Simulating collider stratification bias and an application to the inverse obesity paradox in prostate cancer

2024/02/15 – LUPOP seminar, Methodological issues in epidemiology and population research

Josef Fritz

josef.fritz@i-med.ac.at, josef.fritz@med.lu.se

Institute of Medical Statistics and Informatics
Medical University of Innsbruck
Austria

Register-based epidemiology
Department of Translational Medicine, Malmö
Lund University
Sweden

Body mass index/obesity and prostate cancer (PCa)

Directions of associations

	Localised PCa	Advanced PCa	PCa-specific death
Genkinger et al. (2020)	neg ↓	null →	pos ↑
Jochems et al. (2020)	neg (HR~0.9)* ↓	null →	pos (HR~1.1-1.2)* ↑

*: per 5-kg/m^2 increase

Less effective asymptomatic PSA screening in obese men (non-biological reason)

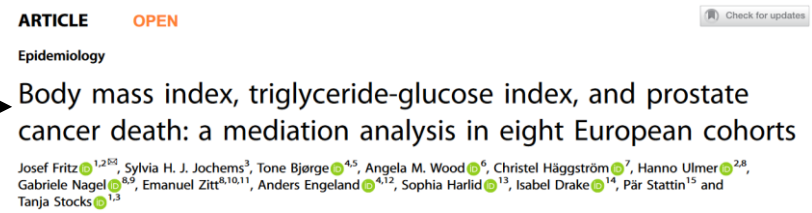
Case-only analysis:

BMI measurement → PCa → PCa death

Inverse obesity paradox

Obesity and prostate cancer death – why?

- ✓Biological mechanisms
  - ❑Insulin resistance
  - ❑Sex hormones
- ✓Less successful treatment in obese men
- ✓Methodological shortcomings
  - ❑Analysis of a composite endpoint (PCa development + PCa death) by following up all participants from study entry in full cohort analyses
  - ❑Collider stratification bias in PCa case only analyses



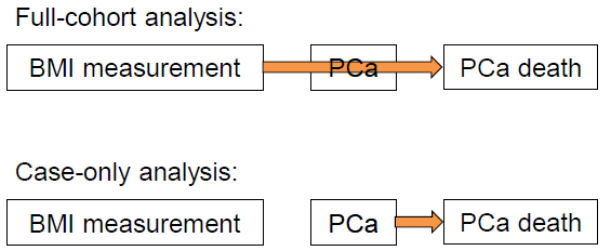
ARTICLE OPEN

Check for updates

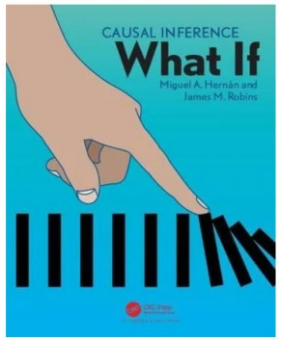
Epidemiology

Body mass index, triglyceride-glucose index, and prostate cancer death: a mediation analysis in eight European cohorts

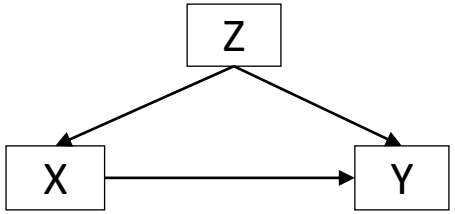
Josef Fritz[1,2✉], Sylvia H. J. Jochems[3], Tone Bjørge[4,5], Angela M. Wood[6], Christel Häggström[7], Hanno Ulmer[2,8], Gabriele Nagel[8,9], Emanuel Zitt[8,10,11], Anders Engeland[4,12], Sophia Harlid[13], Isabel Drake[14], Pär Stattin[15] and Tanja Stocks[1,3]

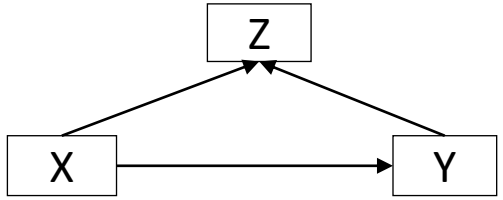


Types of biases according to Miguel Hernán (Causal Inference: What If)

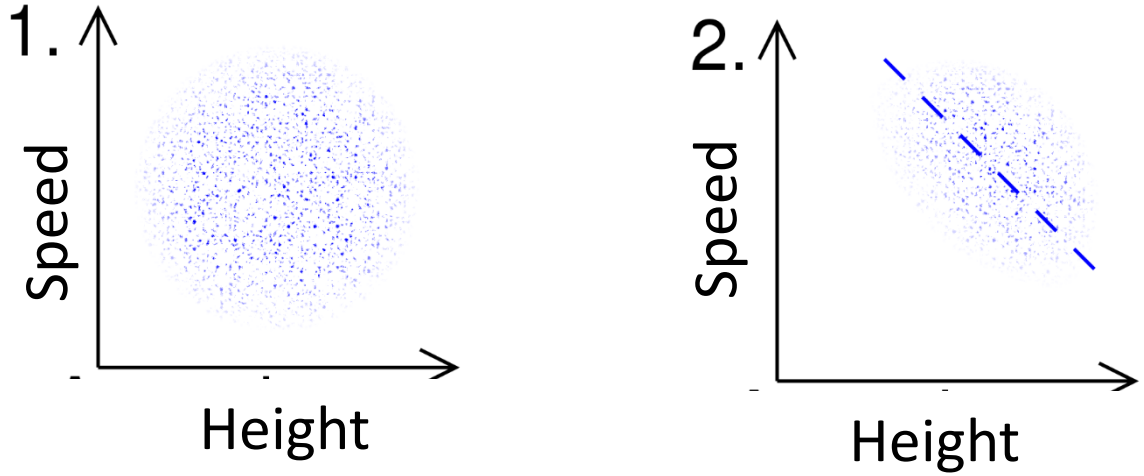


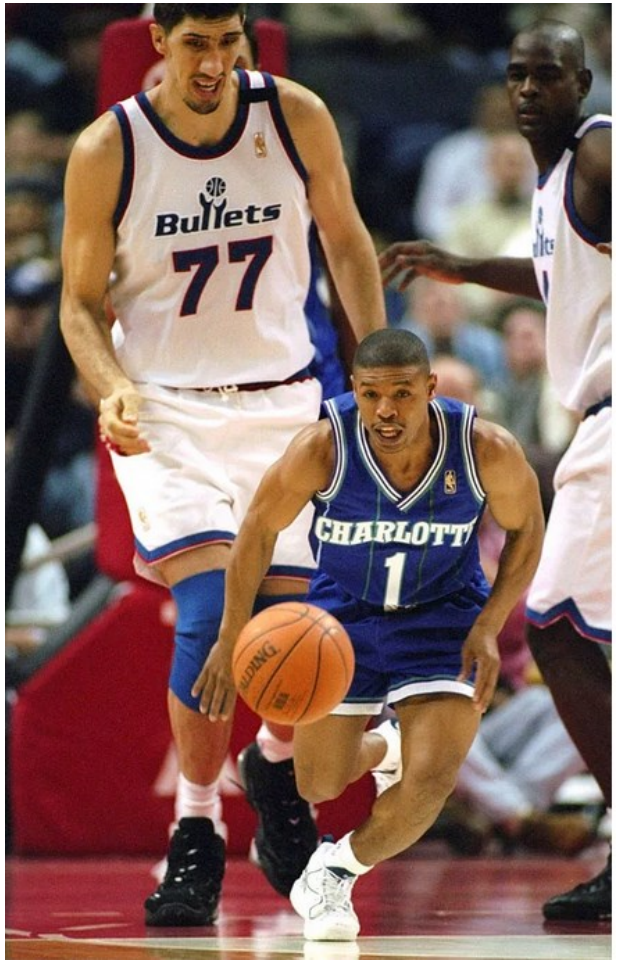
- ❑ Bias due to confounding
  - ✓ Exposure and outcome share common causes
  - ✓ Controlling for Z eliminates bias
- ❑ Bias due to measurement error
- ❑ Bias due to selection
  - ✓ Collider bias
  - ✓ Collider = variable both affected by exposure and outcome
  - ✓ Controlling for Z introduces bias



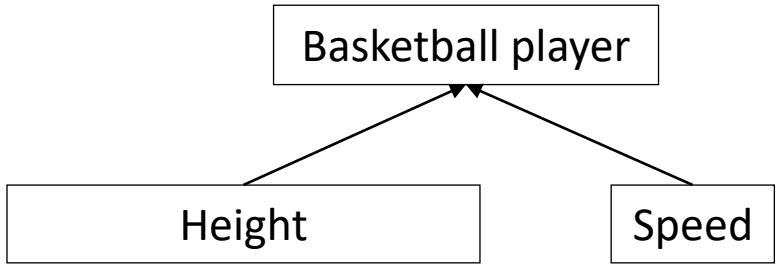


Collider bias - illustration





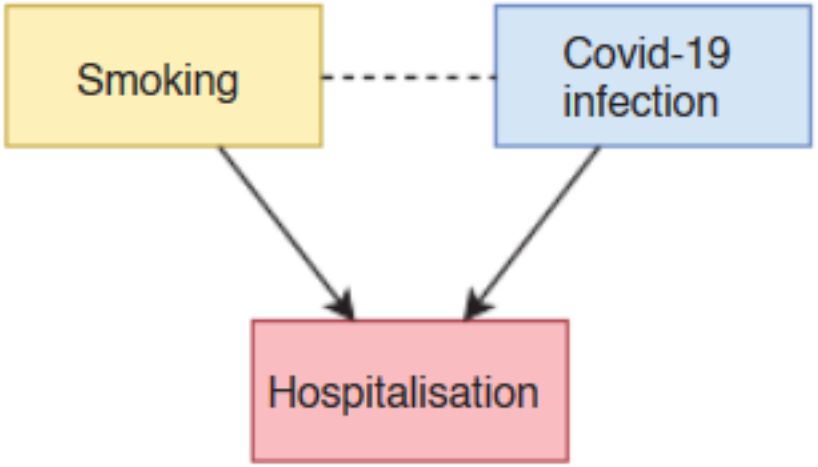
- Height and speed are uncorrelated in the population (Fig. 1)
- However, NBA basketball players must be either tall, or fast
- Thus height and speed become (negatively) correlated in them (Fig. 2)



Smoking might protect against Covid-19 infection – collider bias

The spectre of Berkson's paradox: Collider bias in Covid-19 research (Significance, 2020):

- https://rss.onlinelibrary.wiley.com/doi/10.1111/1740-9713.01413



UK healthcare workers appear to have milder COVID-19 infections compared to the general UK population

nature > nature communications > articles > article

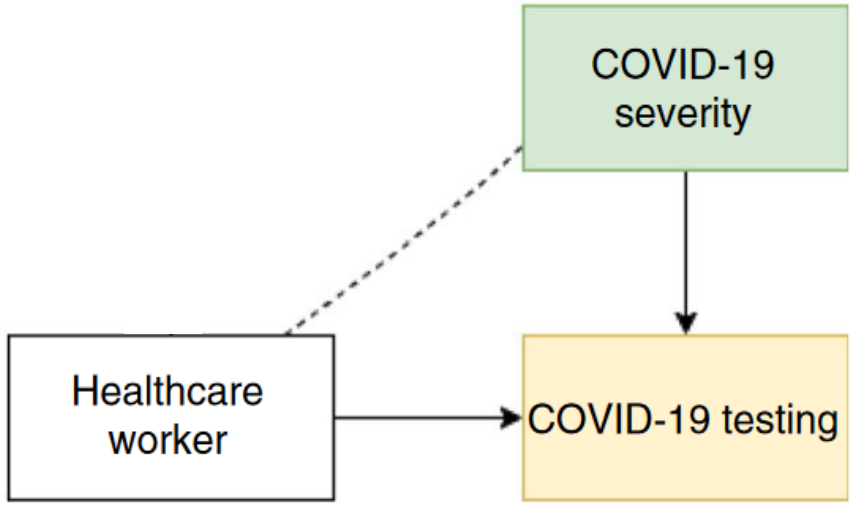
Article | Open Access | Published: 12 November 2020

Collider bias undermines our understanding of COVID-19 disease risk and severity

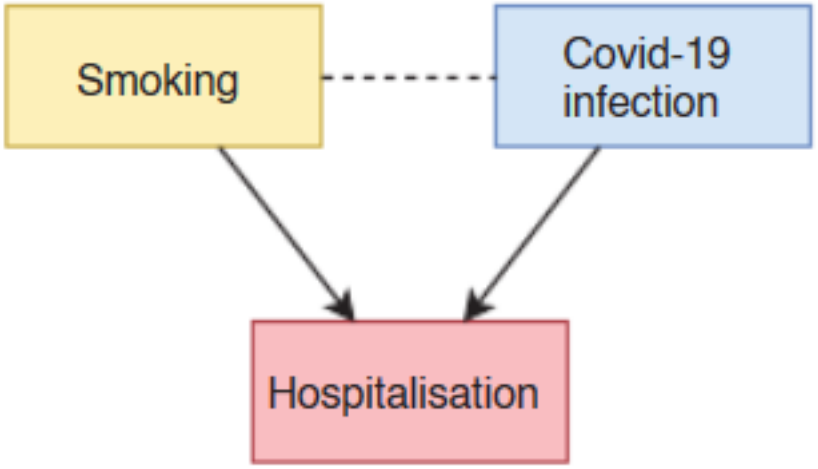
Gareth J. Griffith, Tim T. Morris, Matthew J. Tudball, Annie Herbert, Giulia Mancano, Lindsey Pike, Gemma C. Sharp, Jonathan Sterne, Tom M. Palmer, George Davey Smith, Kate Tilling, Luisa Zuccolo, Neil M. Davies & Gibran Hemani

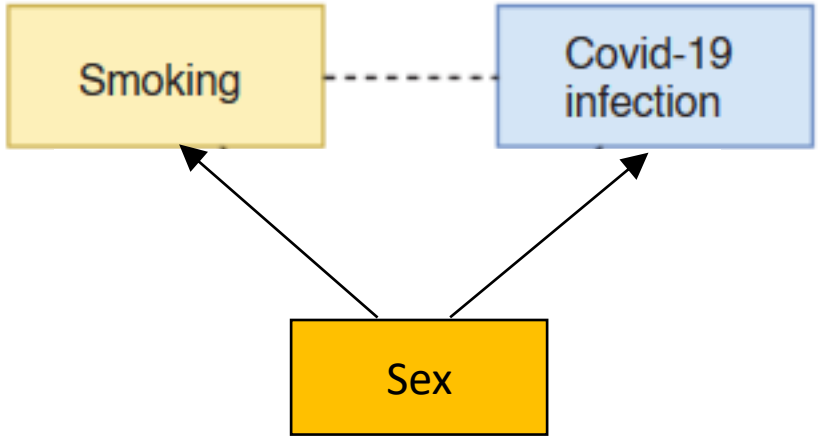
Nature Communications **11**, Article number: 5749 (2020) | Cite this article

58k Accesses | **357** Citations | **336** Altmetric | Metrics



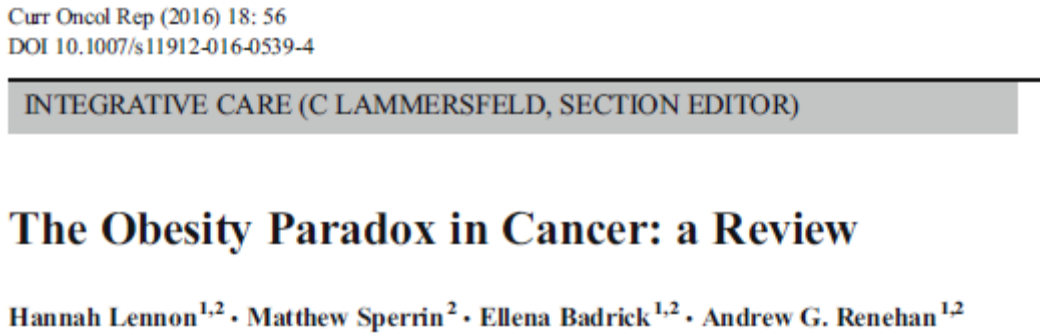
Subgroup analysis and collider bias





Obesity paradox

- ❑Refers to the counterintuitive observation that, for some diseases, overweight or obese individuals may have better outcomes or survival rates compared to those with normal weight, although initially overweight or obesity are risk factors for developing the disease
- ❑Observed for cardiovascular disease, heart failure, myocardial infarction, and several kinds of cancer, e.g. renal cancer, hematological cancers
- ❑Explanations:
  - ✓True biological reasons
  - ✓Artefact due to collider stratification



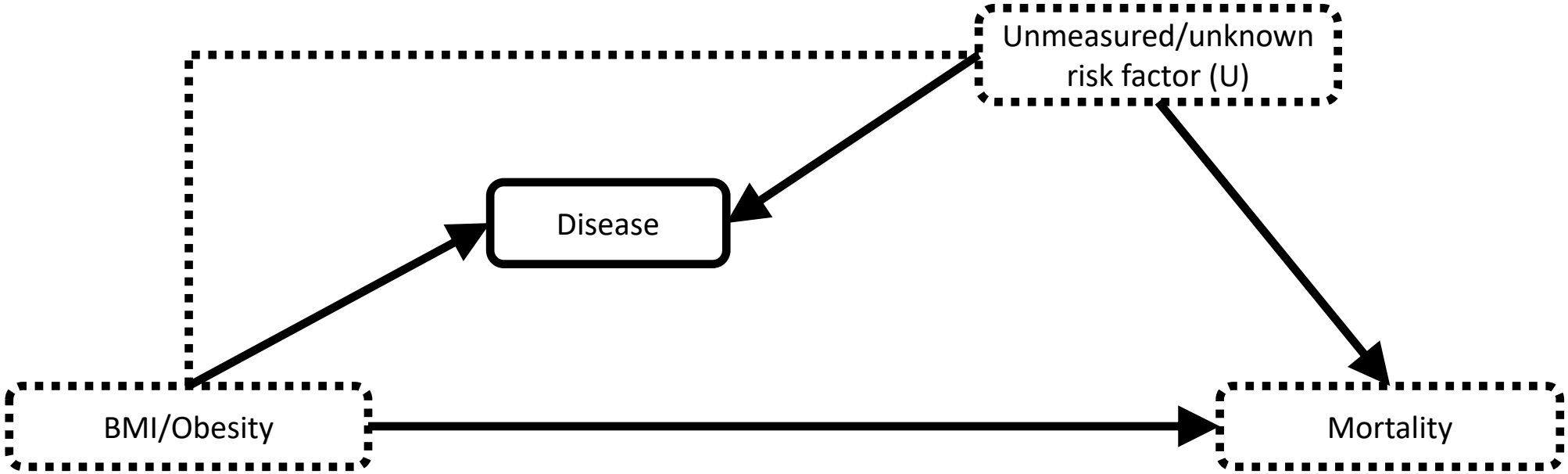
Curr Oncol Rep (2016) 18: 56
DOI 10.1007/s11912-016-0539-4

INTEGRATIVE CARE (C LAMMERSFELD, SECTION EDITOR)

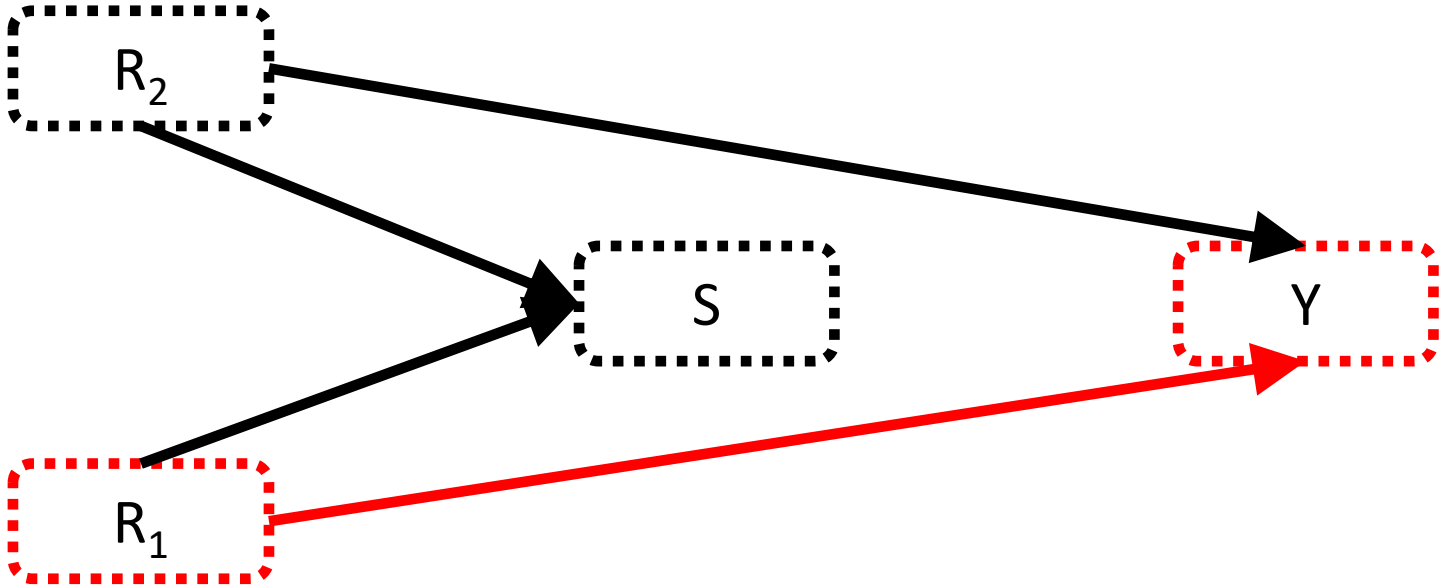
The Obesity Paradox in Cancer: a Review

Hannah Lennon[1,2] · Matthew Sperrin[2] · Ellena Badrick[1,2] · Andrew G. Renehan[1,2]

Collider bias in the obesity paradox



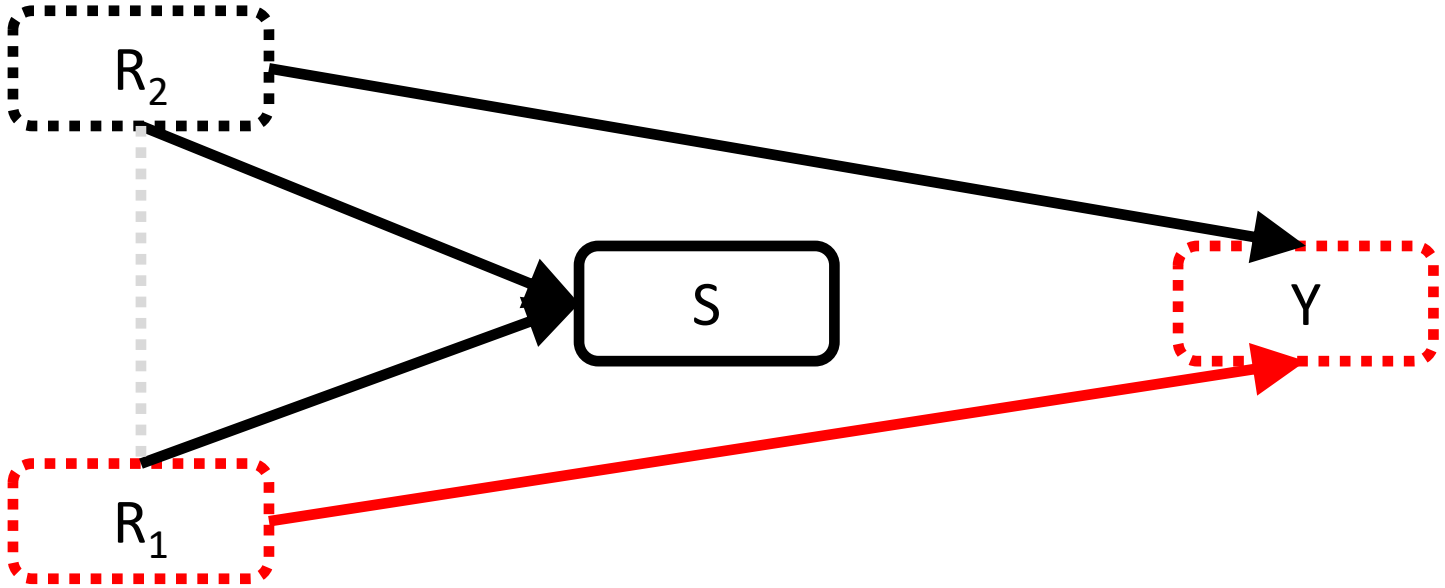
Collider bias – conceptual framework



<u>In the full sample:</u>

- regressing Y on R_1 w/o adjustment for R_2 gives an unbiased estimate of the causal effect

Collider bias – conceptual framework



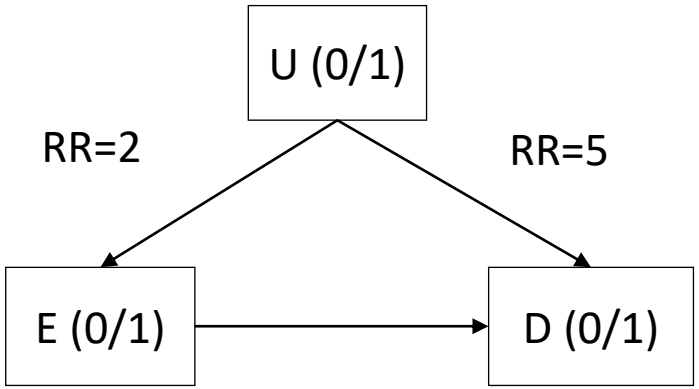
<u>In the full sample:</u>

- regressing Y on R_1 w/o adjustment for R_2 gives an unbiased estimate of the causal effect

<u>Data selected based on S:</u>

- Regressing Y on R_1 w/o adjustment for R_2 introduces bias
- Collider (stratification) bias
- By adjusting for R_2, this collider bias can be eliminated

Quantitative bias assessment via formulas



RR_{obs}=3

RR_{true}≥1.8 (i.e., 3/(2*5/(2+5-1)))

› Ann Epidemiol. 2008 Aug;18(8):637-46. doi: 10.1016/j.annepidem.2008.04.003.

Bias formulas for external adjustment and sensitivity analysis of unmeasured confounders

Onyebuchi A Arah [1], Yasutaka Chiba, Sander Greenland

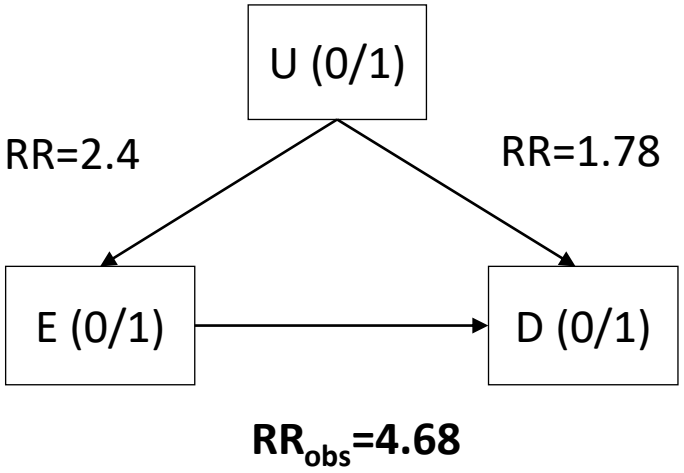
› Epidemiology. 2016 May;27(3):368-77. doi: 10.1097/EDE.0000000000000457.

Sensitivity Analysis Without Assumptions

Peng Ding [1], Tyler J VanderWeele

$$\mathrm{RR}_{ED}^{\mathrm{true}} \geq \mathrm{RR}_{ED}^{\mathrm{obs}} \Big/ \frac{\mathrm{RR}_{EU} \times \mathrm{RR}_{UD}}{\mathrm{RR}_{EU} + \mathrm{RR}_{UD} - 1}.$$

Quantitative bias assessment via simulations



RR_{obs}=4.68

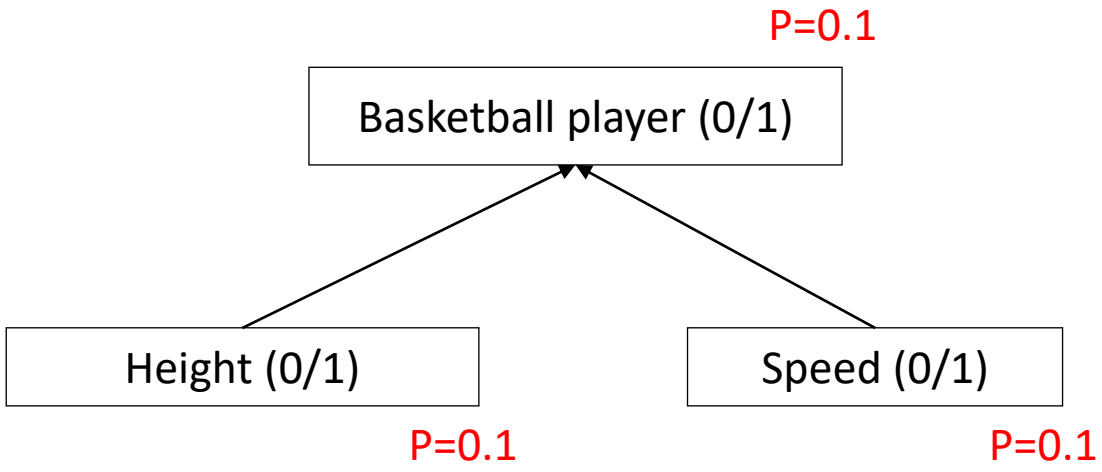
E	U	D	Frequency
0	0	0	90
0	0	1	10
0	1	0	35
0	1	1	15
1	0	0	70
1	0	1	30
1	1	0	40
1	1	1	360

RR_{true}≥2.61 (i.e., 4.68/1.80)

- ❑ Dataset with binary variables E, U, and D
- ❑ RRs as specified above
- ❑ Log-binomial regression model yields:
- ✓ **$RR(E,D)_{obs}$=4.68**
- ✓ **$RR(E,D)_{true}$=3.00**

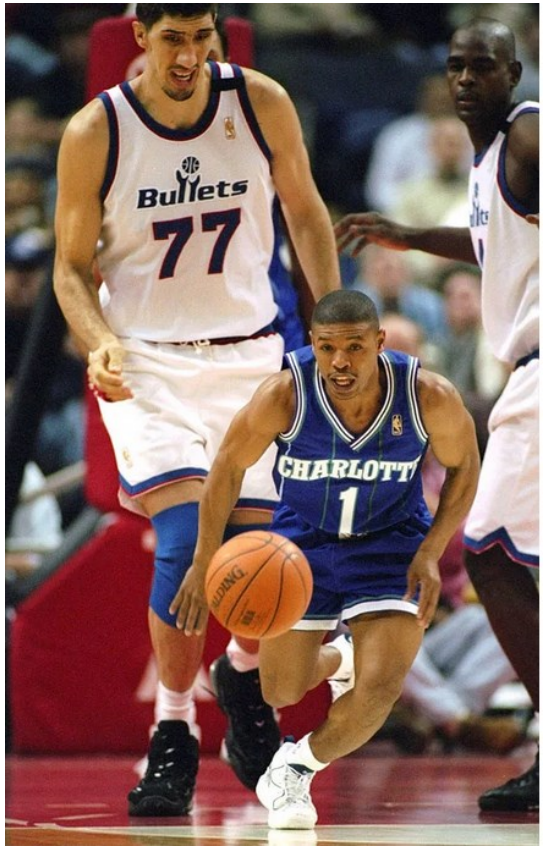
Simulation-based approach

Flexibility for a wide range of scenarios



- OR (Height, Basketball player) = 5
- OR (Speed, Basketball player) = 5
- **No interaction** on OR scale
- Then, in basketball players, height and speed are correlated with **ρ=-0.09** (assuming no correlation in the general population)

- OR (Height, Basketball player) = 5
- OR (Speed, Basketball player) = 5
- **Interaction** between Height and Speed of OR = 0.4
- Then, in basketball players, height and speed are correlated with **ρ=-0.22** (assuming no correlation in the general population)



Simulation-based approaches

The Obesity Paradox in Survival after Cancer Diagnosis: Tools for Evaluation of Potential Bias

Elizabeth Rose Mayeda and M. Maria Glymour

CEBP, 2017

Selection Bias as an Explanation for the Obesity Paradox

Just Because It's Possible Doesn't Mean It's Plausible

M. Maria Glymour and Eric Vittinghoff

Epidemiology, 2014

International Journal of Epidemiology, 2018, 226–235
doi: 10.1093/ije/dyx206
Advance Access Publication Date: 27 September 2017
Original article

Methods

Collider scope: when selection bias can substantially influence observed associations

Marcus R Munafò,[1,2]* Kate Tilling,[1,3] Amy E Taylor,[1,2] David M Evans,[1,4] and George Davey Smith[1,3]

Formula-based approaches

Bounding Bias Due to Selection

Louisa H. Smith[a] and Tyler J. VanderWeele[a,b]

Epidemiology, 2019

Collider Bias Is Only a Partial Explanation for the Obesity Paradox

Matthew Sperrin,[a] Jane Candlish,[a] Ellena Badrick,[a] Andrew Renehan,[b] and Iain Buchan[a]

Epidemiology, 2016

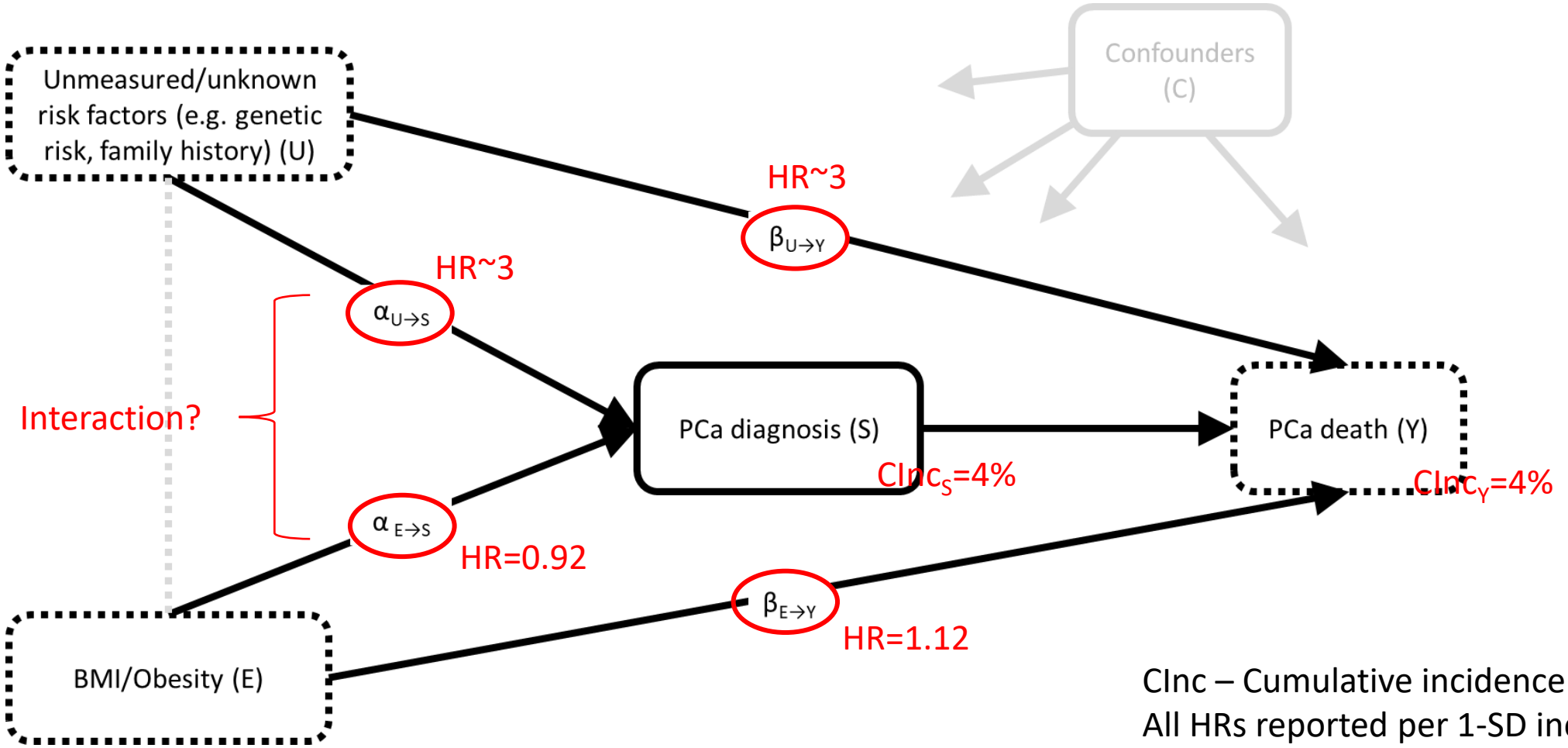
Original article

Does selection bias explain the obesity paradox among individuals with cardiovascular disease?

Hailey R. Banack MA*, Jay S. Kaufman PhD

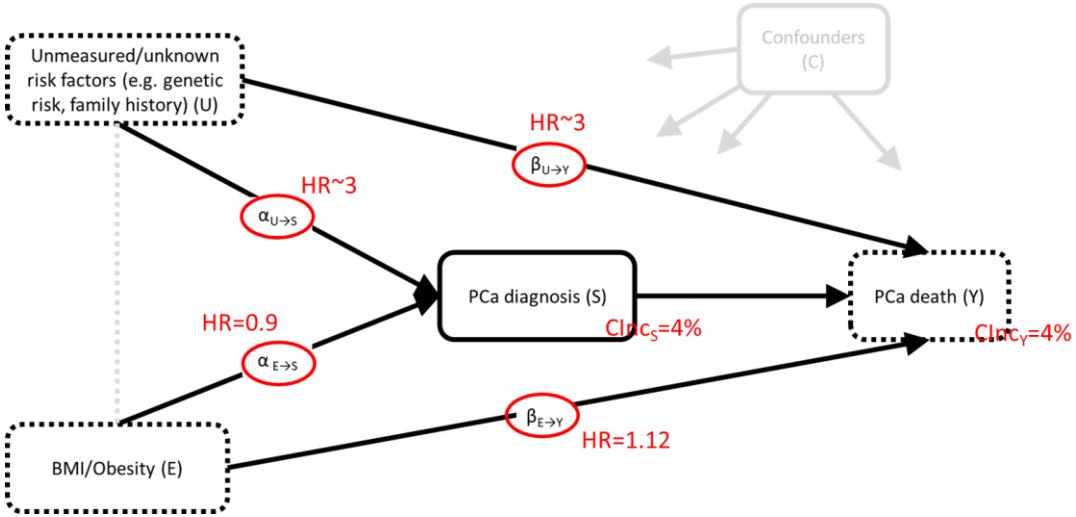
Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada

Ann. Epidemiol., 2015



CInc – Cumulative incidence
All HRs reported per 1-SD increase

- Jochems et al., Int J Cancer 2020
- Darst et al., Eur Urol 2021
- Pagadala et al., J Natl Cancer Inst 2023



- ✓ Simulate data according to plausible input parameters and do the following:
- ✓ Amongst the subgroup (PCa cases):
  - ❑ Calculate HR of Y on E – unadjusted for U (biased): HR_1
  - ❑ Calculated HR of Y on E – adjusted for U (unbiased): HR_2
- ✓ Non-collapsibility of the HR: Marginalization of HR_2 (Daniel et al., *Biometrical J* 2021) → $HR_{2,Marg}$
- ✓ Percentage bias (PB): $(HR_1 - HR_{2,Marg})/HR_{2,Marg} \times 100$
- ✓ In the simulations: binary variables instead of time-to-event data (ORs instead of HRs)

Simulation of plausible BMI – PCa scenarios

- E – normally distributed (continuous BMI)
- S and Y – binary with cumulative incidences of 4% each
- OR(E,S)=0.92 (per 1-SD)
- OR(E,Y)=1.12 (per 1-SD)
- U normally distributed

Scenario	OR(U→S)	OR(U→Y)	Interaction OR	Collider-biased OR	Un-biased OR	Percentage bias
1	3	3	No	1.124	1.116	0.8%
2	10	10	No	1.121	1.104	1.6%
3	3	3	1.04	1.150	1.114	3.2%
4	3	3	2	1.407	1.121	25.6%
5	3	3	0.5	0.886	1.120	-20.8%

Conclusion

- **Collider stratification bias is unlikely to relevantly affect the positive association between BMI and PCa-specific mortality as observed in analyses of localized PCa cases only**
- Main reason: the association of BMI with risk of localized PCa (HR per 5-kg/m2 ~0.9) is too small
  - For renal cancer: collider bias might be large enough to explain the obesity paradox there (Mayeda & Glymour, CEBP 2017)

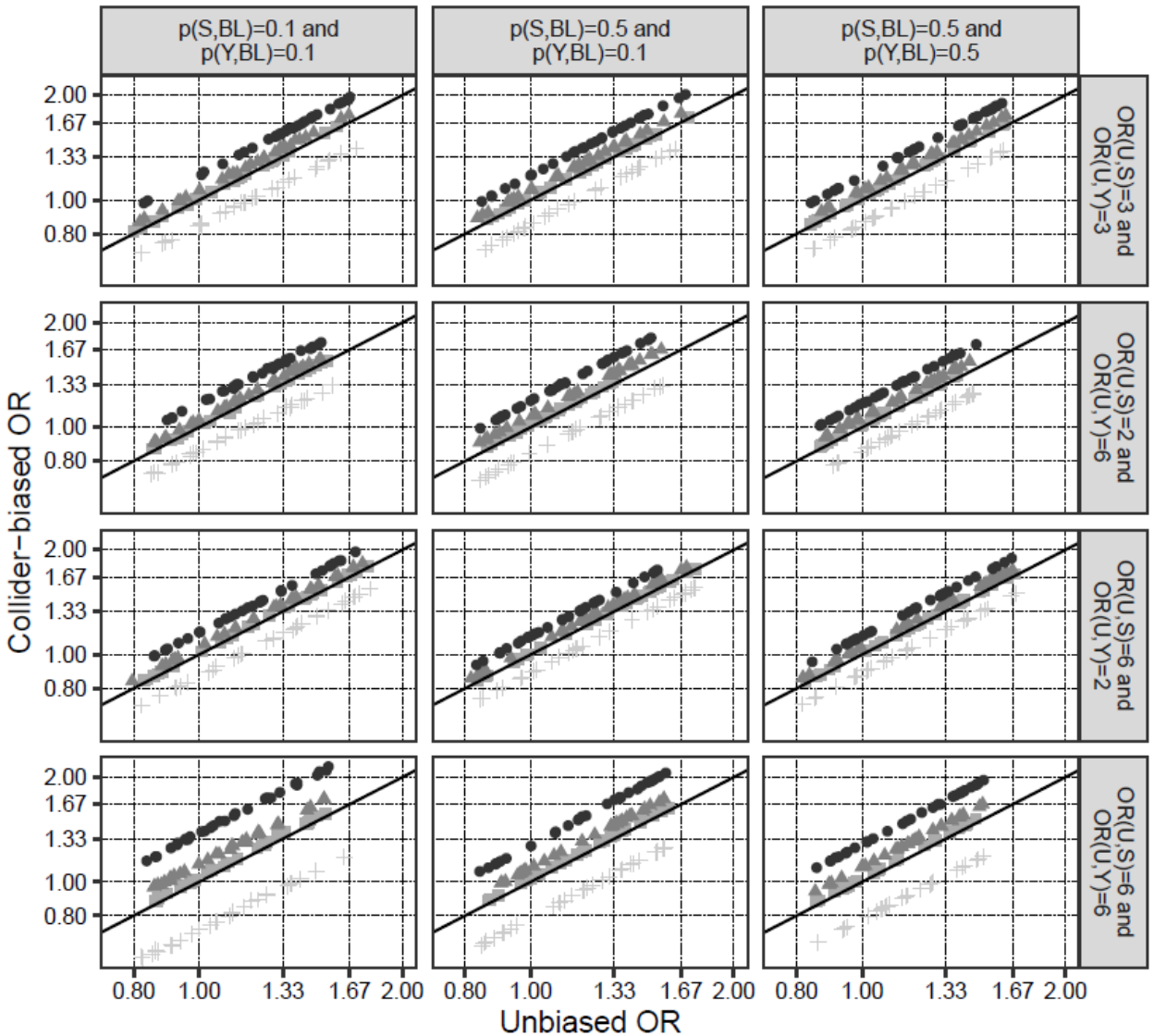
Summary

- Collider stratification bias
- Qualitative vs. quantitative assessment
- Analytical formulas
- Flexibility of simulation-based approach
  - Different data types
  - More complicated relationships (e.g. interactions)
- Testing for robustness of findings & sensitivity analyses
- Study about collider bias simulation and inverse obesity paradox in PCa soon to be published

Any questions?

Back-up slides

$CInc_S$, $CInc_Y$



$OR_{U\to S}$, $OR_{U\to Y}$

Sensitivity of results regarding presence of and/or strength of interaction