Overadjustment bias and unnecessary adjustment
Comment on Schisterman E et al, *Epidemiology*, 2009

Rémy Slama
CLEMS – Club de Lecture en Epidemiologie et Méthodes Statistiques
IAB
22 février 2010

Etiological epidemiological studies
3 main types of bias in epidemiology

<table>
<thead>
<tr>
<th>Type of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement error</td>
</tr>
<tr>
<td>Selection bias</td>
</tr>
<tr>
<td>Confounding</td>
</tr>
</tbody>
</table>

Overview

- The **observational** nature of epidemiological studies entails issues in terms of comparability of exposed and non-exposed groups.
- As a consequence, **confounding** bias can occur.
- Statistical **adjustment** (in a regression model) is one of the main approaches used to correct for confounding bias.
- Although adjustment is meant to limit confounding bias, in some cases it can lead to:
  - A decrease in the precision of the estimated effect of E on D, without affecting bias (**unnecessary adjustment**)
  - A bias in the estimated impact of E on D (**overadjustment**)
Dans quelles situations un facteur d'exposition E est-il associé à la maladie M ?

Il y a 5 situations principales :

1. Cause et conséquence :
   — Existence d'une association causale entre E et M

2. Conséquence commune :
   — E et M partagent une conséquence commune sur laquelle on stratifie l'association entre E et M

3. Cause commune :
   — E et M ont une cause commune

4. Fluctuations d'échantillonnage

5. Erreur de mesure (sur E et/ou M)
Dans quelles situations un facteur d'exposition E est-il associé à la maladie M ?

Il y a 5 situations principales :

1. **Cause et conséquence** :
   - Existence d'une association causale entre E et M

2. **Conséquence commune** :
   - E et M partagent une conséquence commune sur laquelle on stratifie l'association entre E et M

3. **Cause commune** :
   - E et M ont une cause commune

4. **Fluctuations d'échantillonnage**

5. **Erreur de mesure (sur E et/ou M)**

Biais de causalité inverse

Biais de sélection

Biais de confusion

Biais de classement

Overadjustment
Overadjustment due to an intermediate variable

Ex. 1: BMI  Triglyceride levels  Preeclampsia
Ex. 2: Diet  Cholesterol levels  Heart diseases

Adjusting for M will entail a bias (usually towards the null) in the estimated association between E and D.

However, adjusting for M can make sense if one is interested in estimating the direct impact of E on D (i.e., the impact not mediated by M).

Overadjustment due to a “descending” proxy for an intermediate variable

Ex: Smoking (current and past)  Uterine abnormality (unmeasured)  Spontaneous abortion

Adjusting for M will entail bias (usually towards the null) in the estimated association between E and D (Weinberg, AJE, 1993).
Overadjustment: a theoretical model

- Model assumptions:
  \[ U = \alpha_U + \beta_U E + \varepsilon_U, \]
  \[ M = \alpha_M + \beta_M U + \varepsilon_M, \]
  \[ D = \alpha_D + \beta_D U + \varepsilon_D. \]

“Real” (unbiased) effect of E on D:
\[ D = c_1 + \beta_D E + \varepsilon_1 \]

Estimated effect of E on D if one adjusts for M:
\[ D = c_2 + \frac{\beta_D \times \beta_U}{1 + \beta_M} + \frac{\beta_D}{(1 + \beta_M)^2} \]

Amplitude of the resulting bias in the association between D and E:
\[ \frac{\beta_D \times \beta_U}{(1 + \beta_M)} - \beta_D \times \beta_U \]

If M is an ascending (and not descending) proxy of U, then adjusting for M will not cause overadjustment.
Here $C_4$ is a risk factor for $D$ that is not directly associated with $E$. Adjusting for $C_4$ will not bias the estimated association between $E$ and $D$ (bias-neutral adjustment). However, adjusting for $C_4$ will possibly entail a gain or loss in precision.
Variation in precision due to adjustment for a risk factor of D that is not associated with E (C₄)

Simulation based on a linear regression model, n<1000:
• Adjusting for C₁, C₂, ... C₄ does not entail a strong bias.
• Adjusting for C₄ entails a decrease in the variance of the estimate of the parameter quantifying the association between E and D

Other situations corresponding to unnecessary adjustment

Adjusting for C₅ is not only unnecessary but harmful (induces a bias)
Additional remarks

• Adjusting for a consequence $C_3$ of $E$ does not correspond to overadjustment (Schisterman et al.)
  – However, if $C_3$ is also a consequence of $D$, then a bias is expected
    (“selection bias”, see Hernan, Epidemiology, 2004)

• Distinguishing unnecessary adjustment from confounding may in practice not always be easy
  – How can you be sure that $C_4$ does not affect $E$?

Terminology

• Confounding:
  Distorsion in the association between $E$ and $D$ due to extraneous factors.

• Residual confounding:
  Confounding due to confounding factors poorly taken into account (e.g., because of an inappropriate coding of the variable).

• Overadjustment:
  Control for an intermediate variable (or a descending proxy for an intermediate variable) on a causal pathway between $E$ and $D$.

• Unnecessary adjustment:
  Control for a variable that does not affect bias but that does affect precision.

Question: is this adjustment really unnecessary if precision is increased after adjustment (for linear models)?
Critères de confusion
(Rothman, Greenland, Lash, *Modern Epidemiology *)

3 critères nécessaires pour que X soit un facteur de confusion dans la relation entre E et M :

Dans la *population source* de l’étude :

1. X est associé à l’exposition dans la population source
2. X est associé à la maladie chez les non exposés
3. X n’est ni une conséquence de E ni de D

![Diagram](image)

Conclusion

• Not adjusting for a confounder can entail a bias
• Adjusting for a factor that does not satisfy the criteria for being a confounder can also entail a bias.
• When selecting confounders, it is important to exclude a priori factors that are likely to be on the causal pathway between exposure and disease
• It is also important **not to adjust for consequences of D**
• Adjusting for a factor that is only associated to E or D can impact precision
  – This impact can sometimes be beneficial
Like any treatment, statistical adjustment can have serious side-effects...