Causal mediation analysis of observational, population-based cancer survival data

Bernard Rachet & Ruoran Li
Cancer Survival Group, Faculty of Epidemiology and Population Health, LSHTM

Twitter: @CSG_LSHTM

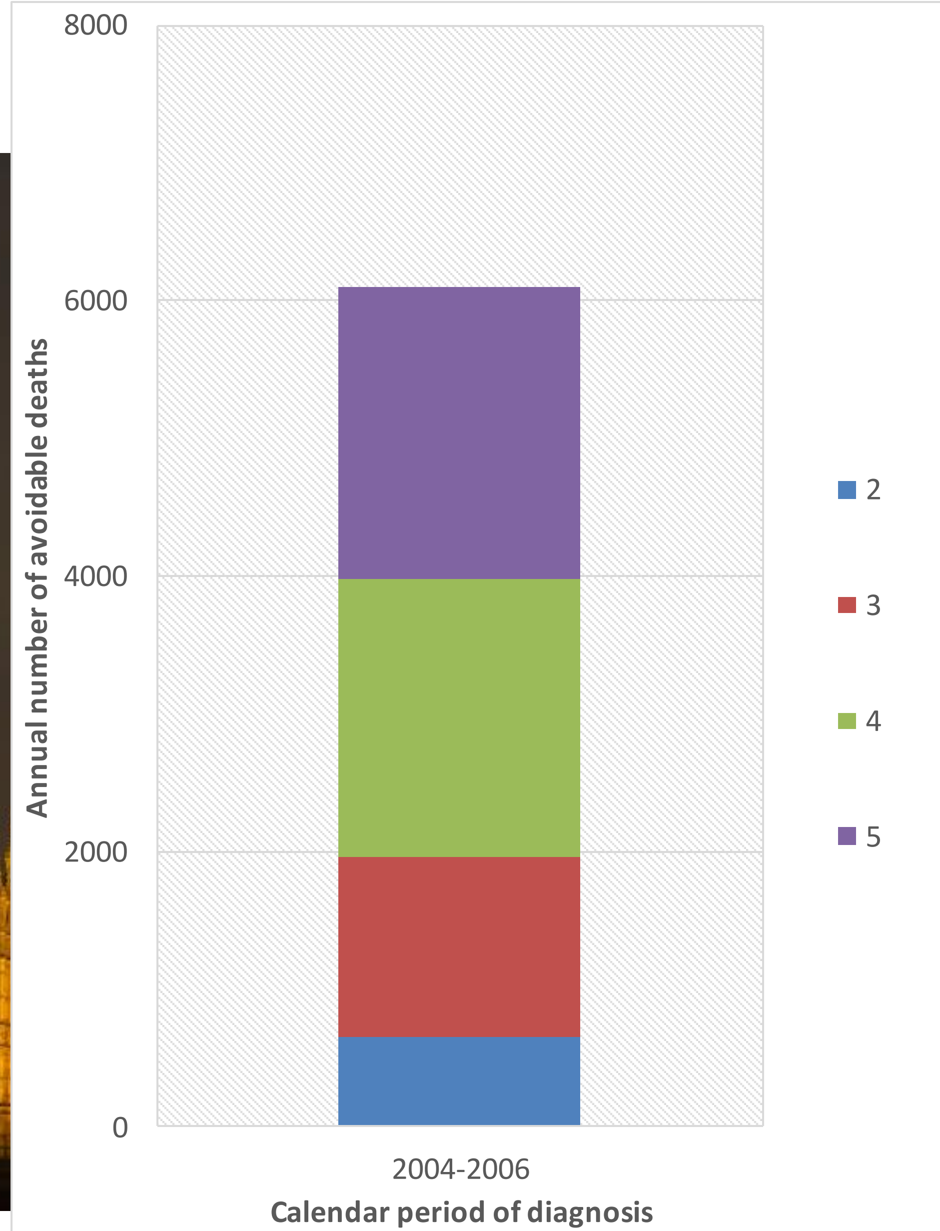
LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

Outline

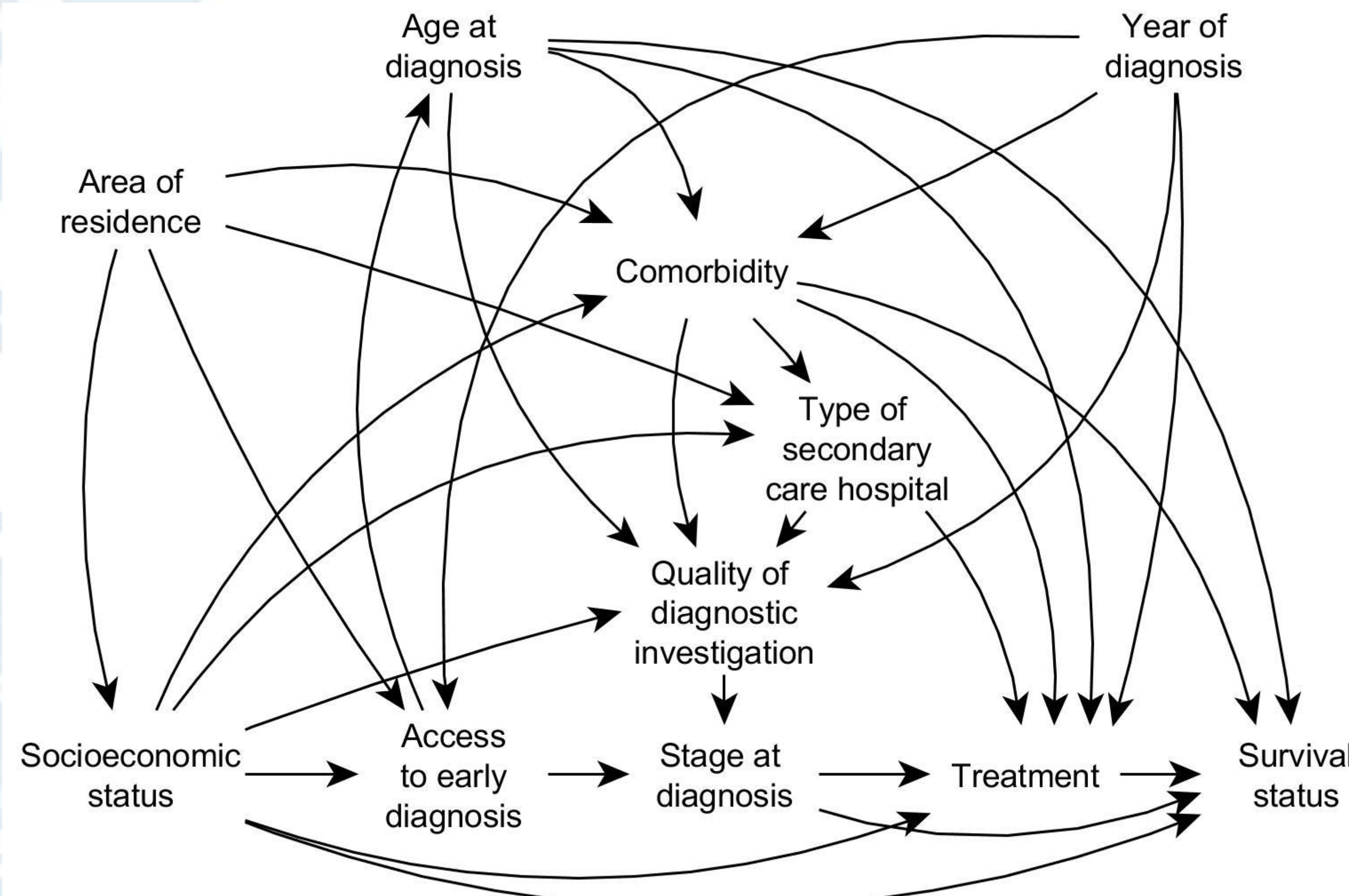
- Main questions
- Challenges with conventional approaches – results from the past
- An example of applying mediation to cancer survival data
- Problems and discussions
  - Misclassification of mediators
    - Treatment missing for more affluent – sensitivity analysis
    - Under-staged deprived patients – sensitivity analysis
  - Biases for mediation analyses
  - Controlled and natural effects
  - Conceptual frameworks – Suggestions?
    - Including issue of diagnosed stage *versus* real stage

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Inequalities in cancer survival



Explaining inequalities



Challenges in the past

- **More deprived patients:**
  - More comorbidity
  - More advanced cancer at diagnosis (colon, rectum, breast)
  - More often diagnosed during emergency admission
  - More often treated in non-specialised hospital and by non-specialised surgeon
  - Received more often sub-optimal and delayed treatment (colon, rectum)

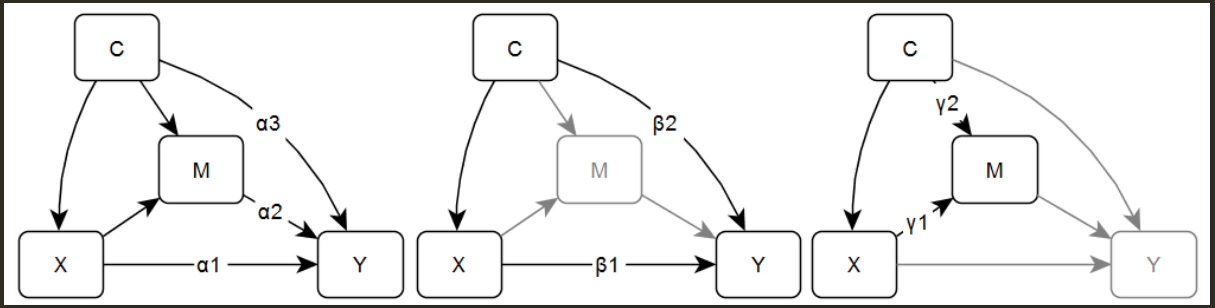
- **Past conventional analysis (colon, rectum, breast)**
  - No excess mortality hazard for deprivation among those treated within one month since diagnosis
  - Adjusting for comorbidity did not modify the excess mortality hazard for deprivation
  - Adjusting for stage reduced the excess mortality hazard for deprivation by less than a third

  - Limited stage and treatment data and conventional analytic approaches did not enable identification of mechanisms underlying deprivation gap in survival

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

NB: explain DAG

TRADITIONAL MEDIATION ANALYSIS

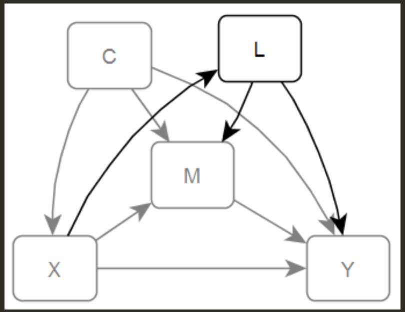


Difference method (Baron and Kenny, 1986)

Product method (Wright,1921)

Problems

- Definition of effects model-dependent
- Inflexible: interaction & non-linearity
- Intermediate confounder



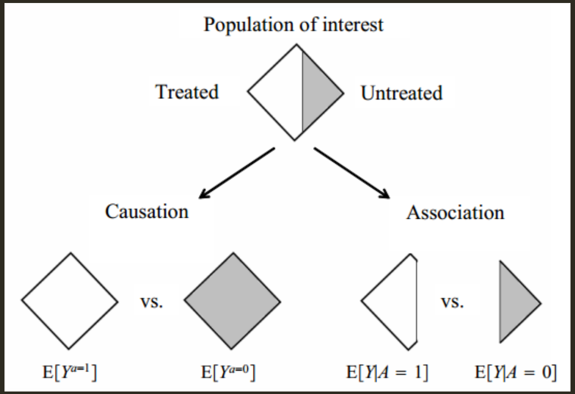
COUNTERFACTUAL APPROACHES

	A	Y
Rheia	0	0
Kronos	0	1
Demeter	0	0
Hades	0	0
Hestia	1	0
Poseidon	1	0
Hera	1	0
Zeus	1	1
Artemis	0	1
Apollo	0	1
Leto	0	0
Ares	1	1
Athena	1	1
Hephaestus	1	1
Aphrodite	1	1
Cyclope	1	1
Persephone	1	1
Hermes	1	0
Hebe	1	0
Dionysus	1	0

Observed

	$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1
Kronos	1	0
Demeter	0	0
Hades	0	0
Hestia	0	0
Poseidon	1	0
Hera	0	0
Zeus	0	1
Artemis	1	1
Apollo	1	0
Leto	0	1
Ares	1	1
Athena	1	1
Hephaestus	0	1
Aphrodite	0	1
Cyclope	0	1
Persephone	1	1
Hermes	1	0
Hebe	1	0
Dionysus	1	0

Counterfactual worlds



- Notations: Y(x), Y(x,m), Y(x,M(x))

From: http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/

CAUSAL APPROACHES ALLOWS MODEL-FREE DEFINITION OF EFFECTS…

Total causal effect

- TCE=E(Y[1,M(1)])-E(Y[0,M(0)])

Natural direct effect

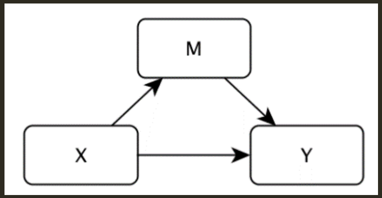
- NDE(0)=E(Y[1,M(0)])-E(Y[0,M(0)])

Natural indirect effect

- NIE(1)=E(Y[1,M(1)])-E(Y[1,M(0)])

Controlled direct effect

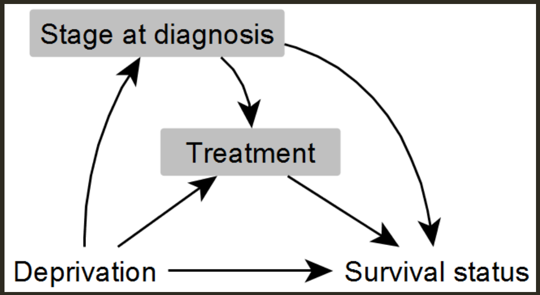
- CDE(m)=E(Y[1,m])-E(Y[0,m])



Assumptions for identification

1: no unmeasured confoundings
2: no exposure induced M/Y confounder (L)

BUT WE HAVE L…



- Important mediator-outcome confounders affected by exposure
- Likely presence of many interactions
- Binary outcome

One of the solutions proposed in VanderWeele, Vansteelandt and Robins (Epidemiology 2014)

Interventional effect

- Randomized interventional analogues of natural direct and indirect effects
- Estimated with an extension of Robins' g-computation formula implemented using Monte Carlo simulation
- Similar definitions to NIE and NDE

Outline

- Main question
- Challenges with conventional approaches – results from the past
- An example of applying mediation to cancer survival data
- Problems and discussions
  - Misclassification of mediators
    - Treatment missing for more affluent – sensitivity analysis
    - Under-staged deprived patients – sensitivity analysis
  - Biases for mediation analyses
  - Controlled and natural effects
  - Conceptual frameworks – Suggestions?
    - Including issue of diagnosed stage vs real stage

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

How much of the socioeconomic differences in breast cancer patient survival can be explained by stage at diagnosis and treatment?

Application of causal mediation analysis to routine data

Ruoran Li, Rhian Daniel, Bernard Rachet
Faculty of Epidemiology and Population Health
London School of Hygiene & Tropical Medicine

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

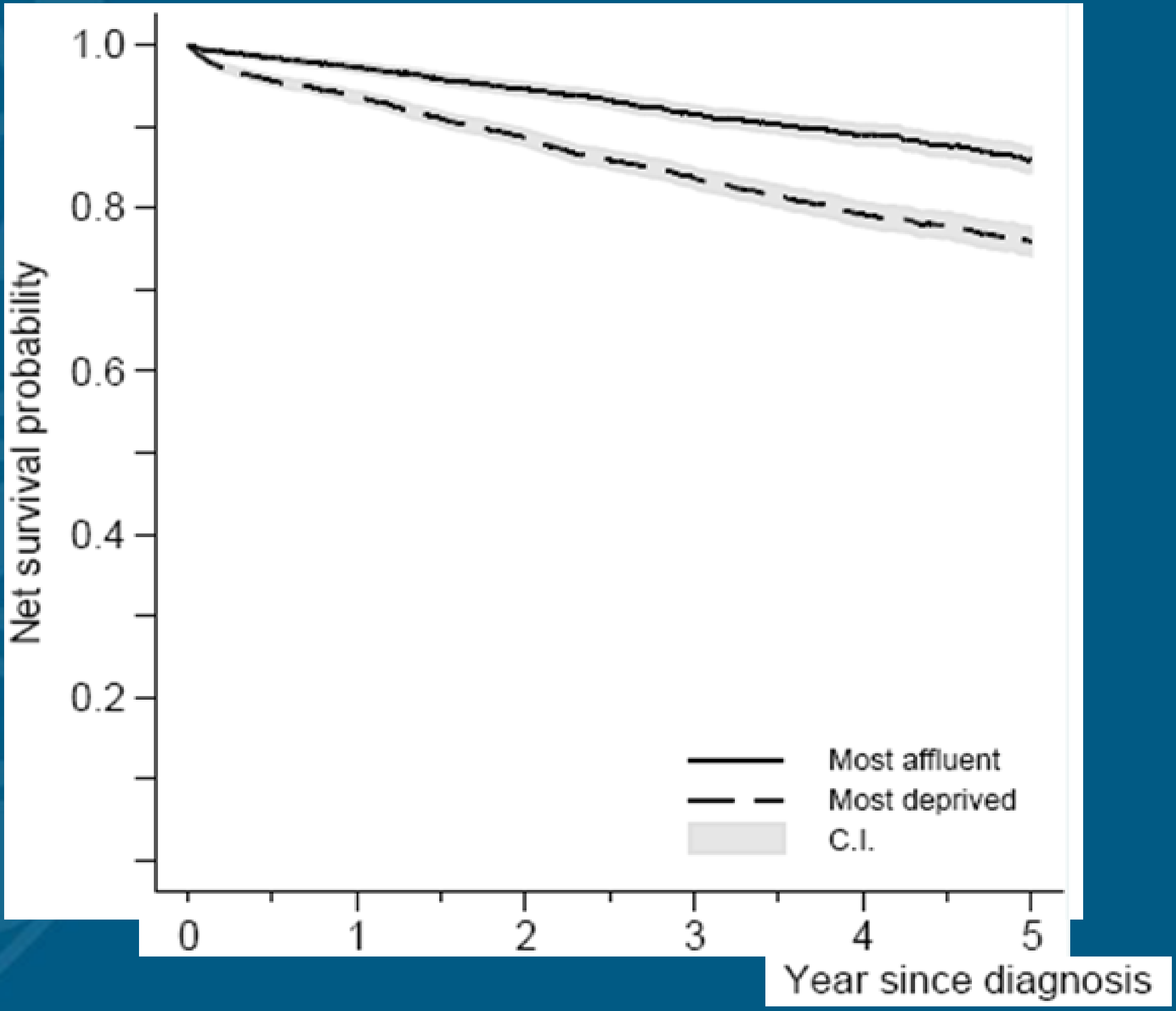
Introducing breast cancer

- Most common cancer in the UK
- Screening (50-70)
- Treatment with strict guidelines

- Northern and Yorkshire Cancer Registry, population-based, covering 12% of the English population
- Women with malignant breast cancers (N=36,793)
  - Diagnosed during the period 2000–2007
  - Followed up until 31 December 2007

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

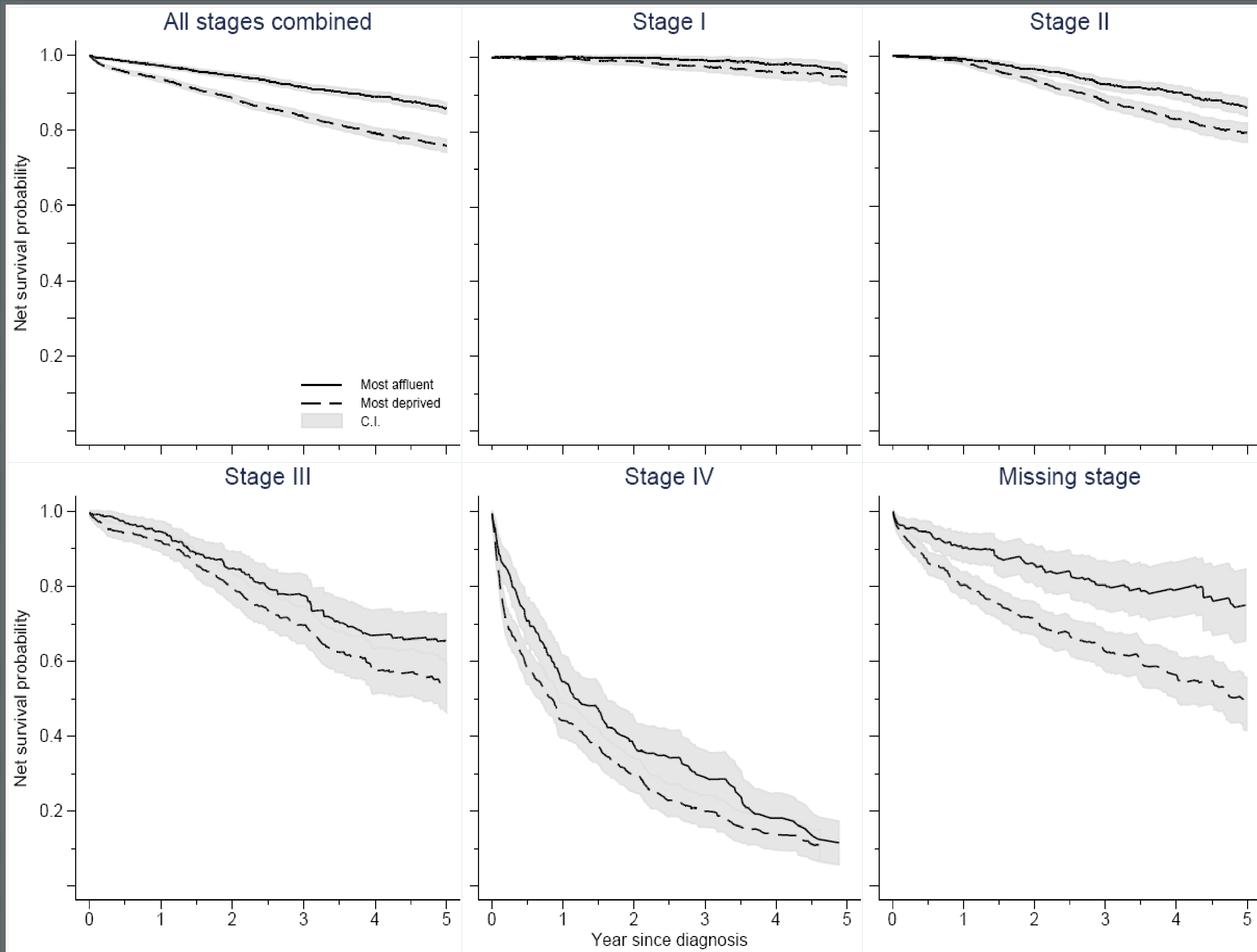
Large deprivation gap in survival from breast cancer…



Possible explanations

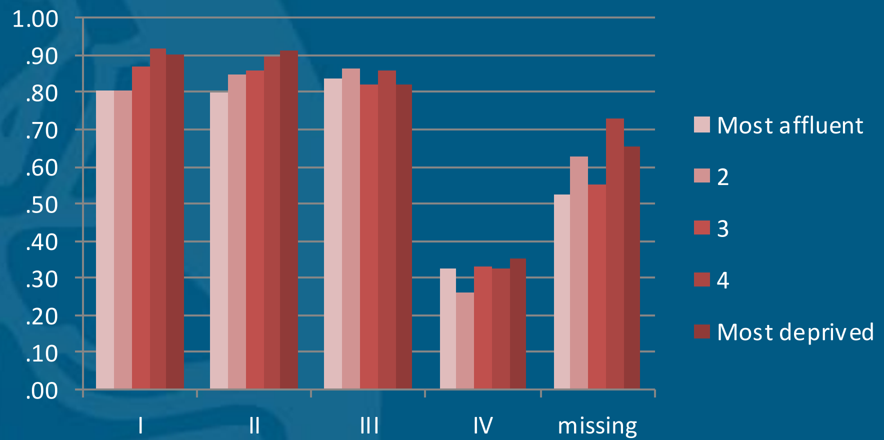
- Differential stage at diagnosis?
- Differential treatment?

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE


All stages combined
Stage I
Stage II
Stage III
Stage IV
Missing stage
Net survival probability
1.0
0.8
0.6
0.4
0.2
Most affluent
Most deprived
C.I.
0
1
2
3
4
5
Year since diagnosis

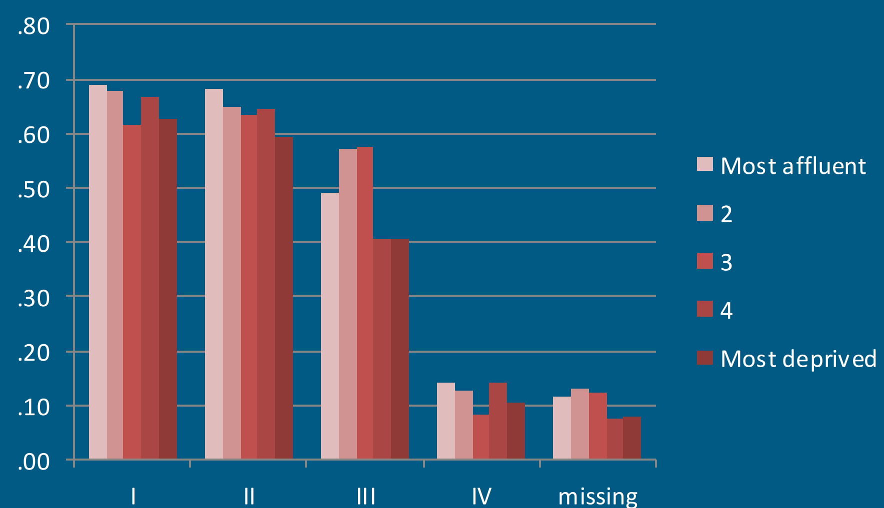
Differential treatment?
– probability of getting major surgery

15-49 pre-screening



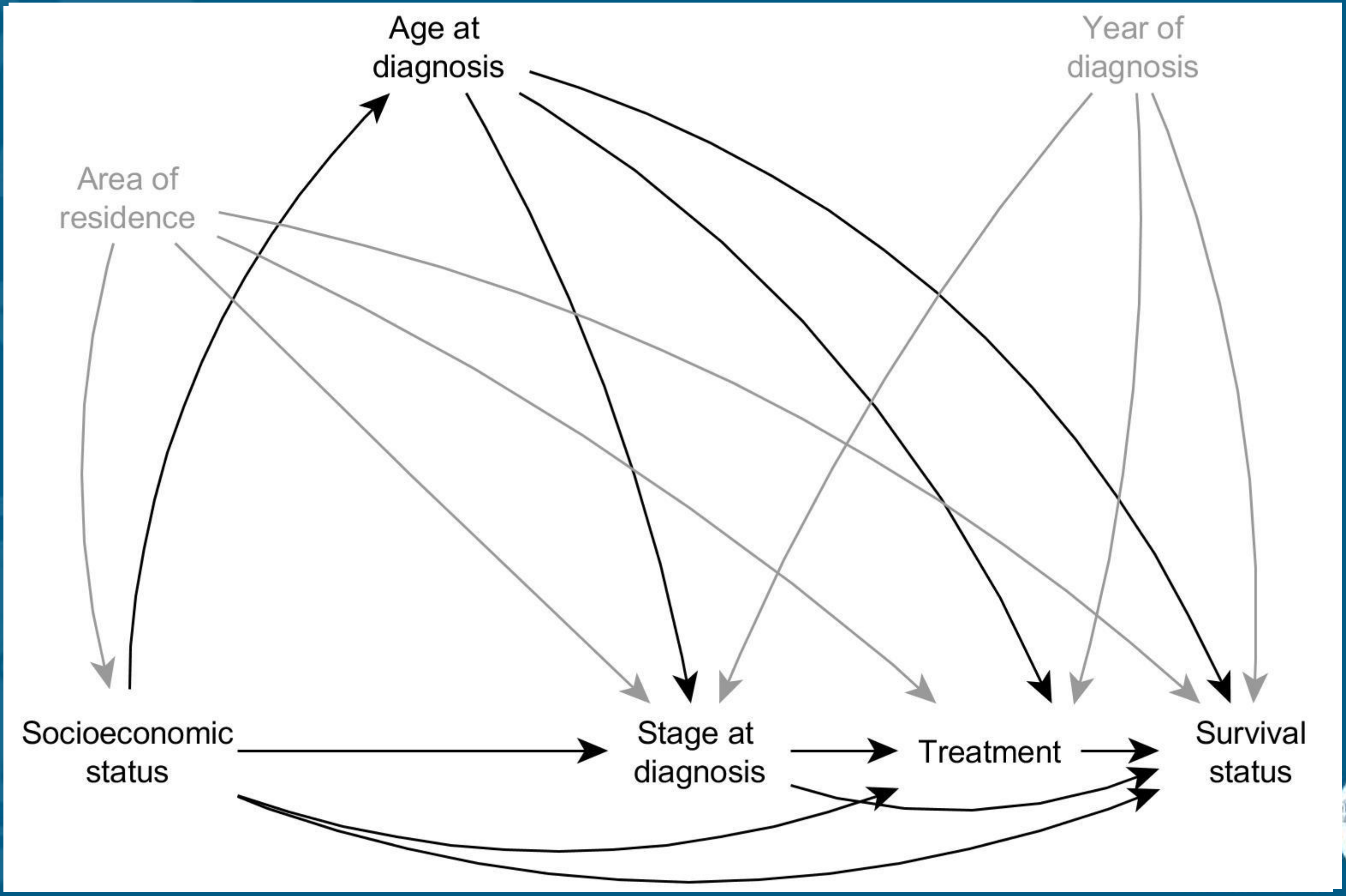
50-69 screening

70+
post-screening

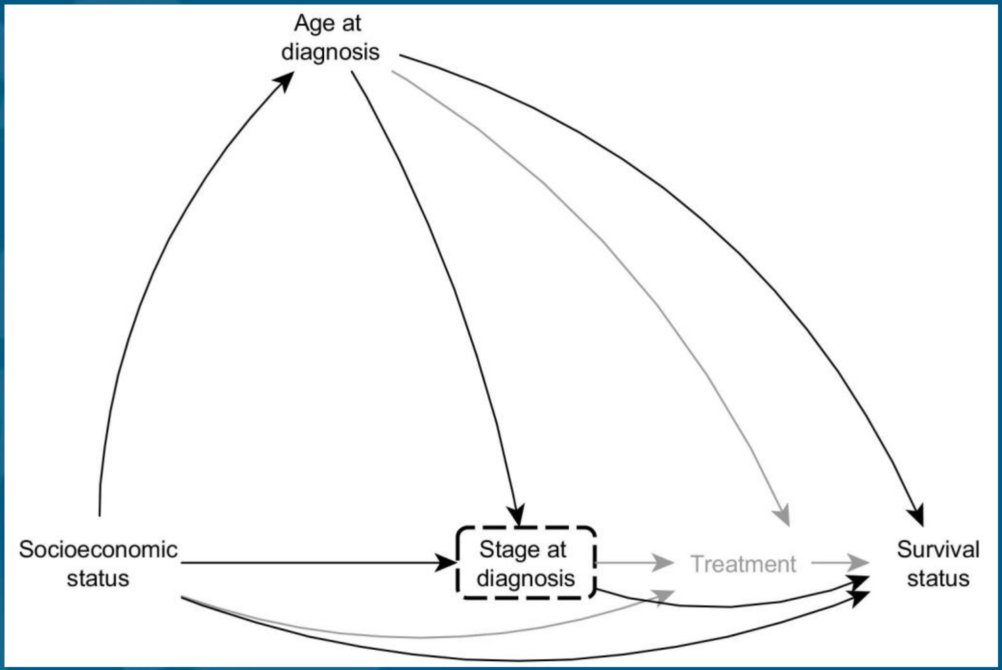


LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Linking to the conceptual diagram…



If we look at stage



We can **decompose** the total effect (TCE) of socioeconomic status (deprivation) on mortality into…

- Those mediated by stage (The indirect effect, NIE)

- Those not mediated by stage (The direct effect, NDE)

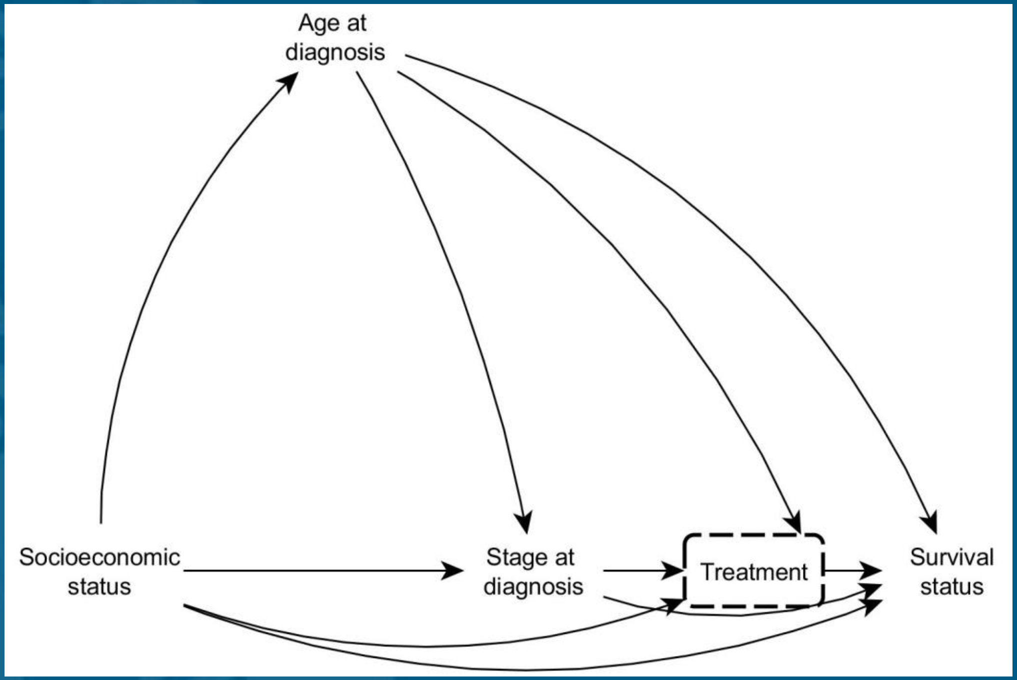
TCE = log(odds(Y(Dep=*most*, Stage(Dep=*most*)))) – log(odds(Y(Dep=*least*, Stage(Dep=*least*))))

NIE = log(odds(Y(Dep=*most*, Stage(Dep=*most*)))) – log(odds(Y(Dep=*most*, Stage(Dep=*least*))))

NDE = log(odds(Y(Dep=*most*, Stage(Dep=*least*)))) – log(odds(Y(Dep=*least*, Stage(Dep=*least*))))

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

If we look at treatment



We can **decompose** the total effect (TCE) of deprivation on mortality into…

- Those mediated by treatment (The indirect effect, NIE)

- Those not mediated by treatment (The direct effect, NDE)

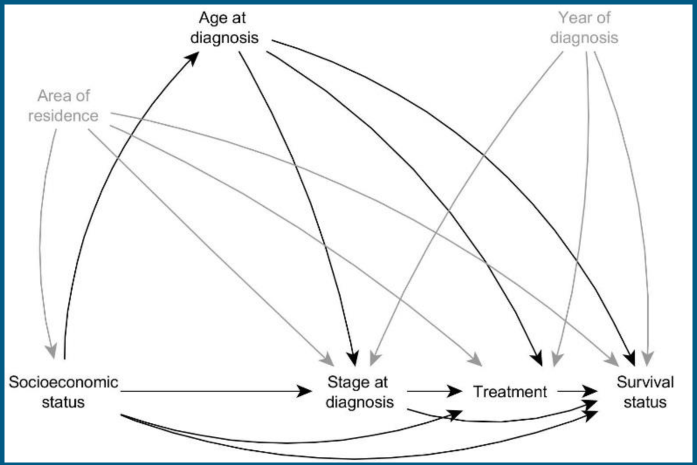
TCE = log(odds(Y(Dep=*most*, Treat(Dep=*most*)))) – log(odds(Y(Dep=*least*, Treat(Dep=*least*))))

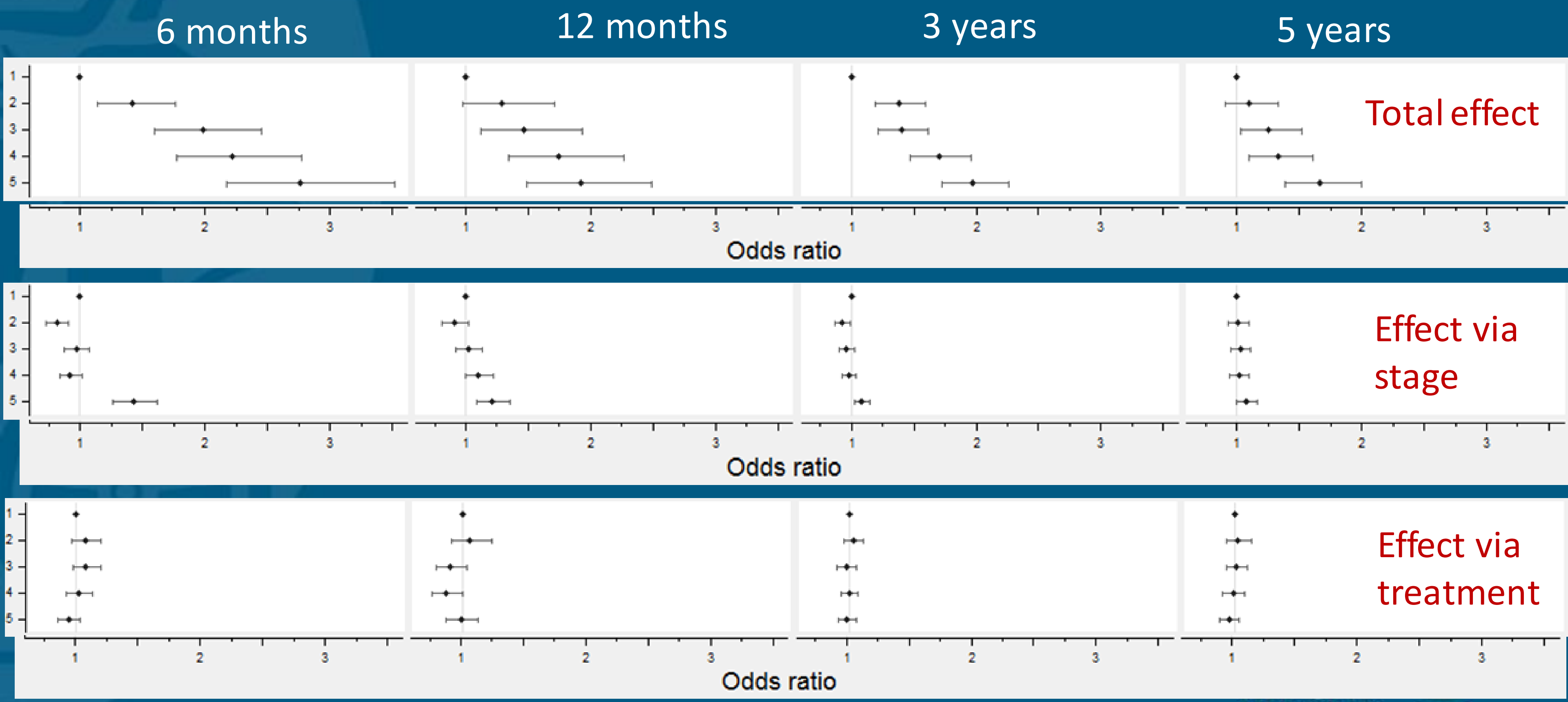
NIE = log(odds(Y(Dep=*most*, Treat(Dep=*most*)))) – log(odds(Y(Dep=*most*, Treat(Dep=*least*))))

NDE = log(odds(Y(Dep=*most*, Treat(Dep=*least*)))) – log(odds(Y(Dep=*least*, Treat(Dep=*least*))))

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

G-formula results





Preliminary conclusions

- Breast cancer survival differed between the most deprived and most affluent patients

- Effect of deprivation on mortality:
  - Large total effect **FOR ALL DEPRIVATION CATEGORIES**:
    - Increasing with deprivation
    - Decreasing with time since diagnosis
  - Mediated via stage **ONLY FOR MOST DEPRIVED CATEGORY**:
    - One third of at six months
    - One tenth at three/five years since diagnosis
  - Mediated via treatment:
    - None

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

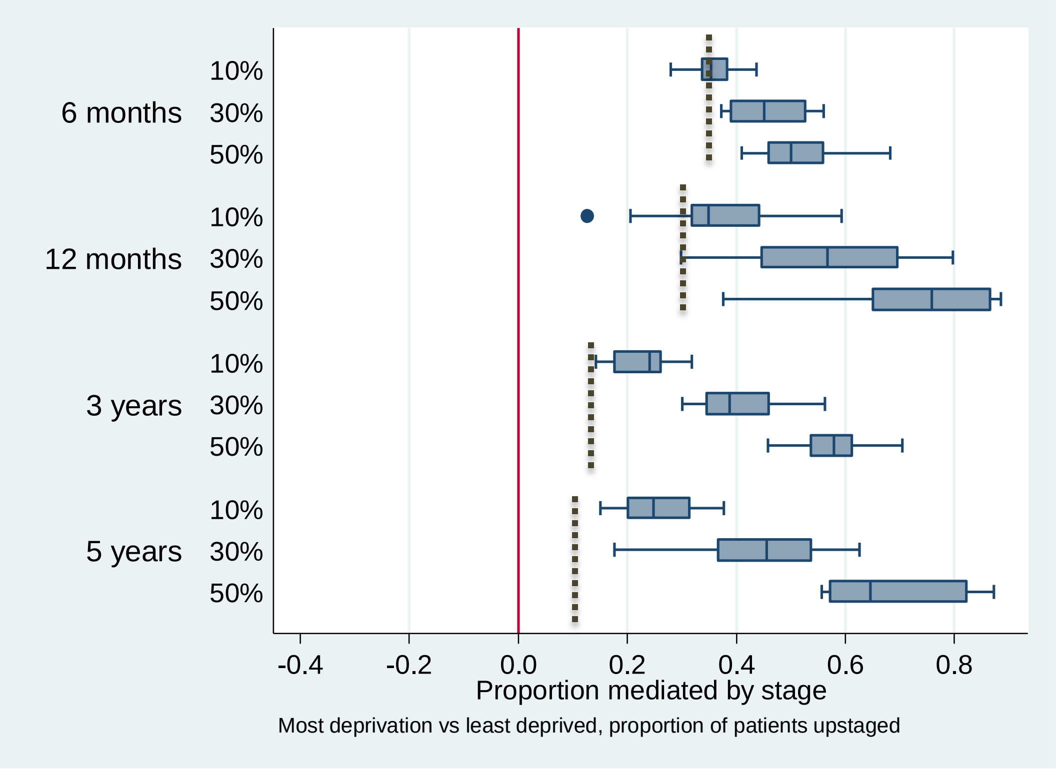
Outline

- Main questions
- Challenges with conventional approaches – results from the past
- An example of applying mediation to cancer survival data
- Problems and discussions
  - Misclassification of mediators
    - Treatment missing for more affluent – sensitivity analysis
    - Under-staged deprived patients – sensitivity analysis
  - Biases for mediation analyses
  - Controlled and natural effects
  - Conceptual frameworks – Suggestions?
    - Including issue of diagnosed stage vs real stage

Misclassification of stage

- More deprived patients may be under-staged?
- Randomly “up-staging” 10%, 30% and 50% of most deprived patients…

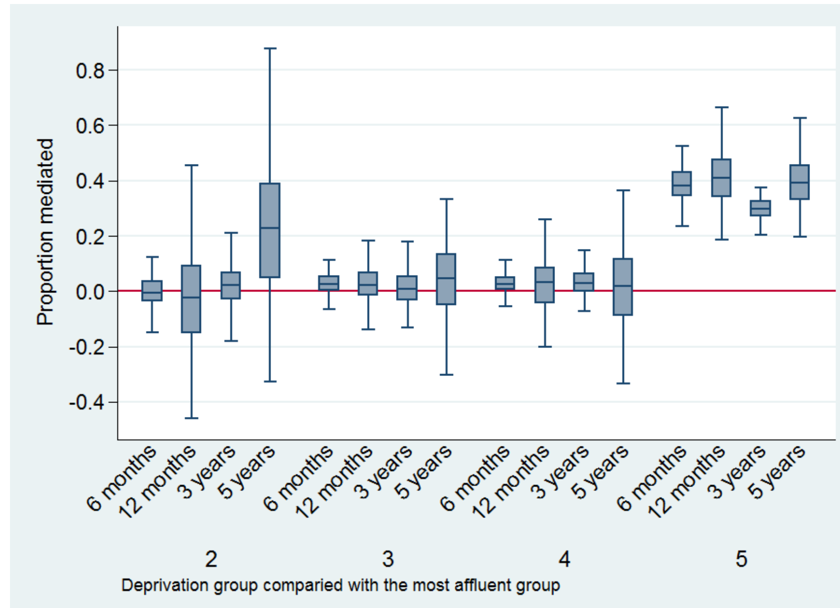
- 10% up-staging did not change results much
- After 30%-50% upstaging, stage would mediate more than half of the survival differences
- Longer-term survival is more affected …



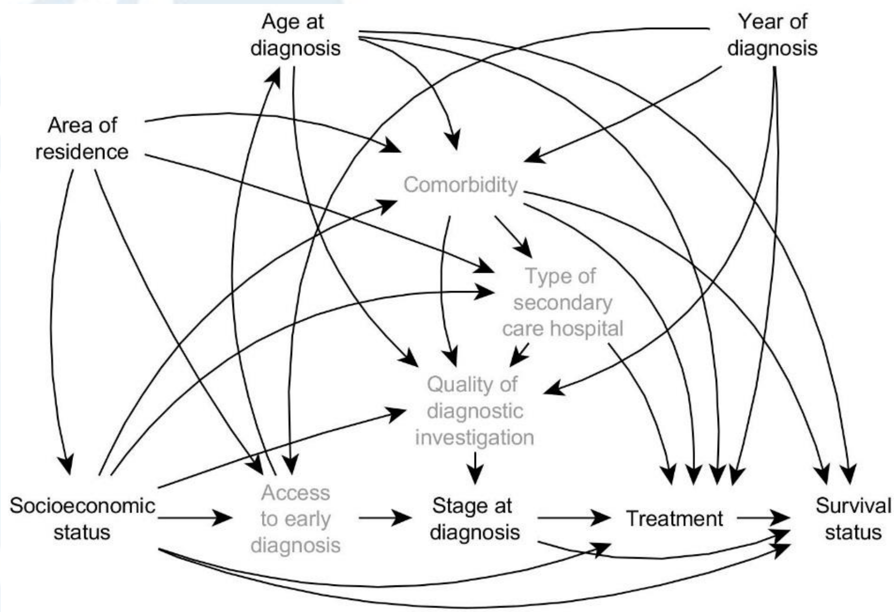
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Misclassification of treatment

- One report showed that 4% of surgical treatment for breast cancer were made in private hospital
- Sensitivity analysis:
  - Assumption: all missing surgery is among most affluent patients
  - Randomly adding “major surgery” to 4% of women, all from the most affluent category
- Now treatment mediates survival differences for the most deprived!

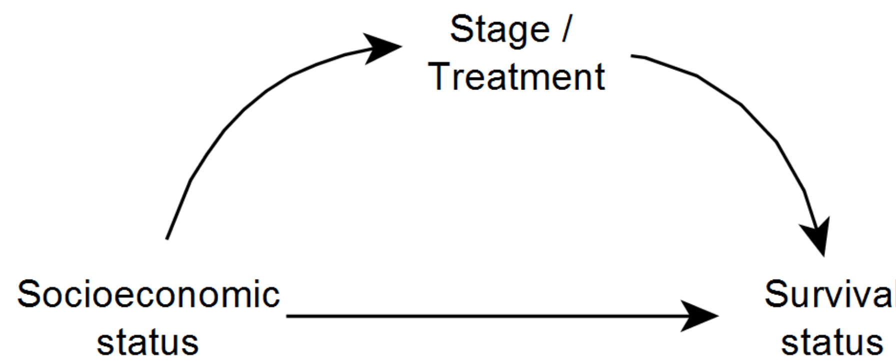


Biases for mediation analysis



- Unmeasured or poorly measured confounders, e.g. between mediator and outcome?
- Presence of confounder(s) between mediator and outcome affected by exposure?

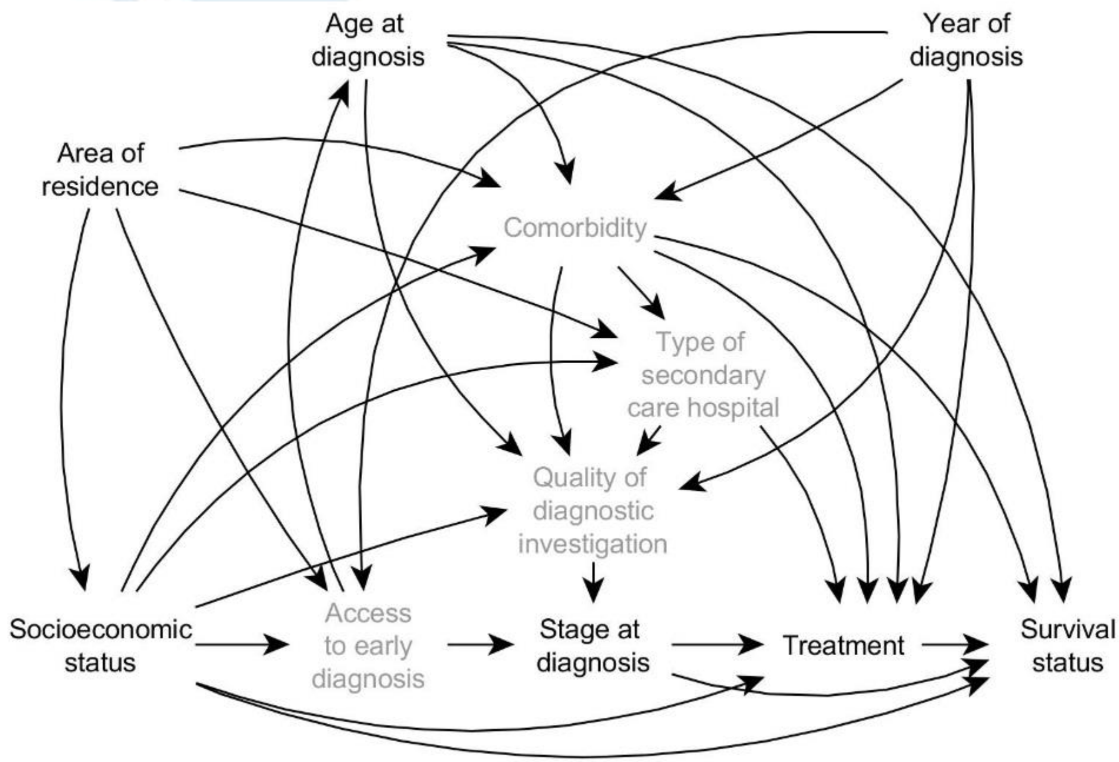
Controlled *vs* natural effects



- Natural direct effect
  - What effect would SES have on survival status if the more deprived patients had the stage/treatment distribution of the most affluent patients?
  - It measures delays in diagnosis (stage) or inequities in management (treatment)
- Controlled direct effect
  - What effect does SES have on survival status if we intervened on everyone's diagnosed stage/treatment and set it to a particular level?
  - More sensible to estimate CDE for compliance to treatment guideline?
    - Classify treatment as *compliant to guideline (Yes/No)* according to detailed patient and tumour characteristics
    - Mediator = compliance to guideline

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

Conceptual framework



- Among more deprived patients:
  - Sub-optimal diagnostic investigation
  - Wider discrepancy between true and observed stage
- How to account for this stage misclassification?

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Summary

- First application of the causal mediation tool in study of cancer registry data
- Population-based data
- Drawbacks
  - Data quality and detail
  - Unmeasured confounder, e.g. comorbidity
- Useful for answering questions related to causality
  - Resource allocation

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

References

- Woods L. M., Rachet B., Coleman M. P. 2005 Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol 17(1):5-19
- Daniel, R. M., De Stavola, B. L., and Cousens, S. N. 2011. gformula: Estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula. The Stata Journal 11(4):479-517.
- Hernán M. A., Robins J. M. Causal Inference. Part II Causal inference with models http://www.hsph.harvard.edu/miguel-hernan/files/2013/10/hernanrobins_v2.15.02.pdf [updated 15 October 2013]
- VanderWeele TJ, Vansteelandt S, Robins JM. Effect Decomposition in the Presence of an Exposure-Induced Mediator-Outcome Confounder. Epidemiology. 2014;25(2):300-6.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE