Harnessing Big Data in Nephrology: What Makes Your Data-Driven Research Publishable?

Presented by
Germaine Wong
Professor of Clinical Epidemiology
Sydney School of Public Heath
Email: germaine.wong@health.nsw.gov.au

THE UNIVERSITY OF SYDNEY

Disclosures

- Associate Editor of Kidney International
- Past Associate Editor of Transplantation

Promise of Big Data

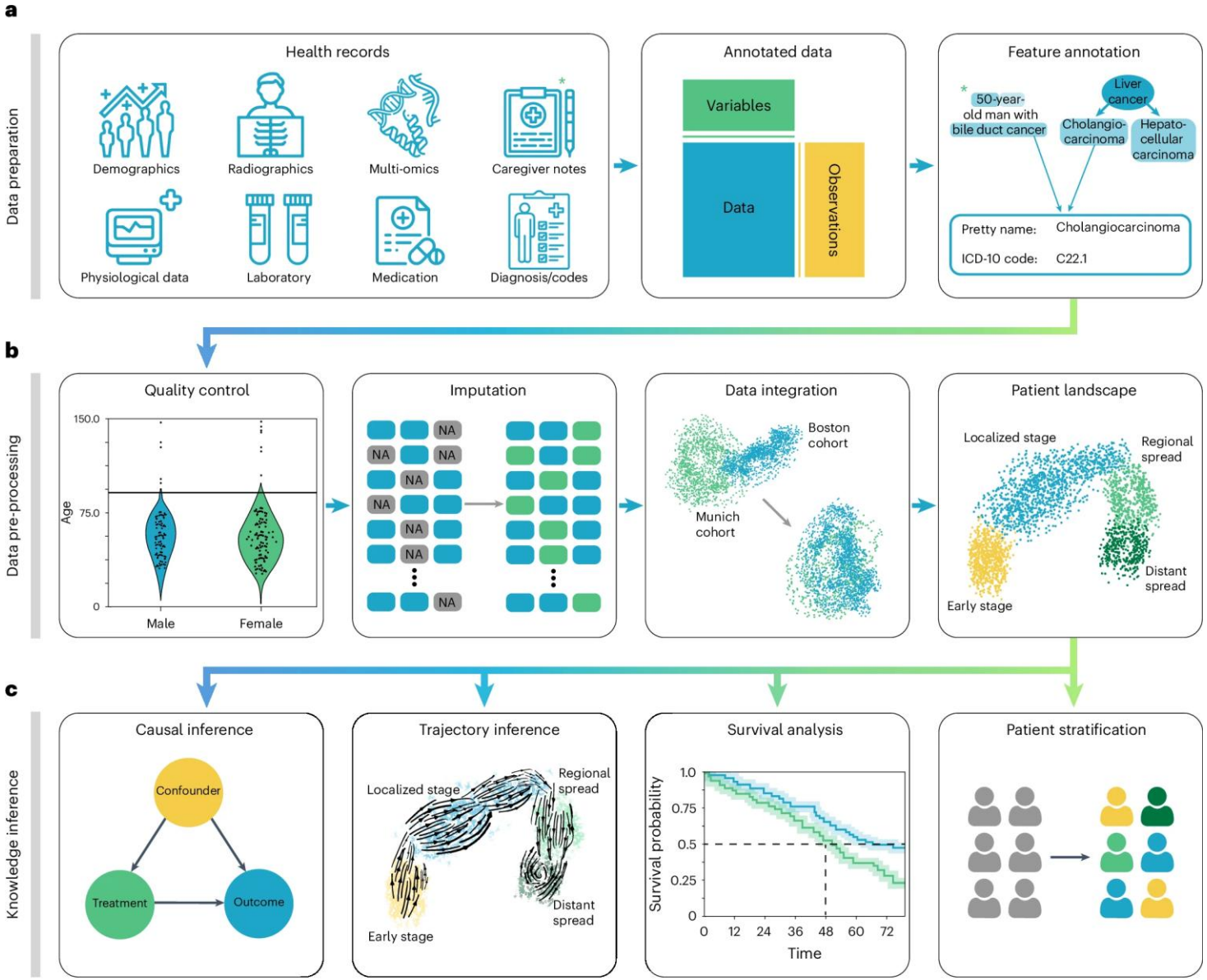
Data sources: Omics

Omics	Sample	Outcomes
Genomics	Blood and biopsies	Acute and chronic rejection, long-term allograft function
Transcriptomics	Biopsies, whole blood and urine	IFTA, acute and chronic rejection, tolerance, graft function
Proteomics	Urine	Acute and chronic rejection, IFTA
Metabolomics	Urine, serum, biopsies	Allograft function, rejection
Multi-omics	Blood, biopsies, plasma	Acute and chronic rejection

Registry data

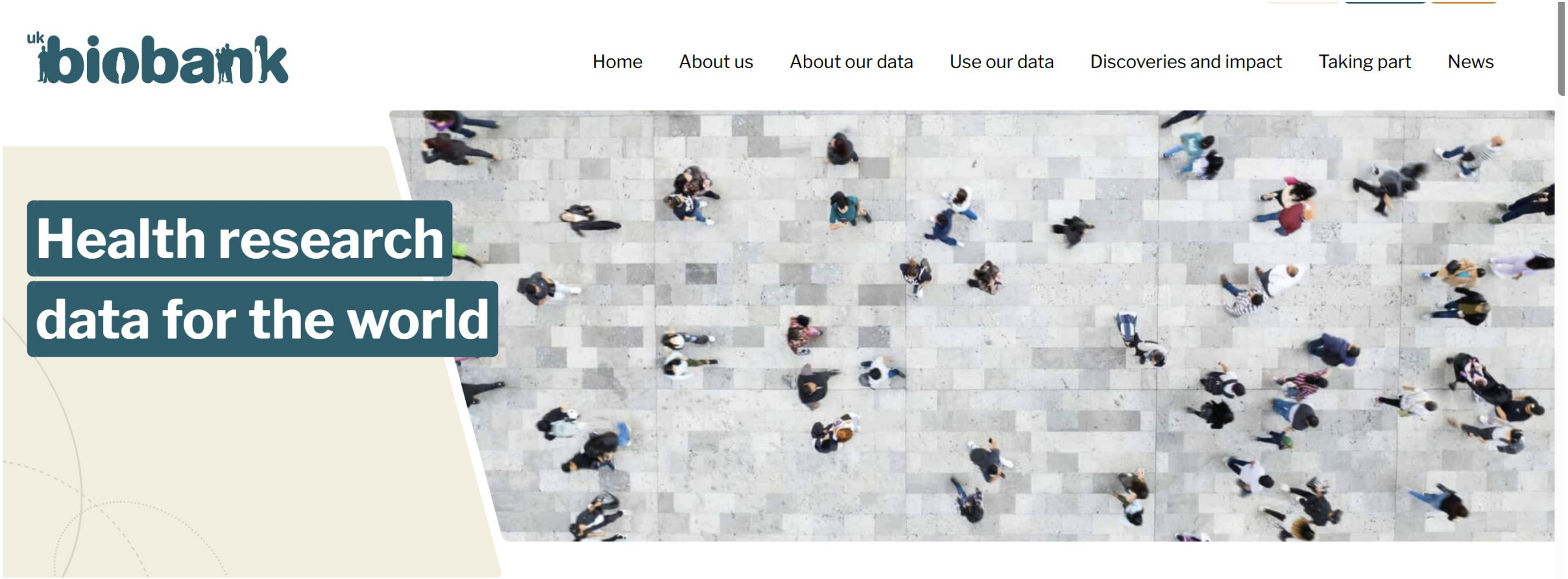
- International Registry in Organ Donation and Transplantation (IRODaT)
  www.irodat.org/
- WHO Global Observatory on Donation and Transplantation (GODT)
  www.transplant-observatory.org
- Council of Europe: Newsletter Transplant: International figures on Donation and Transplantation 2014
  https://www.edqm.eu/sites/default/files/newsletter_transplant_2015_2.pdf
- Bone Marrow Donors Worldwide
  www.bmdw.org
- Familial Amyloidotic Polyneuropathy World Transplant Register
  www.fapwtr.org
- International Intestinal Transplant Registry
  www.intestinaltransplant.org/itr
- International Islet Transplant Registry
  www.med.uni-giessen.de/itr
- International Pancreas & Islet Transplant Association (IPITA)
  www.tts.org/ipita/valuable-resources-and-links/registries
- International Registry for Heart and Lung Transplantation (ISHLT)
  www.ishlt.org/registries/heartLungRegistry.asp
- Australia and New Zealand: Organ Donation and Transplantation Registry
  www.anzdata.org.au
- Collaborative Transplant Study (CTS) (Heidelberg University)
  http://www.ctstransplant.org
- European Liver Transplant Registry
  www.eltr.org
- European Renal Association - European Dialysis and Transplant Association Registry - ERA-EDTA
  www.era-edta.org/
- Eurotransplant statistics (Austria, Belgium, Germany, Luxemburg, the Netherlands, Slovenia and Croatia)
  http://www.eurotransplant.org/cms/index.php?page=yearlystats
- Indian Transplant Registry
  www.transplantindia.com
- La Sociedad de Transplante de América Latina y el Caribe (STALYC) registry
  http://www.stalyc.net/en/registry.html
- Scandiatransplant (Island, Norway, Finland, Denmark and Sweden)
  www.scandiatransplant.org/data/scandiatransplant-figures

Electronic Health Record



Huemos et al Nature Medicine 2024

Open-source biobank



Research question

- Well-formulated
- Address an important clinical question
- Addresses a knowledge gap
- Research impact

Not data mine

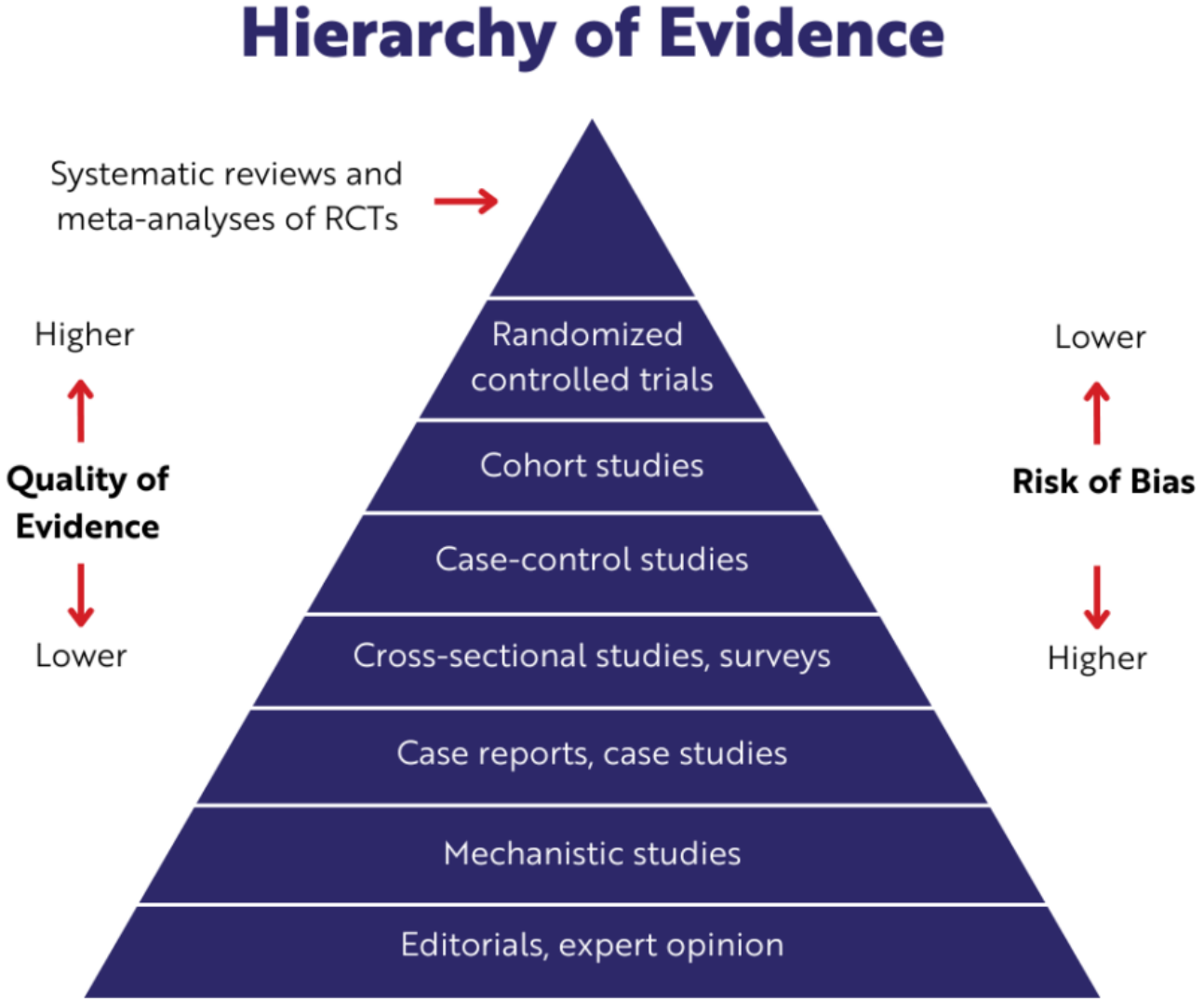
Hypotheses?

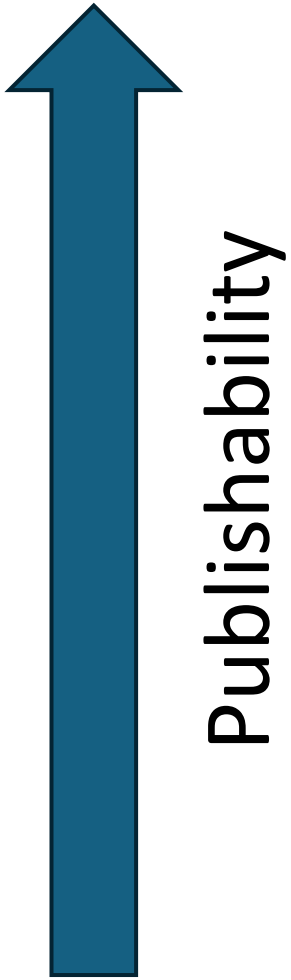
Beware of testing too many of them!

Testing too many may exaggerate the significance of the findings (Type 1 error)

Or can also lead to falsely accepting a null hypothesis (Type 2 error)

Study designs





An appropriate study design to address the research question of interest

Gust de Backer

Methodology - Quantitative research – Editors' pet hate!

DO NOT claim	DO NOT conduct	DO NOT forget about
DO NOT claim causality when you have found an association	DO NOT conduct multiple tests without correction	DO NOT forget about missing values

Methodology - Quantitative research – Editors' pet hate!

DO NOT	DO NOT misinterpret	DO NOT ignore	ENSURE	DO NOT ignore
DO NOT over-adjust for mediators	DO NOT misinterpret the P values and the Confidence Intervals	DO NOT ignore study power and precision	ENSURE appropriate measures of exposure and outcomes	DO NOT ignore time-varying covariates

Methodology - Quantitative research – Editors' pet hate!

DO NOT violate	DO NOT	DO NOT	DO NOT ignore	DO NOT treat
DO NOT violate model assumptions	DO NOT data dredge and p-hacking	DO NOT over-generalize findings	DO NOT ignore competing risk	DO NOT treat observational like randomized controlled trial

Methodology - Quantitative research – Editors' pet hate!

DO NOT

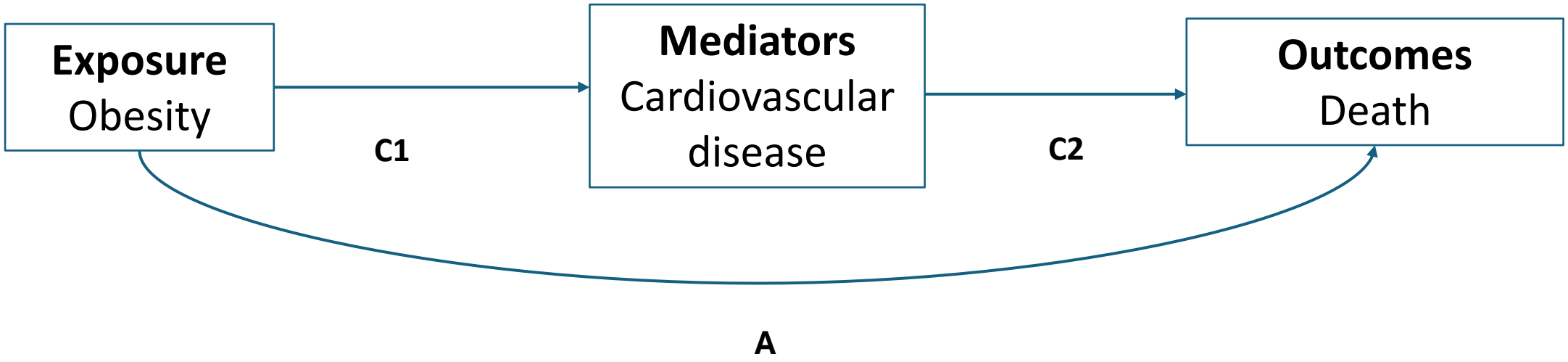
DO NOT <u>over-adjust</u> for mediators

Over-adjustment bias

Research Question:

In patients treated with dialysis, does being obese increase the risk of overall mortality? (**Direct effect - A**)

What are the factors that may contribute to this relationship? (**Indirect effect C1 and C2)**

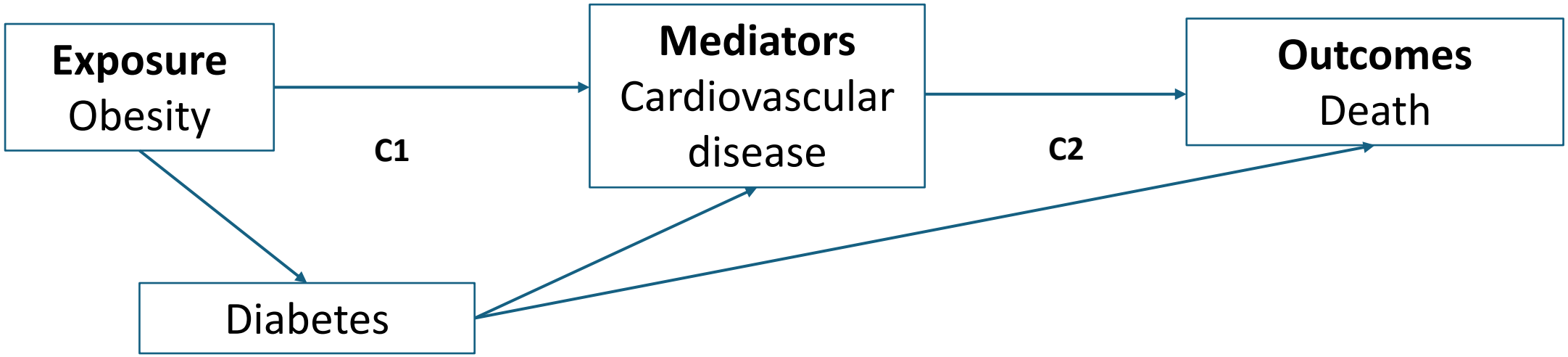


Over-adjustment bias

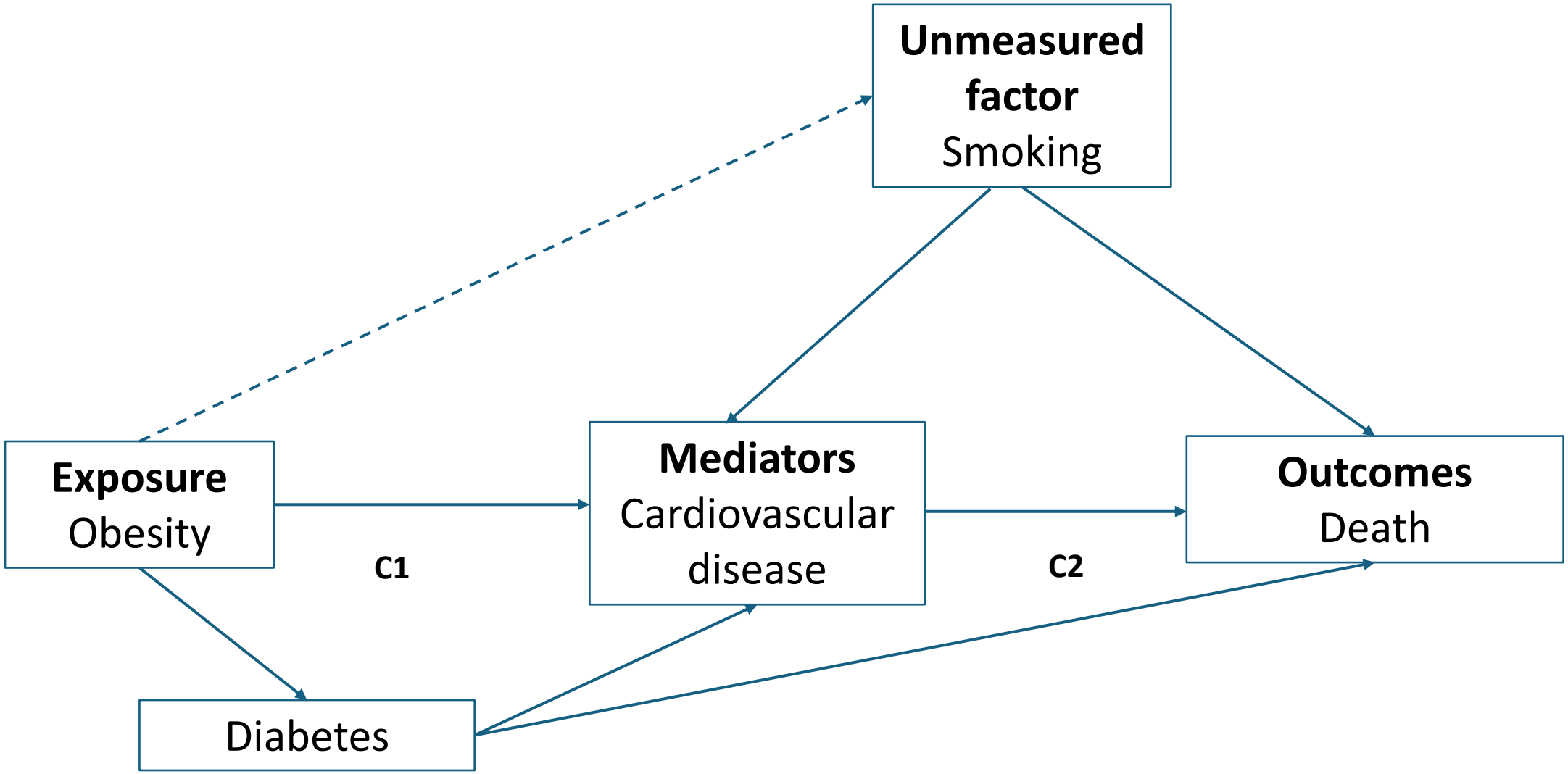
Research Question:

In patients treated with dialysis, does being obese increase the risk of overall mortality? (**Direct effect - A**)

What are the factors that may contribute to this relationship? (**Indirect effect C1 and C2)**



Collider bias



Journal of Clinical Epidemiology 149 (2022) 127–136

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Avoiding overadjustment bias in social epidemiology through appropriate covariate selection: a primer

Anita van Zwieten[a,b], Peter W.G. Tennant[c,d,e], Michelle Kelly-Irving[f,g], Fiona M. Blyth[a,h], Armando Teixeira-Pinto[a,b], Saman Khalatbari-Soltani[a,h,*]

[a]*The University of Sydney School of Public Health, Faculty of Medicine and Health, Sydney, New South Wales, Australia*
[b]*Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, New South Wales, Australia*
[c]*Leeds Institute for Data Analytics, University of Leeds, Leeds, UK*
[d]*Faculty of Medicine and Health, University of Leeds, Leeds, UK*
[e]*Alan Turing Institute, London, UK*
[f]*Equity Research Team, CERPOP, Université de Toulouse, Inserm, Université Paul Sabatier, Toulouse, France*
[g]*Institut fédératif d'études et recherche interdisciplinaire santé société (Iferiss), Université Toulouse III Paul Sabatier, Toulouse, France*
[h]*ARC Centre of Excellence in Population Aging Research (CEPAR), University of Sydney, Sydney, Australia*

Accepted 25 May 2022; Published online 2 June 2022

Methodology - Quantitative research – Editors' pet hate!

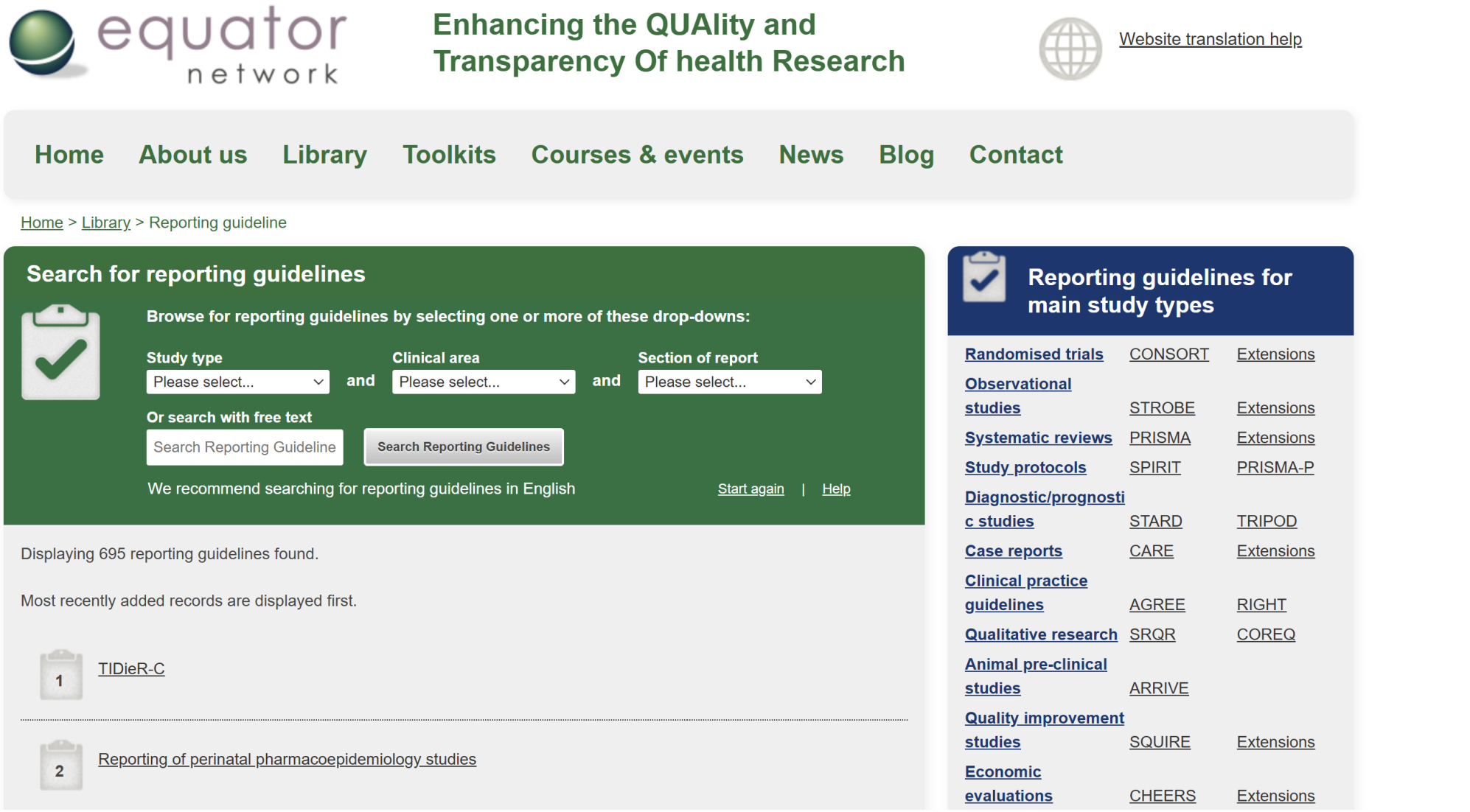
DO NOT ignore

DO NOT ignore time-varying covariates

Time-Dependent Anemia, EPO Use, and Mortality in Dialysis

- Anemia is a dynamic, time-varying exposure in dialysis.
- EPO dosing responds to anemia → creating feedback loops.
- Structure: Anemia(t) → EPO(t) → Anemia(t+1) → Mortality.
- EPO is both a mediator and a time-dependent confounder affected by prior anemia.
- Standard Cox models introduce:
- – Over-adjustment (blocking anemia→EPO→mortality pathway)
- – Collider bias when conditioning on EPO or Hb
- Correct causal method: Marginal Structural Models (MSMs) with IPTW.

Reporting guidelines



What makes a quantitative big-data research paper publishable?

- Clear, clinically meaningful research question
- Casual framework for guidance
- Analytic rigour and transparency
- Avoid big-data pitfalls
- Data provenance, quality and validation
- Interpretable, clinically plausible findings
- Coherent narrative