Why *causality* is not such an impossible word

Rhian Daniel, LSHTM

Centre for Statistical Methodology (CSM) Open Meeting

29 September 2010

Outline

Why bother?

What has causal inference research (since Rubin 1978) given us? (1)

1. A formal language (counterfactuals, hypothetical interventions) so that age-old epidemiological concepts can be nailed down mathematically, eg
   - causal effect
   - direct effect
   - indirect effect
   - confounding
   - selection bias
   - effect modification
2. Tools for making explicit the assumptions under which our analysis (eg regression) gives estimates that can be interpreted causally, eg
   - causal diagrams (DAGs)

Why bother?

What has causal inference research (since Rubin 1978) given us? (2)

3. When the assumptions needed for ‘standard’ analyses to be causally-interpretable are too far-fetched, alternative methods have been proposed that give causally-interpretable estimates under a weaker set of assumptions, eg (for problems of intermediate confounding)
   - g-computation formula
   - inverse probability weighting of marginal structural models
   - g-estimation of structural nested models

   [Would this have been possible without 1 & 2?]
4. Sensitivity analyses can be performed to see how robust our (causal) conclusions are to violations of these assumptions
   [Not possible without explicit assumptions]

Example: the birthweight “paradox” (1)

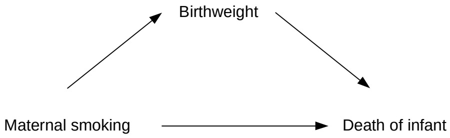
- Many epidemiological studies from the 1960s onwards found that low birthweight (LBW) infants have lower infant mortality in groups in which LBW is most frequent.
- “The increase in the incidence of LBW among infants of smoking mothers was confirmed. However, a number of paradoxical findings were observed which raise doubts as to causation. Thus, no increase in neonatal mortality was noted. Rather, the neonatal mortality rate and the risk of congenital anomalies of LBW infants were considerably lower for smoking than for nonsmoking mothers. These favourable results cannot be explained by differences in gestational age. . . ” (Yerushalmy, AJE 1971)

Example: the birthweight “paradox” (2)

Example: the birthweight “paradox”

A ‘causal inference’ view (1)

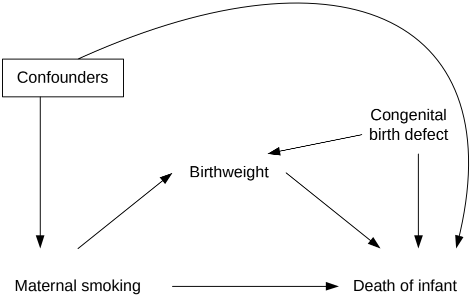
- Hernández-Díaz et al (AJE, 2006) explained this “paradox” using simple causal thinking.



- Birthweight is on the causal pathway from maternal smoking to the death of the child.
- If we wanted the total causal effect of maternal smoking on infant mortality, we shouldn’t adjust for BW.
- By adjusting, we are trying to estimate a direct effect. (Point 1).

Example: the birthweight “paradox”

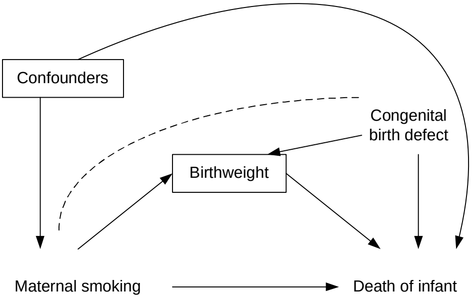
A ‘causal inference’ view (2)



- But there are common causes of LBW and infant mortality, eg congenital birth defects, and confounders of smoking and infant mortality. (Point 2).

Example: the birthweight “paradox”

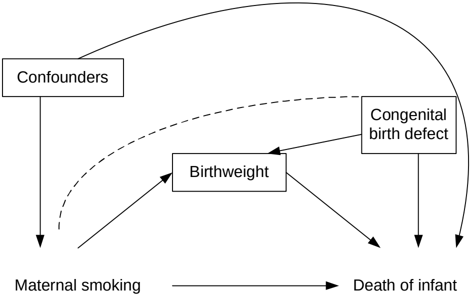
A ‘causal inference’ view (3)



- Stratifying on the common effect of two independent causes induces an association between the causes. (Why?)
- Congenital birth defects plays the role of a confounder in this analysis.
- This explains the “paradoxical” findings.

Example: the birthweight “paradox”

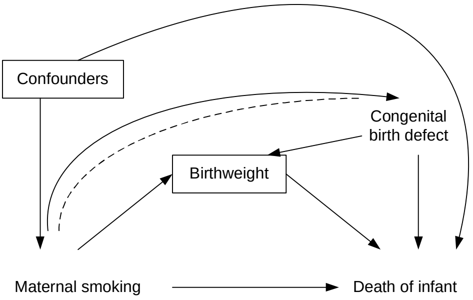
A ‘causal inference’ view (4)



- So we should adjust for it when looking within strata of birthweight. (Still point 2).

Example: the birthweight “paradox”

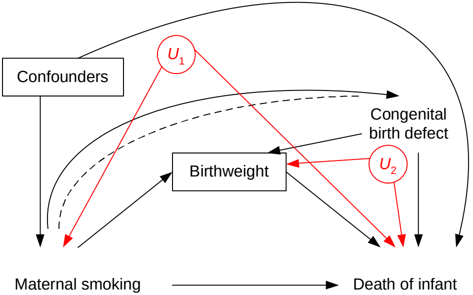
A ‘causal inference’ view (5)



- But what if maternal smoking also causes congenital birth defects?
- Now it is an intermediate confounder.
- Alternative methods (g-computation, ipw, g-estimation) can be used. (Point 3).

Example: the birthweight “paradox”

A ‘causal inference’ view (6)



- And what if there are other (unmeasured) common causes of birthweight and infant mortality?
- Sensitivity analyses. (Point 4).

Why bother?

In conclusion...

- If you know the language of causal inference, you will be able to:
  - know exactly what you mean when talking about causal effect/direct effect/confounding etc
  - be honest about the assumptions under which association=causation
  - try to use analyses based on more plausible assumptions
  - report how sensitive your causal conclusions are to these assumptions
- If you don't know the language of causal inference, you risk:
  - getting into a muddle when talking about causal concepts
  - sticking to analyses which can be causally-interpretable only under highly implausible assumptions
  - that people will interpret your estimates causally even when you warn them that association$\neq$causation

Final thought

- Always saying "...but association is not causation" is like putting "this product may contain nuts" on all food packaging.
- It's true and absolves us of all responsibility.
- But is it useful? Is it ethical?
- Causality is not an impossible word. It's challenging, important, interesting, fun...

If you want to know more. . .

Short course

- Causal Inference in Epidemiology: Recent Methodological Developments
- November reading week.
- `http://www.lshtm.ac.uk/prospectus/short/causal_inference.html`

If you want to know more. . .

Seminars/discussion groups/workshops

- Join our causal inference mailing list (email me: Rhian.Daniel@LSHTM.ac.uk)
- Upcoming seminars:
  - November 1st, Manson Theatre, 1pm: "Intermediate confounding, measurement error and missing data: a way through the epidemiologist's reality?"
  - November 19th, 12:45pm (room tbc): "The hazards of hazard ratios" (Jonathan Bartlett)
  - December 1st, 12:45pm (room tbc): "The regression discontinuity design: redesigned for epidemiology" (Gianluca Baio, UCL & Sara Geneletti, LSE)