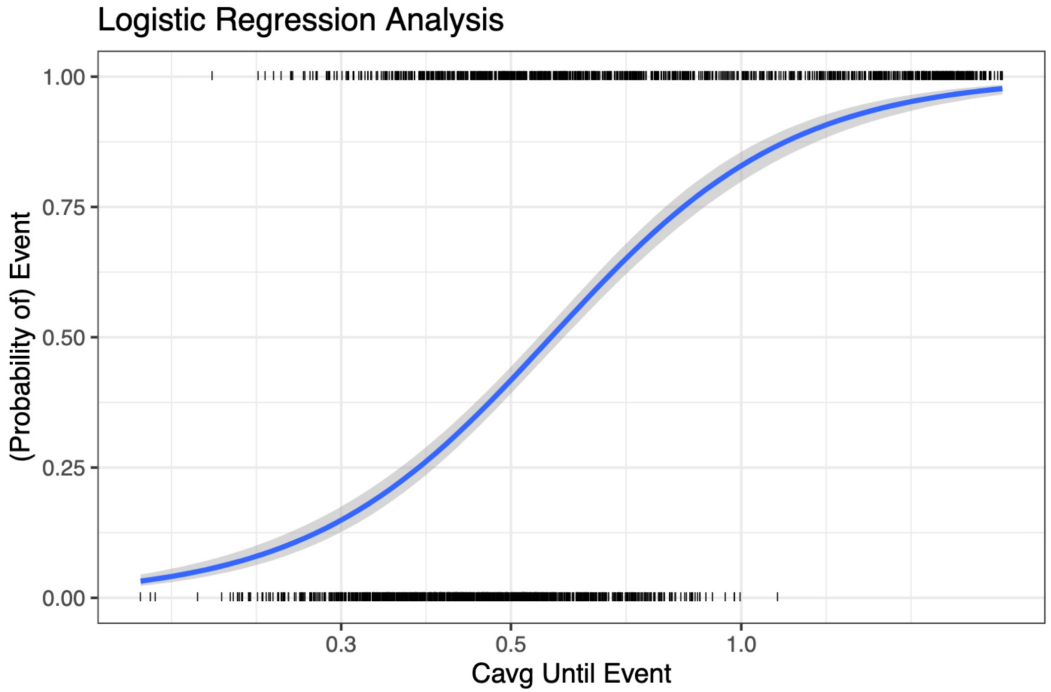


In this case, backdoor path (in red) tells us we *should* include both age and EGFR in the model :

- Decision should not be based on the correlation between the covariates (except in extreme cases)
- Decision should not be based on significance or width of confidence intervals (non-significance and wider confidence intervals are appropriate if we can't isolate the direct effects of EGFR and/or age)
- Decision should be based on an understanding of the causal structure, not statistical properties

METRUM
RESEARCH GROUP

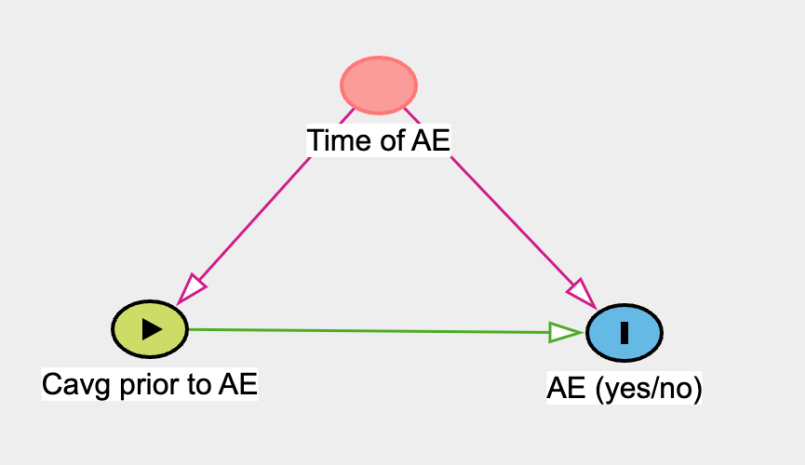
Respecting DAGs and Target Trials When We Choose Exposure Metrics

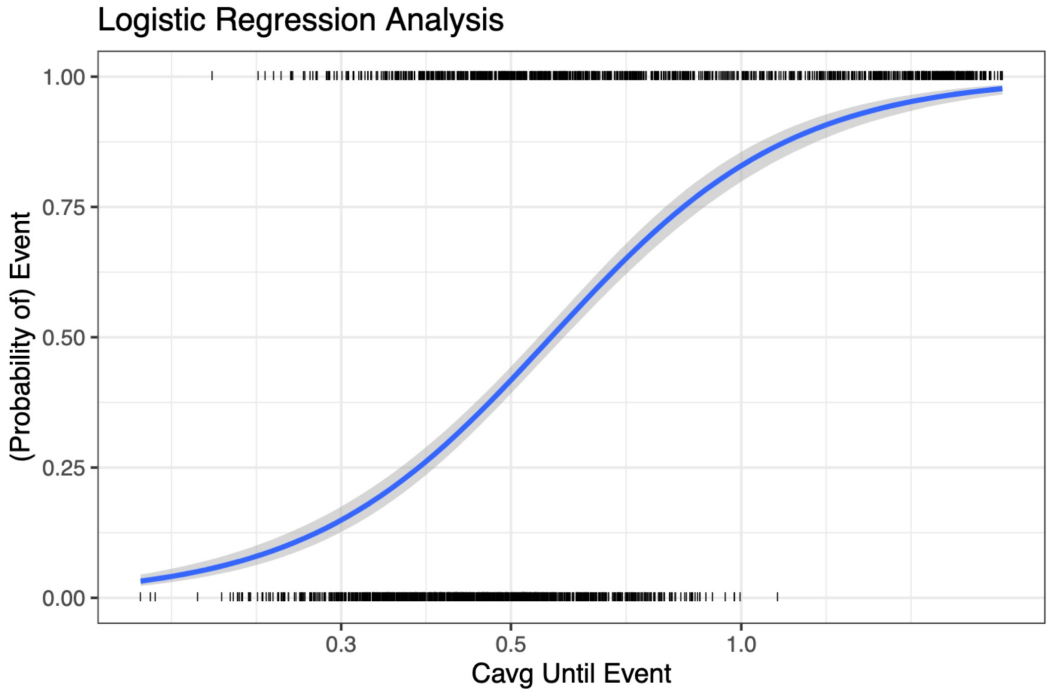


Target trial question:
Q: Setting aside all practical constraints, would it be logically possible to randomly assign patients to a level of “Cavg prior to AE”
A: Not possible even in principle, because the randomization would have to be prior to the AE, so “Cavg prior to AE” would not be computable at randomization

METRUM
RESEARCH GROUP

Respecting DAGs and Target Trials When We Choose Exposure Metrics





~~High exposures caused earlier adverse events~~

Earlier adverse events caused high exposures

METRUM RESEARCH GROUP

Stay tuned on this;
Manuscript submitted: Wiens, French, Rogers

A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks

Miguel A. Hernán, John Hsu, and Brian Healy

For much of the recent history of science, learning from data was the academic realm of statistics,[1,2] but in the early 20th century, the founders of modern statistics made a momentous decision about what could and could not be learned from data: They proclaimed that statistics could be applied to make causal inferences when using data from randomized experiments, but not when using nonexperimental (observational) data.[3,4,5] This decision classified an entire class of scientific questions in the health and social sciences as not amenable to formal quantitative inference.

Not surprisingly, many scientists ignored the statisticians' decree and continued to use observational data to study the unintended harms of medical treatments, health effects of lifestyle activities, or social impact of educational policies. Unfortunately, these scientists' causal questions often were mismatched with their statistical training. Perplexing paradoxes arose; for example, the famous "Simpson's paradox" stemmed from a failure to recognize that the choice of data analysis depends on the causal structure of the problem.[6] Mistakes occurred. For example, as a generation of medical researchers and clinicians believed that postmenopausal hormone therapy reduced the risk of heart disease because of data analyses that deviated from basic causal considerations. Even today, confusions generated by a century-old refusal to tackle causal questions explicitly are widespread in scientific research.[7]

To bridge science and data analysis, a few rogue statisticians, epidemiologists, econometricians, and computer scientists developed formal methods to quantify causal effects from observational data. Initially, each discipline emphasized different types of causal questions, developed different terminologies, and preferred different data analysis techniques. By the beginning of the 21st century, while some conceptual discrepancies remained, a unified theory of quantitative causal inference had emerged.[8,9]

We now have a historic opportunity to redefine data analysis in such a way that it naturally accommodates a science-wide framework for causal inference from observational data. A recent influx of data analysts, many not formally trained in statistical theory, bring a fresh attitude that does not a priori exclude causal questions. This new wave of data analysts refer to themselves as data scientists and to their activities as data science, a term popularized by technology companies and embraced by academic institutions.

Data science, as an umbrella term for all types of data analysis, can tear down the barriers erected by traditional statistics; put data analysis at the service of all scientific questions, including causal ones; and prevent unnecessary inferential mistakes. We may miss our chance to successfully integrate data analysis into all scientific

I agree!
Let's do it!

Hernán, Hsu, and Healy. 2019. *Chance* 32(1).

METRUM
RESEARCH GROUP

Thank You!
jimr@metrumrg.com

METRUM
RESEARCH GROUP