**PRINCIPAL INVESTIGATOR RESEARCH TRAINING**
**SESSION 2:**

# Study Design, Blinding, Confounding

Fridtjof Thomas, PhD

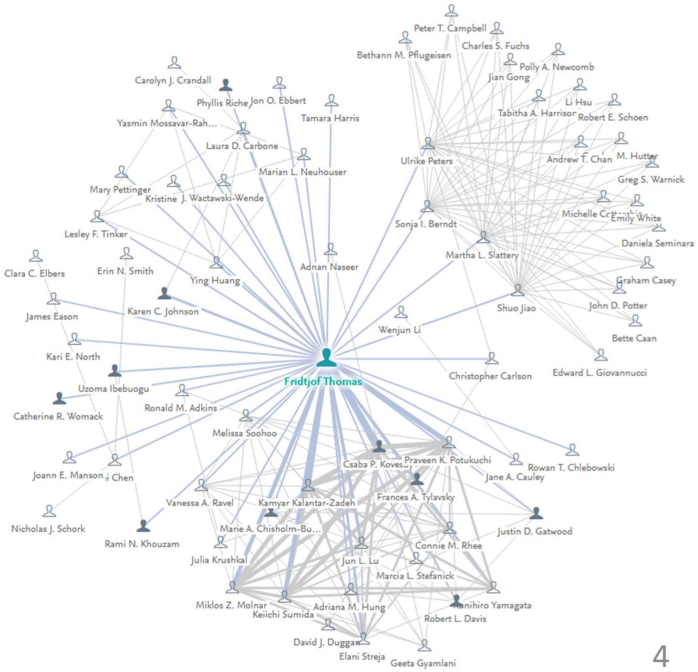
May 06, 2021

# About the presenter

- Professor at the Division of Biostatistics, Dept. of Preventive Medicine
- At UTHSC since 2007
- Design and Analysis Committee of the EARLY trials (2010-2016 – "Early Adult Reduction of weight through LifestYle intervention," a collection of seven randomized clinical trials funded by the National Heart, Lung, and Blood Institute (NHLBI) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health (NIH))
- Member of the Biostatistics Collaborative Core at the Southeast Regional Center of the NIH-NHLBI-funded Women's Health Initiative (WHI) study that has recruited over 160,000 women in over 40 clinical centers nationwide. (2010-2017)
- Grant review experience since 2012 from
  - Department of Defense's Congressionally Directed Medical Research Program (DoD CDMRP)
  - NIH Epidemiology of Chronic and Infectious Disease Study Section
  - NIH Neurological, Aging, and Musculoskeletal Epidemiology (NAME) Study Section

# Network of collaboration (Elsevier Pure "Fingerprint" 10/09/2019)

| Research output network – organizational units | |
|---|---|
| Department of Preventive Medicine | 77 |
| Department of Medicine, Division of Nephrology | 23 |
| Department of Pediatrics | 6 |
| Department of Pharmaceutical Sciences | 5 |
| Department of Clinical Pharmacy – Memphis | 4 |
| Department of Clinical Pharmacy – Nashville | 4 |
| Neuroscience Institute | 4 |
| Department of Ophthalmology | 4 |
| Department of Medicine, Division of Cardiovascular Diseases | 4 |
| Department of Medicine, Division of General Internal Medicine | 3 |
| Department of Acute and Tertiary Care | 2 |
| Department of Health Promotion and Disease Prevention | 2 |



# Outline

- Study Design: what does it accomplish?
- Blinding: why?
- Confounding: what it is and how to handle it

# Study Design

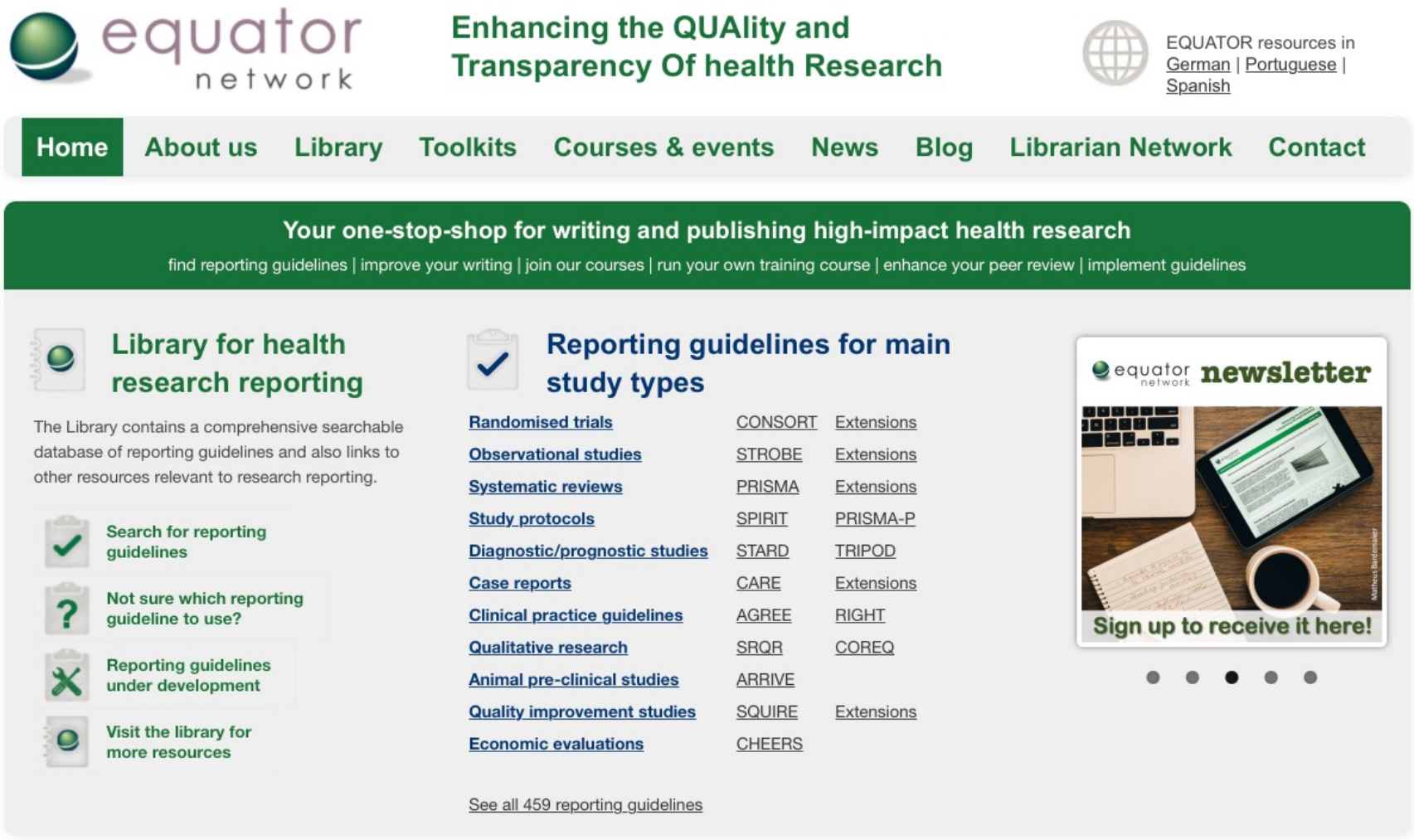
- A well-designed study
  - Will answer the question you want to answer
  - Will often allow for relatively simple data analysis
  - Will convince grant reviewers that only the funding (their positive judgement!) is needed to translate your great idea into truly useful insights/results!
  - Will be easy to write-up for publication

# Study Design (cont.)

- Every study needs to be designed!
  - Randomized clinical trials
  - Secondary data analyses
  - Observational studies

# Tip: Check out early what is needed!

https://www.equator-network.org



## CONSORT 2010 checklist of information to include when reporting a randomised trial*

| Section/Topic | Item No | Checklist item | Reported on page No |
|---|---|---|---|
| **Title and abstract** | | | |
| | 1a | Identification as a randomised trial in the title | |
| | 1b | Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts) | |
| **Introduction** | | | |
| Background and objectives | 2a | Scientific background and explanation of rationale | |
| | 2b | Specific objectives or hypotheses | |
| **Methods** | | | |
| Trial design | 3a | Description of trial design (such as parallel, factorial) including allocation ratio | |
| | 3b | Important changes to methods after trial commencement (such as eligibility criteria), with reasons | |
| Participants | 4a | Eligibility criteria for participants | |
| | 4b | Settings and locations where the data were collected | |
| Interventions | 5 | The interventions for each group with sufficient details to allow replication, including how and when they were actually administered | |
| Outcomes | 6a | Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed | |
| | 6b | Any changes to trial outcomes after the trial commenced, with reasons | |
| Sample size | 7a | How sample size was determined | |
| | 7b | When applicable, explanation of any interim analyses and stopping guidelines | |
| Randomisation: | | | |
| Sequence generation | 8a | Method used to generate the random allocation sequence | |
| | 8b | Type of randomisation; details of any restriction (such as blocking and block size) | |
| Allocation concealment mechanism | 9 | Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned | |
| Implementation | 10 | Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions | |
| Blinding | 11a | If done, who was blinded after assignment to interventions (for example, participants, care providers, those | |



3 pages long check list – look at it early!

Start with the end in mind – your manuscript!

# Study Design (cont.)

- Are you assigning the exposure?

  If ”yes”: Experimental study

  Gold standard: double-blinded randomized controlled trial

  If “no”: Observational study

  For analytical studies (with control groups):

  - Cross-sectional studies
  - Cohort studies (exposure -> outcome)
  - Case-control studies (outcome -> exposure)

# How strong will your derived evidence be?
# Levels of evidence

| | |
|---|---|
| 1a | Systematic review of high quality RCTs with similar results and effect sizes for many different RCTs. |
| 1b | Individual high quality RCT with high precision (narrow confidence interval) |
| 1c | All or none |
| 2a | Systematic review of cohort studies with similar results and effect sizes. |
| 2b | Individual cohort study or low quality RCT (e.g., <80% follow-up) |
| 2c | "Outcomes Research" and ecological studies (based on average exposures etc. of populations of geographical or temporal units) |
| 3a | Systematic review of case-control studies |
| 3b | Individual case-control study |
| 4 | Case-series and poor-quality cohort and case-control studies |
| 5 | Expert opinion (unless critically appraised or based on "first principles") |

*Source: Oxford Centre for Evidence-based Medicine*
*https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/*

# All or none: Example "Bubble Boy" disease

- Babies born without functional immune system.
- SCID-X1: 1 in 50,000-100,000 affected; caused by a mutation in a gene (IL2RG)
- Most die within first year of life. (Only about 20% have access to a suitable sibling for a bone-marrow transplant as the existing cure.)

St. Jude announced April 18, 2019: Gene therapy cure for babies with X-linked severe combined immunodeficiency

"The gene therapy, produced in the Children's GMP, LLC, manufacturing facility on the St. Jude campus, involved use of a virus to transport and insert a correct copy of a gene into the genome of patients' blood stem cells. Following the treatment, the children began producing functioning immune cells for the first time, according to St. Jude, and most patients were discharged from the hospital within one month." [All 8 babies started to produce complete sets of immune cells.]

https://www.stjude.org/inspire/news/bubble-boy-scid-x1-cure.html
https://www.stjude.org/research/news-publications/research-highlights/2019-research-highlights/st-jude-gene-therapy-holds-promise-for-treating-several-diseases.html

# Clinical trials: An Early Example

- **Gustav III of Sweden** (1746–1792) believed that coffee was bad for one's health
- The king decided to demonstrate that believe in a clinical trial
- He picked two identical twins that had been sentenced to death for some crime and commuted their death sentence to life in prison under the following condition:
  - One twin was to drink three pots of coffee a day
  - The other twin was to drink a comparable amount of tea
- The outcome was death
- Two physicians were appointed to verify the outcome of the two participants

Source:

Afshari R. (2017). Gustav III's risk assessment on coffee consumption; A medical history report. *Avicenna journal of phytomedicine*, *7*(2), 99–100.

https://en.wikipedia.org/wiki/Gustav_III_of_Sweden%27s_coffee_experiment

# The outcome of the trial…

The outcome was somewhat unexpected:

- The two physicians died first
- Gustav III was assassinated (at a masquerade ball at the Royal Opera House in Stockholm at midnight on 16 March 1792 – he did not die immediately from the pistol shot but by an infection of the wound 13 days later; Giuseppe Verdi's opera *Un Ballo in Maschera* is based on this event, albeit eventually set in Boston during the colonial era - see https://www.npr.org/2008/03/28/89126026/exiled-to-boston-verdis-a-masked-ball )
- The tea-drinking twin died at the high age of 83; the age of death of the coffee-drinking twin is unknown/lost to history.

Postscript: Coffee was repeatedly banned in Sweden throughout the years but became and remains very popular after the last ban was lifted in 1822.

# What was going wrong in this trial?

**The good:**

- Thoughtful elements like using twins to adjust for confounding/genetics and possibly early childhood experiences etc.
- Clearly defined outcome with adjudication by MDs!

**The bad:**

- Sample size? Randomization?
- Outcome to hard to observe?
- Lack of continuation planning when essential personnel becomes unavailable

**The ugly:**

- Coercive enrollment scheme

# Why is a double-blinded randomized controlled trial the gold standard?

- If neither the investigator nor the participant know which treatment they administer/receive, that knowledge cannot influence the outcome (placebo effects etc.)

- Randomization "controls" (breaks associations) for known and unknown confounders

Thus, we will investigate blinding and confounding in greater detail.

# Blinding (masking)

- Blinding: Keeping the study group assignment hidden after allocation:
  - From the investigators
  - From the participant
  - From the care provider
  - From the assessor of outcomes
  - From the data analyst/statistician doing the analysis

- Study groups are simply referred to by, e.g., A = Group 1 and B = Group 2 (this is not unproblematic as inadvertently unblinding one participant easily can lead to a loss of blinding for all trial participants)
- The blind is often lifted for all after data collection is complete.

Masking is a better term especially when ophthalmologic conditions are studied (like blindness as study outcome).

# Why blinding?

- Maintain the internal validity of the trial by avoiding *bias* (inadvertently skewing results)
  - Due to intervention delivery differences
  - Placebo effects
  - Assessment differences
  - Differences in “data cleaning” etc.

- The protocol should explicitly describe who will be blinded and what measures are taken to keep the blind, e.g.:
  - By using specific flavors to mask distinctive tastes in medications
  - by asking participants who know their group assignment to not reveal that to study personnel during clinic visits when measurements are taken

# Blinding: Some Remarks

- Blinding requires integrity from the PI: It usually means that the PI does not know everything and accepts that, e.g., the biostatistician represents the study in closed parts of DSMB meetings.

- Should DSMBs be blinded to treatment assignment, especially when looking at AEs/SAEs?
  - There are good arguments for either position (DSMB should/should not be blinded)
  - Important: If group assignment is revealed, it should be clearly stated in the minutes (it is usually not possible to go back to "blinded" in later DSMB meetings due to identifying characteristics of the groups)

- Blinding is most often associated with RCTs, but can be important in observational studies as well, e.g., when adjudicating outcomes from EMRs or interpreting free text/progress notes of relevance to the study question.

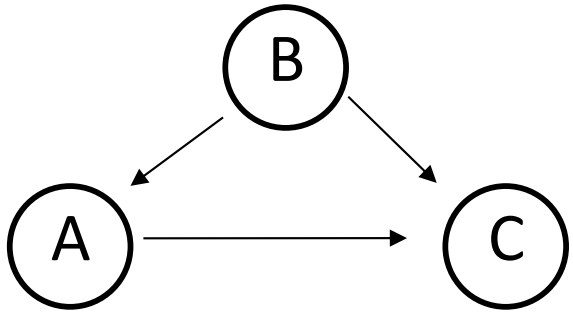
# Blinding: Summary

- Blinding is an essential component to assure the validity of the trial.

- Blinding is not an all-or-nothing concept: E.g., if treatment delivery requires knowledge about the treatment, those evaluating the outcome can still be blinded.

# Confounding: What is it?

## B is a confounder of A -> C




A and B correlate

B and C correlate

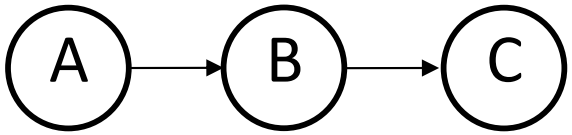
A and C correlate, even when controlling for B

Examples:

1. Smokers are younger and the outcome is risk of stroke.
2. A cardiovascular risk factor (B) leads to prescription of Aspirin and the outcome is myocardial infarction. (Confounding by indication)

# Not everything that is “related” is confounded…

## Example: B is mediator of A -> C




A and B correlate

B and C correlate

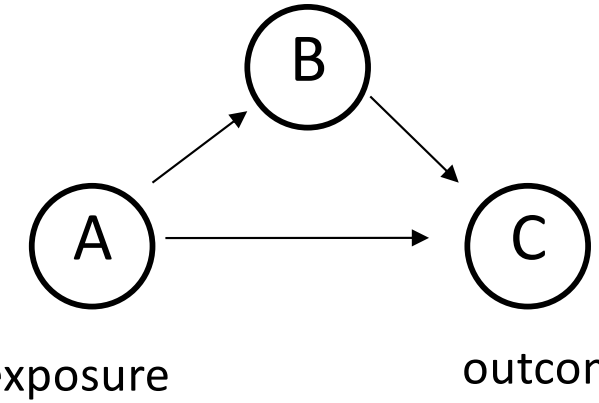
A and C correlate, but not when conditioning on B

Mediator or intermediate variable

Example: Salt (A), blood pressure (B), coronary heart disease (CHD).

# Not everything that is "related" is confounded…

# Example: B is a moderator of A -> C




A and B correlate

B and C correlate

A and C correlate, even when controlling for B

Moderator or effect modifier

Examples:

1. Immunization status (B) with consequences (C) of exposure to pathogenic organisms (A).
2. Alcohol level in blood (C) as a function of alcohol intake (A) and the metabolic process in the liver (by enzymes) (B).

# Confounding (cont.)

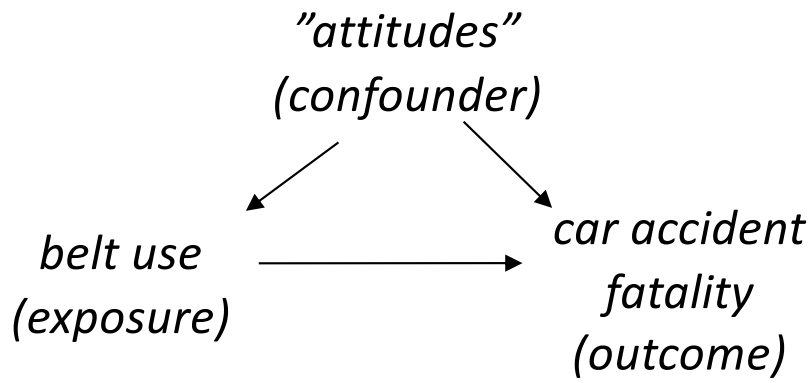
- Can be controlled by
  - Randomization: controls even for unknown confounders
  - Multivariable analysis: adjusting for age, medical history, known risk factors, etc.
  - Matching: making exposed and unexposed comparable in their traits
  - Restriction: controlling for confounding by focusing on subgroups that are not affected by the confounding factor(s)

# Restriction: Example seat belt usage

**Question: Do safety belts prevent fatalities in car accidents?**

Problem: Safety belt usage might be influenced by attitudes that also influence driving with respect to

- speed,
- distance to car ahead,
- driving during adverse weather conditions,
- etc.




Consequence: Naively comparing outcomes from those that used safety belts and those that did not will be flawed (confounded by the underlying attitudes that correlate with belt use and “riskiness” of driving behavior which leads to accidents at possibly different frequencies and at different speeds).

# Restriction: Example seat belt usage

Evans used the U.S. Fatal Accident Reporting System (FARS) and looked at the data in this way:
Among all fatal accidents (involving at least one fatality), some involved cars where two passengers were traveling on the front seats, one belted and one unbelted. Several normally uncontrolled features are identical for driver and passenger: speed, friction on road surface, distance from car ahead, reaction time, etc. The following table (as reported in Rosenbaum, p. 10) summarizes the data:

**Table 1.1** Crashes in FARS 1975–1983 in which the front seat had two occupants, a driver and a passenger, with one belted, the other unbelted, and one died and one survived.

| | Driver<br>Passenger | Not Belted<br>Belted | Belted<br>Not Belted |
|---|---|---|---|
| Driver Died | Passenger Survived | 189 | 153 |
| Driver Survived | Passenger Died | 111 | 363 |

Rosenbaum PR. Design of Observational Studies. New York: Springer; 2010.
Evans L. The effectiveness of safety belts in preventing fatalities. Accident Analysis & Prevention. 1986;18(3):229-41. doi: https://doi.org/10.1016/0001-4575(86)90007-2.

# Confounding: Summary

- Confounding is one reason why naively "looking at the data" can be very misleading
- Age is a confounder for most health conditions: You almost certainly need to control for age in your design and/or analyses.
- Likewise: You need to adjust for all major known confounders and effect modifiers:
  - Observational studies: rich set of covariates; smart designs using restriction; etc.
  - Randomized trials: Randomization will balance observed and unobserved confounders (at least for traits frequent enough)

# Hallmarks of a good study

In addition to blinding of evaluators and confounding, the following aspects must be addressed as well:

- Selection bias
- Measurement bias (misclassification – outcome or covariate; exposure misclassification; systematically missing activities/episodes, e.g., in activity data; recall bias, telescoping bias, etc.)
- Cases and non-cases/controls need identical ways to determine covariates and outcomes!
- "Immortal time bias" in time-to-event analyses
- Differential loss-to-follow-up
- "Artefacts" due to recycling existing data for a different purpose (e.g., billing data; medical prescription data to determine adherence to medication)

**A well-designed study will address all these issues!**

# Summary

- Every study needs a **design**!
- **Blinding** to the extend possible is an essential technique to avoid inadvertently skewing results by knowledge about the exposure status.
- Controlling for **confounders** is needed in all observational studies and many randomized controlled trials (either to increase the precision in the estimates or to adjust for missing data and/or differential loss-to-follow-up; analyses of heterogeneity of treatment effects also involves confounders)