An introduction to confounding, causal inference and DAGs

Anna Lindam, statistician

FoU – enheten Region Jämtland Härjedalen

Arranged by: Forum Norrs lunch method seminars 23rd November 2021

REGION JÄMTLAND HÄRJEDALEN

Overview

- Epidemiological key concepts
- Confounding
- DAGs

Key concepts

Exposure: the variable/factor we think might cause our outcome

Outcome: disease under study for example cancer, heart disease

Covariate: the variable we would like to adjust for in our model

Confounder: a confounding factor, which causes the exposure and outcome

Key concepts

- Mediator: a variable/factor which lies in the path between an exposure and the outcome or between a confounder and the outcome
- Collider: a factor which is caused by the exposure and the outcome (two arrows collide)

Confounding

- Associated with:
  - Control for factors
  - Adjust for factors

- Avoid confounding bias

May reduce bias….

…but might also introduce bias.

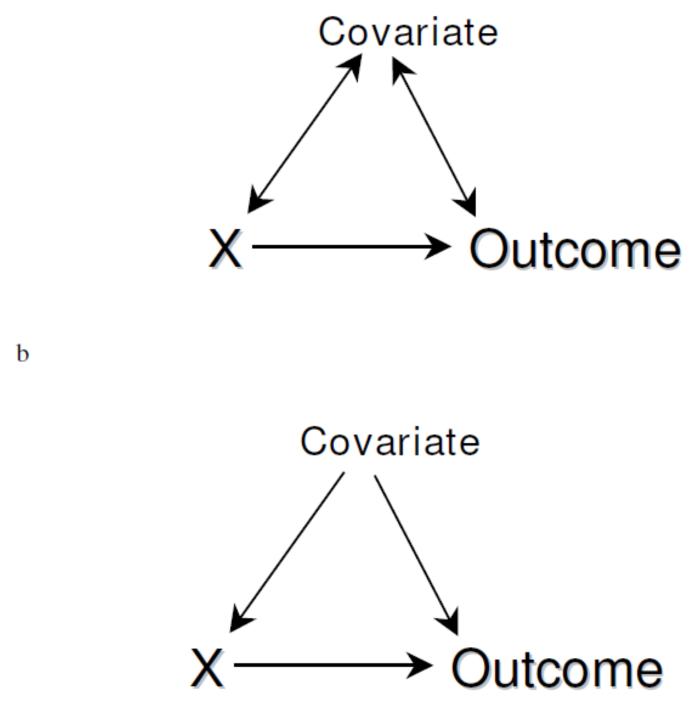
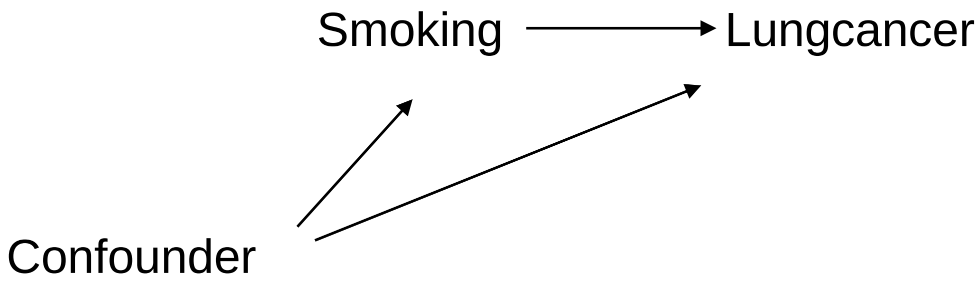


Figure 1
The bi-directional arrows in A show the traditional representation of a confounder as being associated with the exposure (X) and outcome. Because confounders must cause (or be a marker for a cause) of both exposure and outcome (see text for rationale based on basic principles), directed acyclic graphs use only unidirectional arrows to show the direction of causation (B).

Confounders

- Common factors to adjust for
  - age
  - sex
  - education
  - body mass index



DAGs - Directed Acyclic Graphs

A type of causal diagram

Not a statistical method – a way to clarify:

- the relationship between variables
- the assumptions the researchers make in the analyses

Causal inference – to understand causal relationships

How is DAGs used in practice?

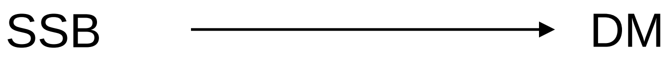
- To plan studies
- To control for confounders in analyses with more than one risk factor
  - For example logistic regression, survival analyses etc.
- To interpret results
- To avoid bias
  - Confounding bias
  - Selection bias
  - Information bias

What is specific for DAGs?

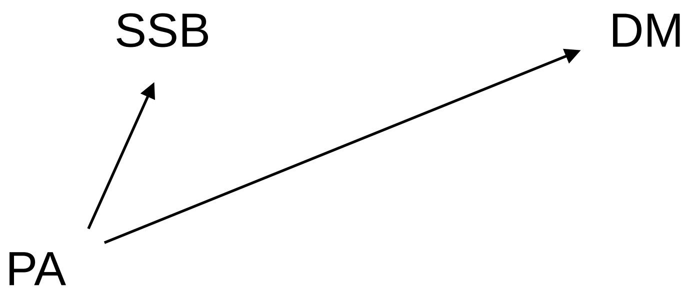
- Illustrates causal relations not predictions or associations
- A confounding factor is a factor which causes exposure and outcome
- A causal relationship is not necessary positive
- Open and closed paths - causal paths
- What we will have left in the graph is causal relationships

Examples

Figur 1.



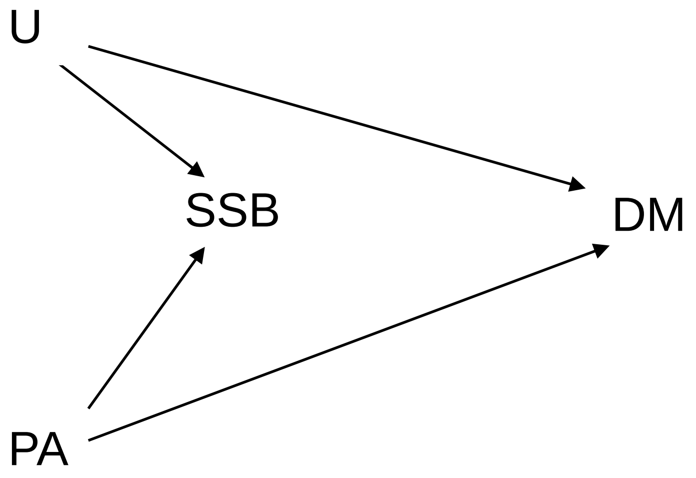
Figur 2.



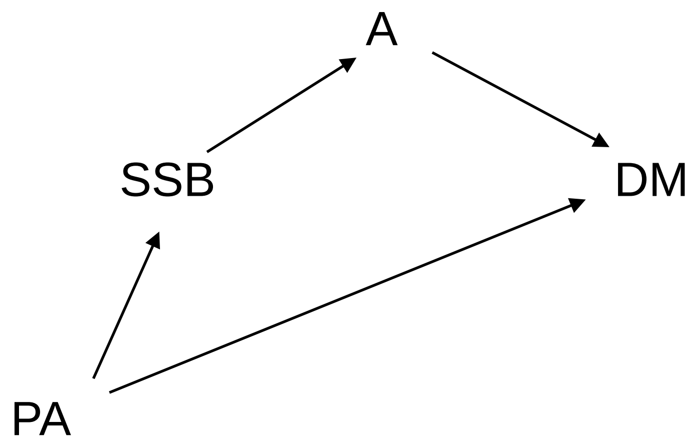
sugar-sweetened beverages (SSB)
diabetes mellitus (DM)
physical activity (PA)

David A. Savitz, *Interpreting Epidemiologic Evidence: Strategies for Study Design and Analysis*

Figur 3.



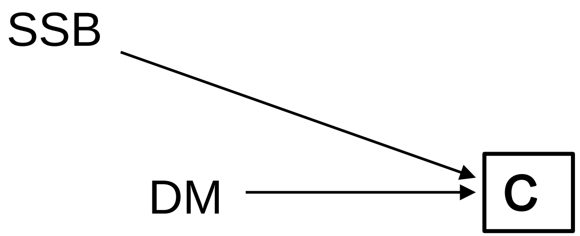
Figur 4.



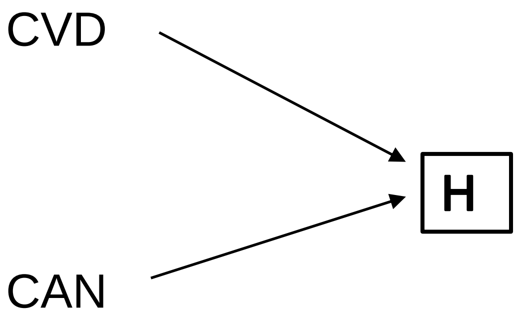
sugar-sweetend beverages (SSB)
diabetes mellitus (DM)
physical activity (PA)

unmeasured factor (U)
adiposity (A)

Figur 5.



Figur 6.



sugar-sweetened beverages (SSB)
diabetes mellitus (DM)
Study design (C)

cardiovascular disease (CVD)
cancer (CAN)
hospitalization (H)

Statistical association

- There is Statistical association in all figures because they show ”open paths”
  - Common cause (figure 2,3,)
  - Share a common effect (figure 5,6) and we examine the conditional association between E and O within levels of that common effect.

- What we would like to do is close paths which does not represent a causal relationship

Confounder vs. collider

Confounder

- Common cause
- If omitted:

  causes bias
- Effect on stratification:

  removes bias

Collider

- Common consequence
- If omitted:

  does not cause bias
- Effect on stratification:

  creates bias

”…expert knowledge is required for the decision to concider a covariate a confounder and to include it in a multivariate model or not.”

Jansky et al. *”The Janus face of statistical adjustment: confounder versus colliders”*

Why important in research?

- You can introduce bias if you adjust for a collider

- Statistical association is not the same as a causal relationship

- In small studies with small samples - not possible to adjust for many variables
  - pick the right ones

- When adjusting for mediators we can introduce overadjustment bias

- We can also decrease the precision of our estimates by adjusting for factors that is not related to the causal relationship under study

Rules

1. A path is blocked by conditioning on any variable on that path that is not a collider
2. If there are no variables being conditioning on, a path is blocked only if there is a collider on that path.
3. A collider that has been conditioned on does not block a path.
4. A collider that has a descendant that has been conditioned on does not block a path. A descendant a variable that is downstream from another on the causal chain.

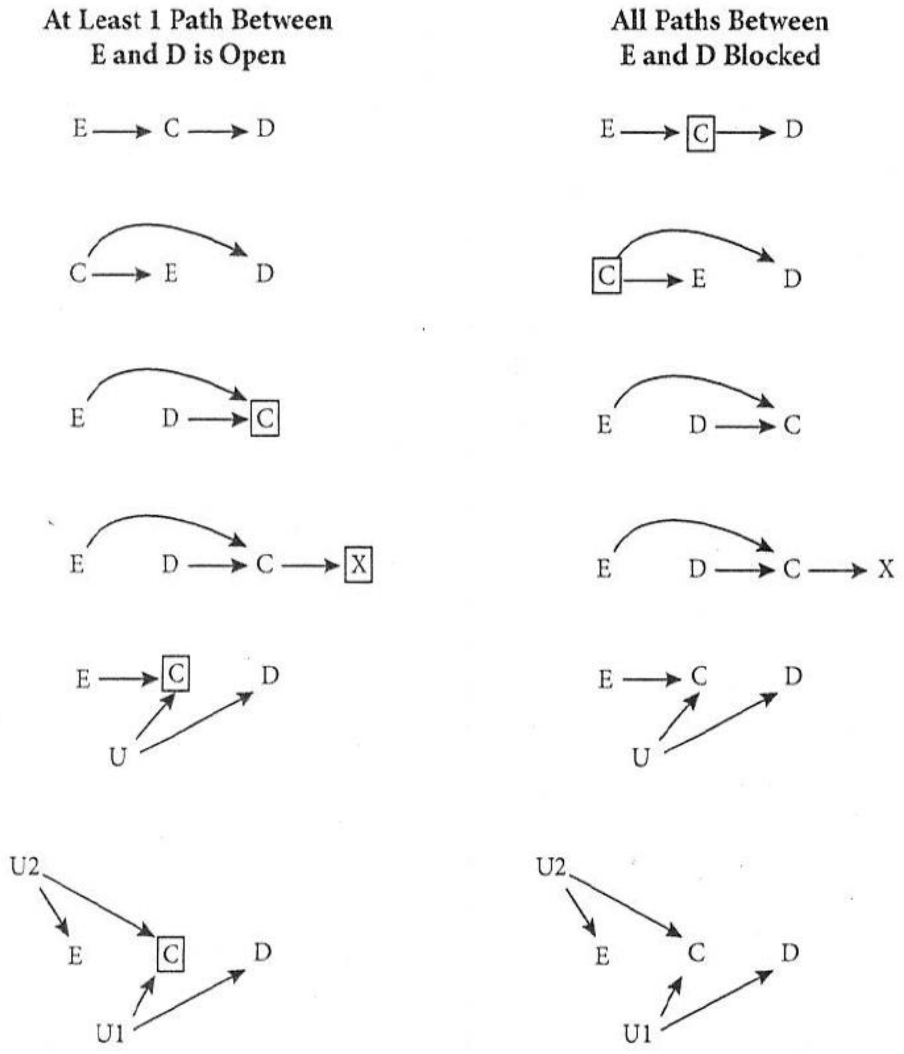


FIG. 3.7 Example causal diagrams. The causal diagrams in the left column each have at least one open path between the exposure (E) and disease outcome (D) of interest. In contrast, the causal diagrams in the right column each have no open paths between E and D.

Exampel – ”The birth weight paradox”

- Low birth weight is associated with infant mortality
- Infants with low birth weight who are exposed to factors which increases low birth weight – decreases the risk for infant mortality
- Infant with low birth weight where the mother smokes have on average lower mortality than infant where the mother does not smoke.

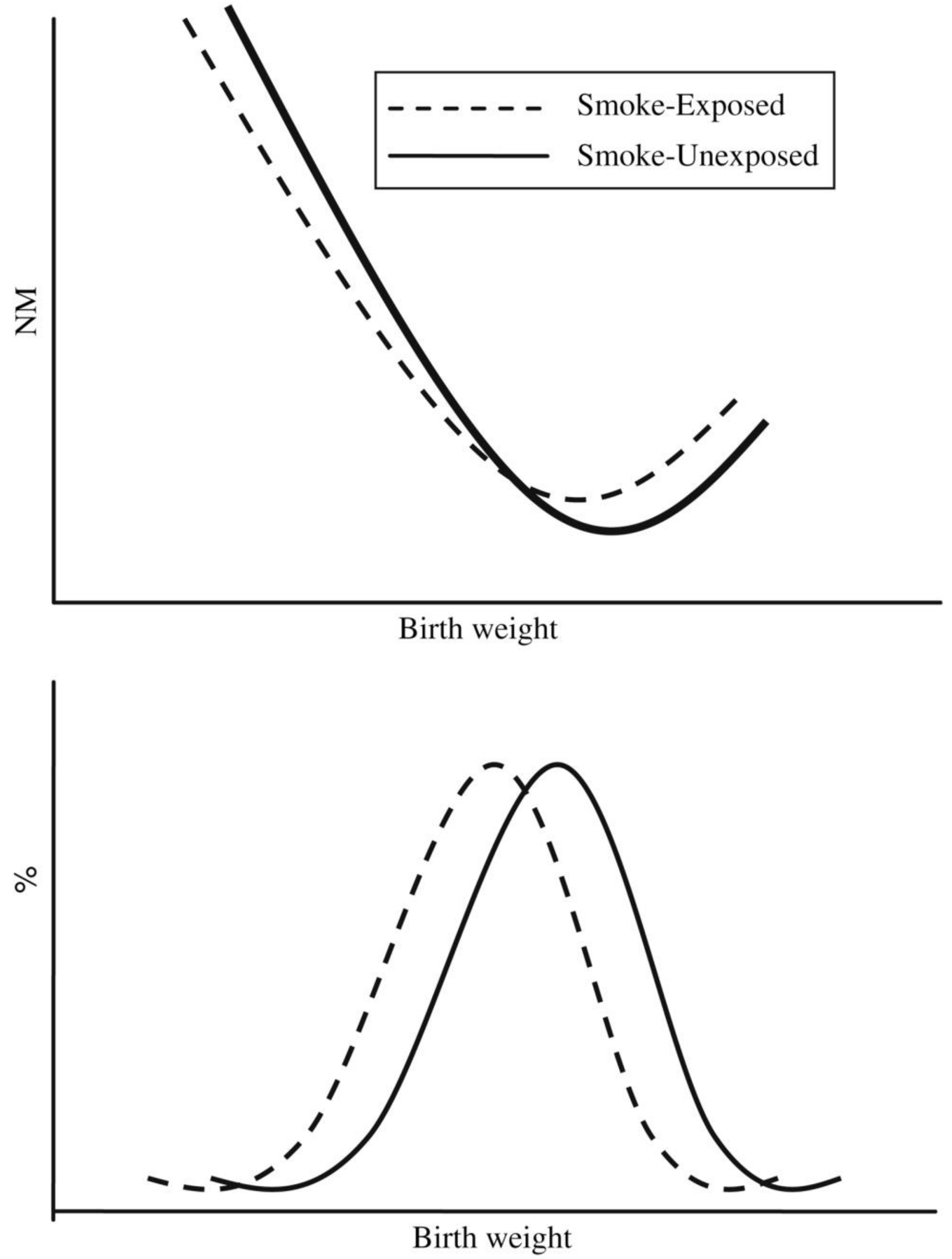
Hernandez-Diaz S, Wilcox AJ, Schisterman EF. et al. From Causal diagrams to birth weight-specific curves of infant mortality

REGION JÄMTLAND HÄRJEDALEN

Birth weight and neonatal mortality (NM) .

Neonatal mortality - death within 28 days from birth.

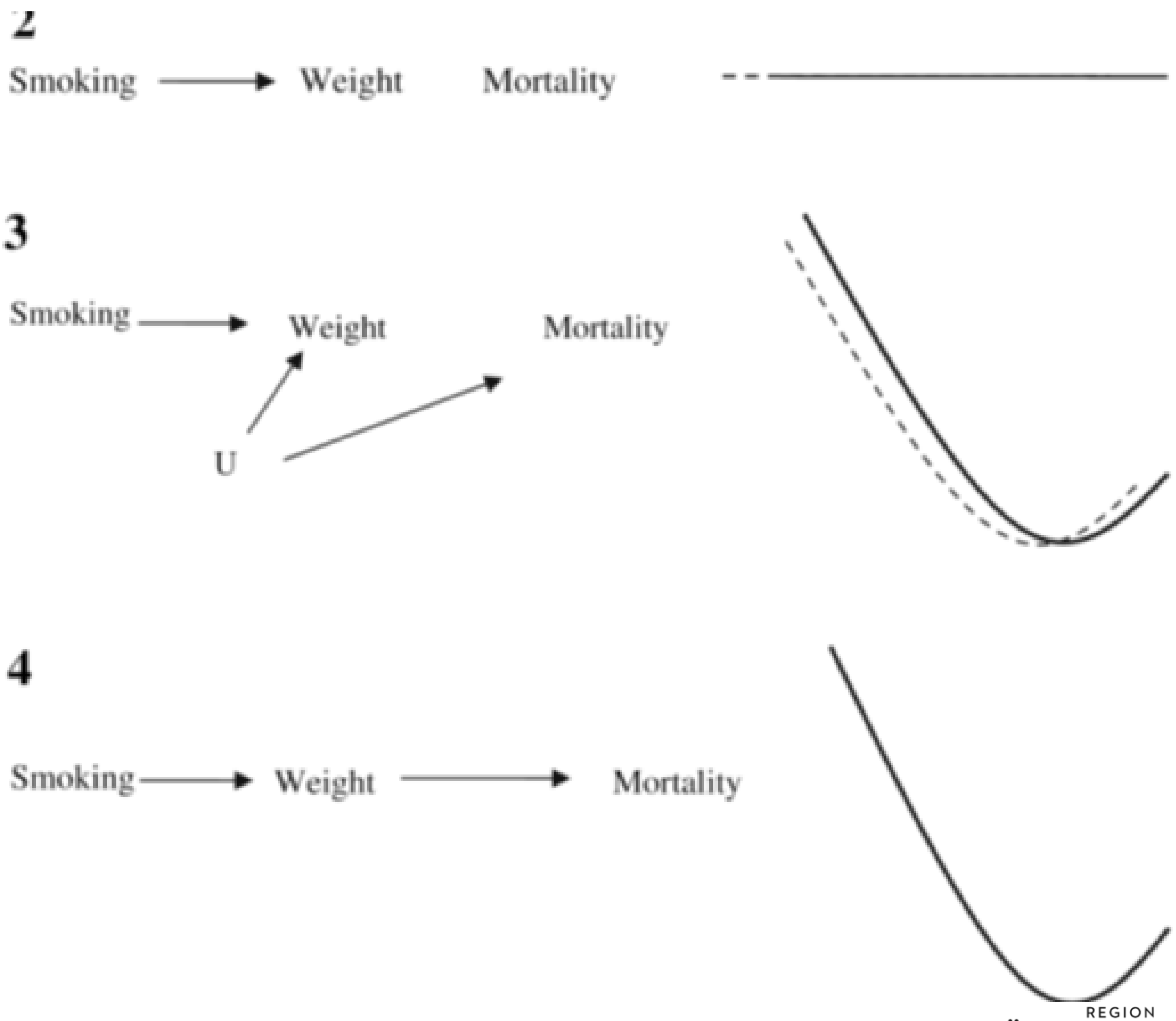
The distribution curve for birth weight for infants of mothers who smoke is lower than for infants of mothers who do not smoke



2. Neither smoking nor birth weight have an effect on mortality

3. In stead there is an unmeasured factor U which causes both low birth weight and mortality

4. Birth weight, but not smoking, has a direct effect on mortality

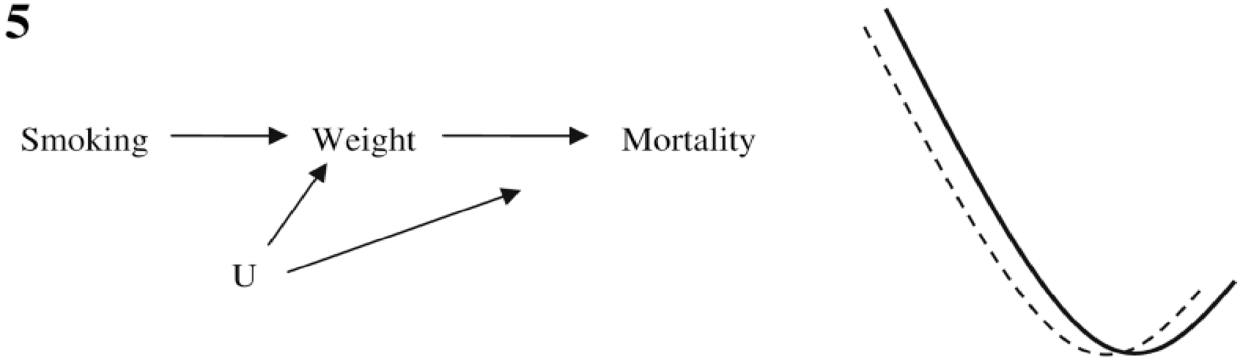


5. An unmeasured factor (U) causes both low birth weight and mortality and there is a direct effect of birth weight on mortality.

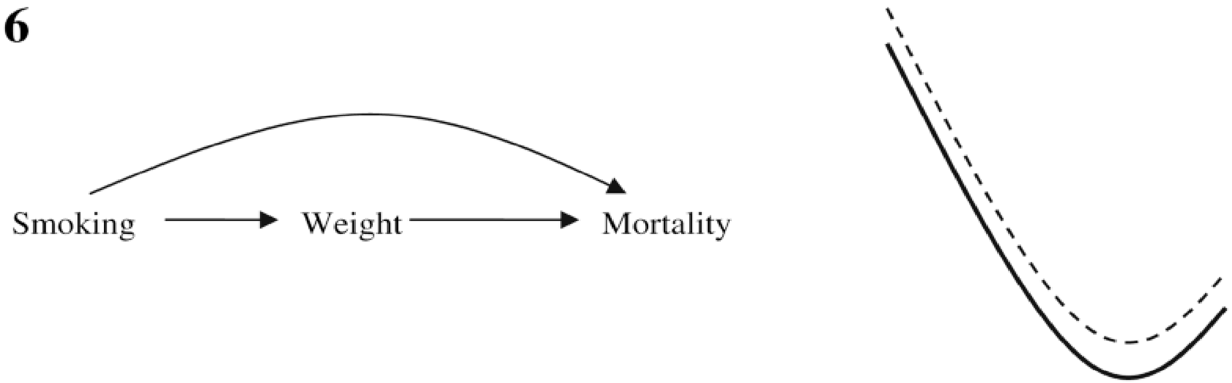
6. Both smoking and birth weight have a direct effect on mortality.

- The effect of smoking is not modified by the effect of birth weight.

5



6

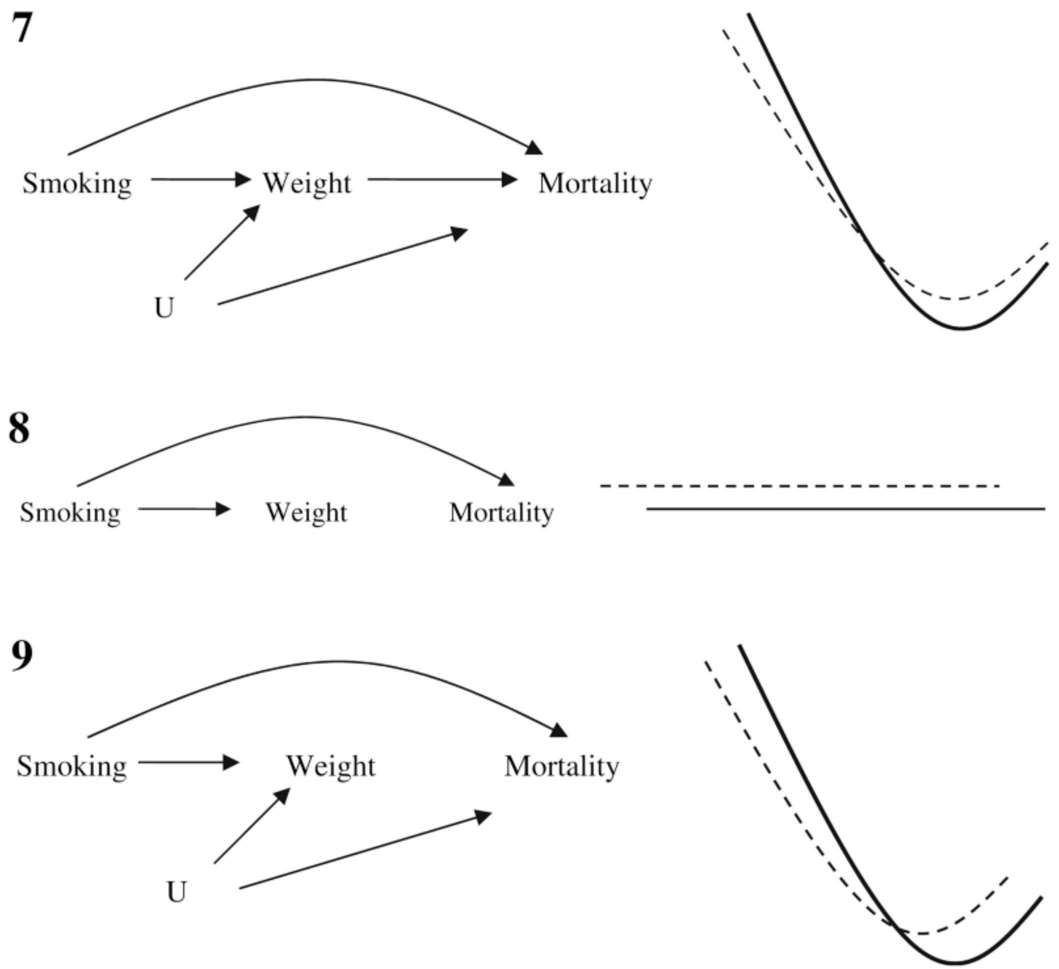


REGION JÄMTLAND HÄRJEDALEN

7. Both smoking and birth weight have a direct effect on mortality and an unmeasured factor (U) is a cause of both low birth weight and mortality.

8. Smoking, but not birth weight, has a direct effect on mortality.

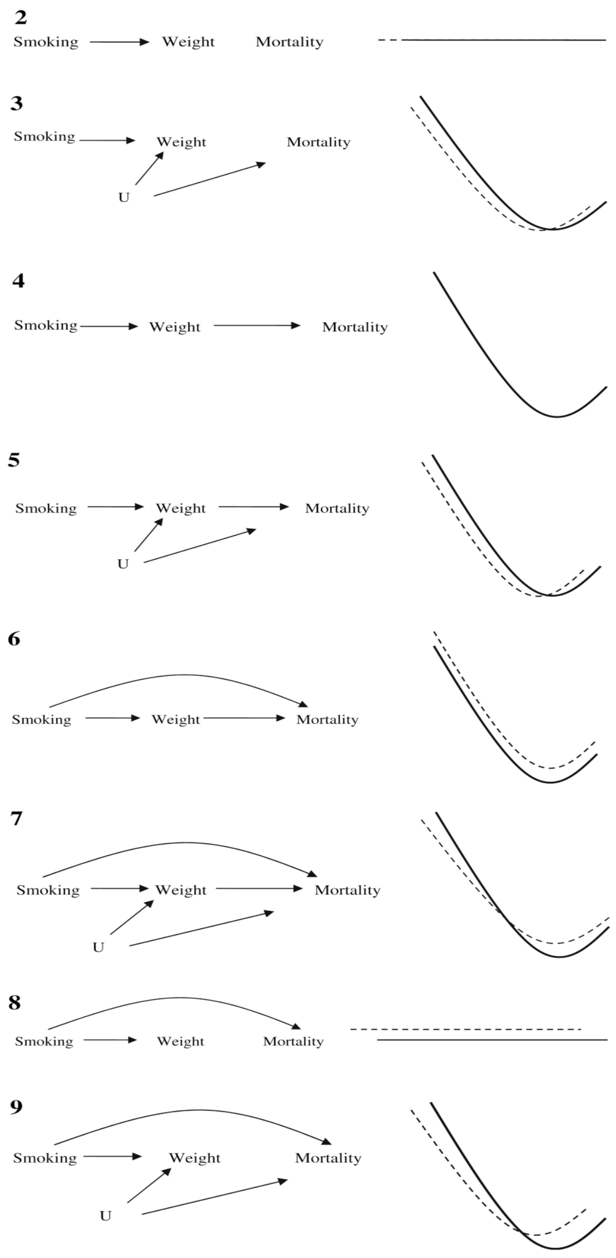
9. An unmeasured factor U is a cause of both low birth weight and mortality which opens a path between birth weight and mortality.



We can disregard figure 2-4, 6 and 8 as they are inconsistent with the observed data.

Figure 5 can be ruled out as it is inconsistent with the commonly accepted hypotheses that smoking is a direct cause of mortality.

We are then left with figure 7 and 9 but as birth weight in itself seem to have a limited effect on mortality, figure 9 is the most likely scenario.



Strengths and weaknesses of DAGs

Strengths

- Clarifies causal relationships
- Facilitates discussions about biological explanations
- Easier to detect bias

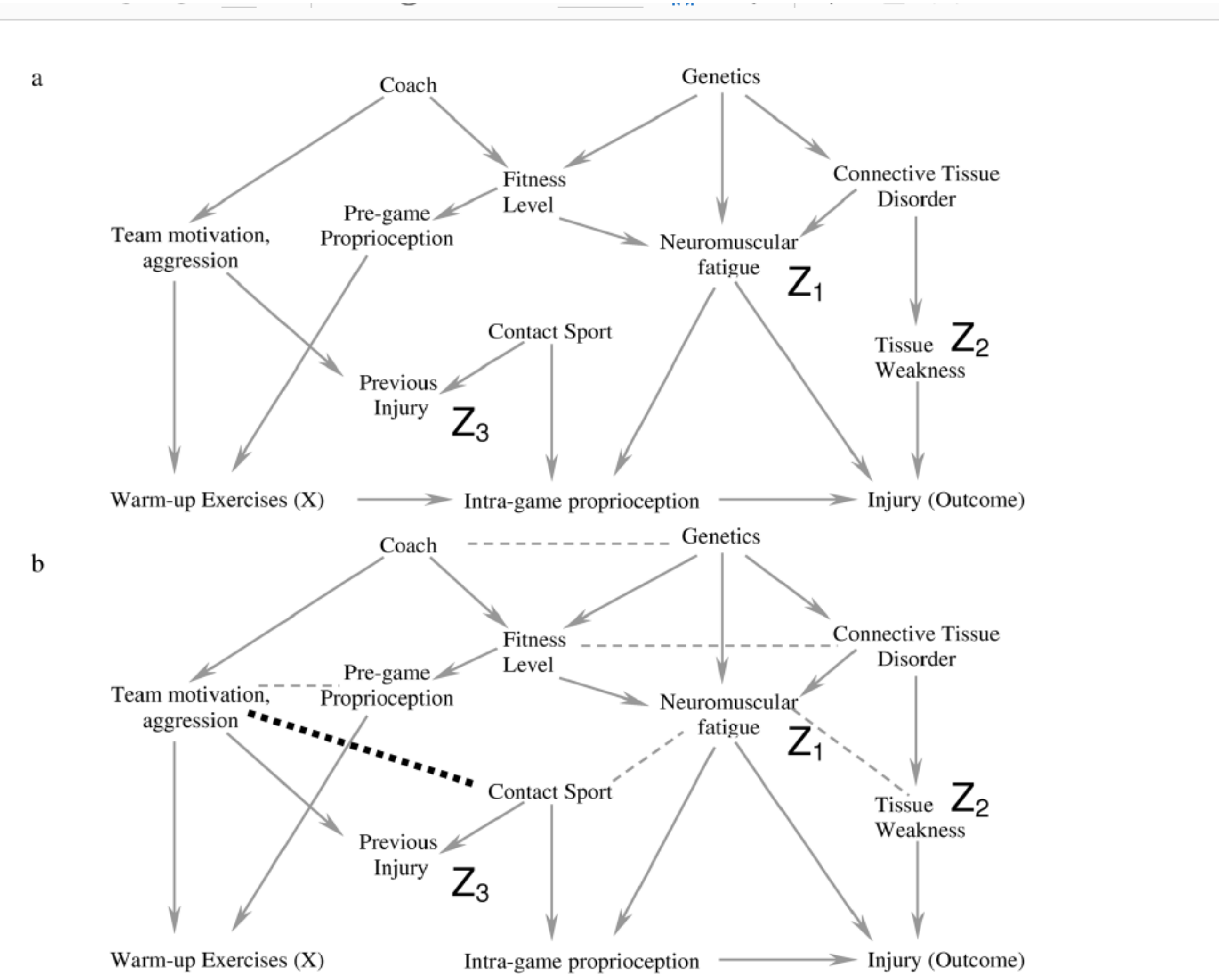
Weaknesses

- Does not show effect size
- Does not handle interactions
- Can not handle to complex situations

5. a-c. This example illustrates the effect of adding the covariate "previous injury" (Z3) to the statistical model

used for the causal diagram in Figure 2a. Note that previous injury is associated with both warming up (through team

motivation/aggression) and the outcome injury (through Contact Sport). After completing steps 1–4, one is left with figure 5b.



Shrier I Platt RW, Reducing bias through directed acyclic graphs BMC Medical Research Methodology 2008, 8:70

Software for DAGs

In the program you can draw your DAG and then the software suggests which variables to adjust for in your model.

Available as a free web version at the link below:

http://dagitty.net/

Referenses

David A. Savitz, *Interpreting Epidemiologic Evidence: Strategies for Study Design and Analysis,* 1st Edition, Oxford University press (2003) p. 21-34

Janszky I, Ahlbom A, Svensson AC, *The Janus face of statistical adjustment: confounder versus colliders;* European journal Epidemiology(2010) 25:361-363

Hernandez-Diaz S, Wilcox AJ, Schisterman EF et al *From Causal diagrams to birth weight-specific curves of infant mortality;* European Journal of Epidemiology. (2008) 23(3): 163–166.

Schisterman EF, Cole SR, Platt RW; *Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies* European Journal of Epidemiology (2009) July ; 20(4): 488–495.

Shrier I Platt RW, *Reducing bias through directed acyclic graphs* BMC Medical Research Methodology 2008, 8:70

REGION JÄMTLAND HÄRJEDALEN

Lunch seminars comming up

15/12 kl. 12-13 Bias in cohort studies (In English)

Christel Häggström, Region Västerbotten/ Institutionen för folkhälsa och klinisk medicin, Umeå universitet

20/1 kl. 12-13 Relativ risk för oddskvotsförvirring? Om olika sätt att uttrycka och jämföra risker.

Per Liv, Region Västerbotten/ Institutionen för folkhälsa och klinisk medicin, Umeå universitet

REGION
JÄMTLAND
HÄRJEDALEN

Thank you for listening!

Anna Lindam

Anna.lindam@regionjh.se

REGION
JÄMTLAND
HÄRJEDALEN