Sensitivity Analysis for Unmeasured Confounding

Tyler J. VanderWeele
Harvard T.H. Chan School of Public Health
Departments of Epidemiology and Biostatistics

Plan of Presentation

(1) Introduction to Sensitivity Analysis

(2) Cornfield Conditions

(3) Bounding Factors for Sensitivity Analysis

(4) E-values for Sensitivity Analysis

(5) Empirical Examples

(6) Other Scales

- Hazard Ratios, Odds Ratios, Continuous Outcomes

(6) Other Sensitivity Analysis Results

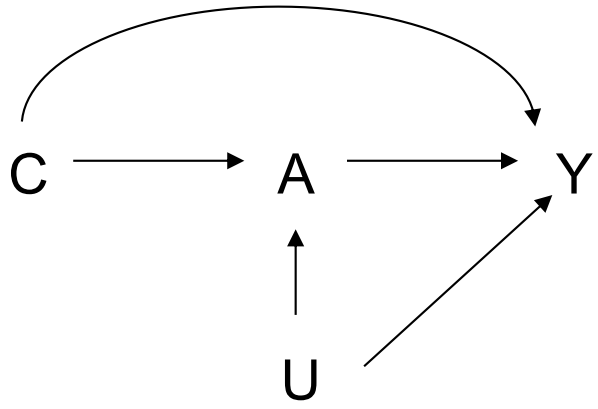
(7) Concluding Remarks

Unmeasured Confounding

Unmeasured/uncontrolled confounding is a common problem in observational studies

Attempts are made to collect data on and control for as many covariates as possible that are related to the exposure and the outcome of interest

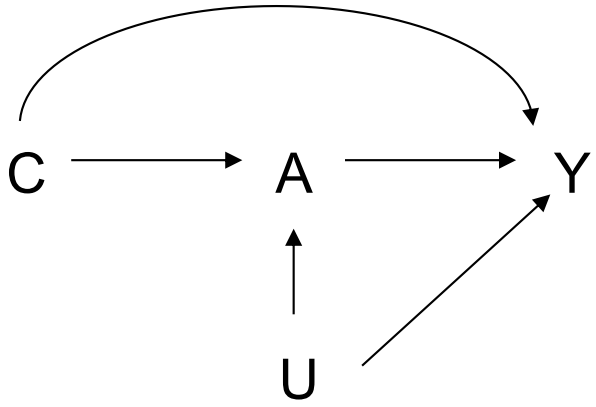
Often, however, one or more important covariates U will remain unmeasured that might confound the relationship between the exposure and the outcome



Such unmeasured confounding variables can bias estimate of the effect of the exposure A on outcome Y

Sensitivity Analysis

”Bias analysis” or “Sensitivity analysis” can help assess the extent to which an unmeasured variable (or variables) U would have to affect both exposure A and outcome Y in order for the observed associations between A and Y to be attributable solely to confounding rather than a causal effect of A on Y



Sensitivity analysis can also be useful in assessing a plausible range of values for the causal effect of A on Y corresponding to a plausible range of assumptions concerning the relationship between the unmeasured confounder U and the exposure A and outcome Y

Example: Breastfeeding is associated with better infant and maternal outcomes… might this be explained by the fact that breastfeeding mothers have higher SES, rather than breastfeeding itself?

Bounds for Total Effects

Sensitivity analysis specifies parameters for the relationships between the unmeasured confounding variables and the observed variables and considers how estimates of effects would change under different values of these sensitivity analysis parameters

Bounds for causal effect essentially consider "worst-case scenarios" i.e. what is the smallest and the largest the causal effect could possibly be given the observed data (Manski, 2003)

Bounds generally give much wider possible ranges for causal effects because much weaker assumptions are used and because more extreme scenarios are considered

Employing additional assumptions for bounds will result in narrower ranges
The stronger the assumptions the tighter the bounds but the less certain one is as to whether the assumptions hold

Sensitivity Analysis

The term “sensitivity analysis” is used outside the context of unmeasured confounding bias to address other types of bias such as selection bias and measurement error

In other fields “sensitivity analysis” is also referred to as “bias analysis” (e.g. Rothman et al., 2008; Lash et al., 2009)

An early application of this approach was the work of Cornfield et al. (1959) who showed that the association between smoking and lung-cancer association was unlikely to be entirely due to unmeasured confounding by an unknown genetic factor

Cornfield Conditions

Let RR denote the estimated risk ratio relating exposure A and outcome Y, possibly conditional on covariates C

Let U denote a binary unmeasured confounding variable

Cornfield et al. (1959) considered how strong the associations between U and A (and between U and Y, Schlesselman, 1978) would have to be to completely explain away the observed RR relating A and Y

Let U_A denote the risk ratio relating A and U
Let U_Y denote the risk ratio relating U and Y

The "Cornfield Conditions" are that U could not be entirely responsible for the observed RR between A and Y unless:

$$U_A > RR \quad \text{and} \quad U_Y > RR$$

i.e. if RR was greater than either U_A or U_Y then U could not be entirely responsible

Cornfield Conditions

Fisher thought the smoking-lung cancer relationship could be explained by a genetic variant U (Fisher, 1958, Nature)

But RR = 10.73 (95% CI: 8.02, 14.36; Hammond and Horn, 1958)

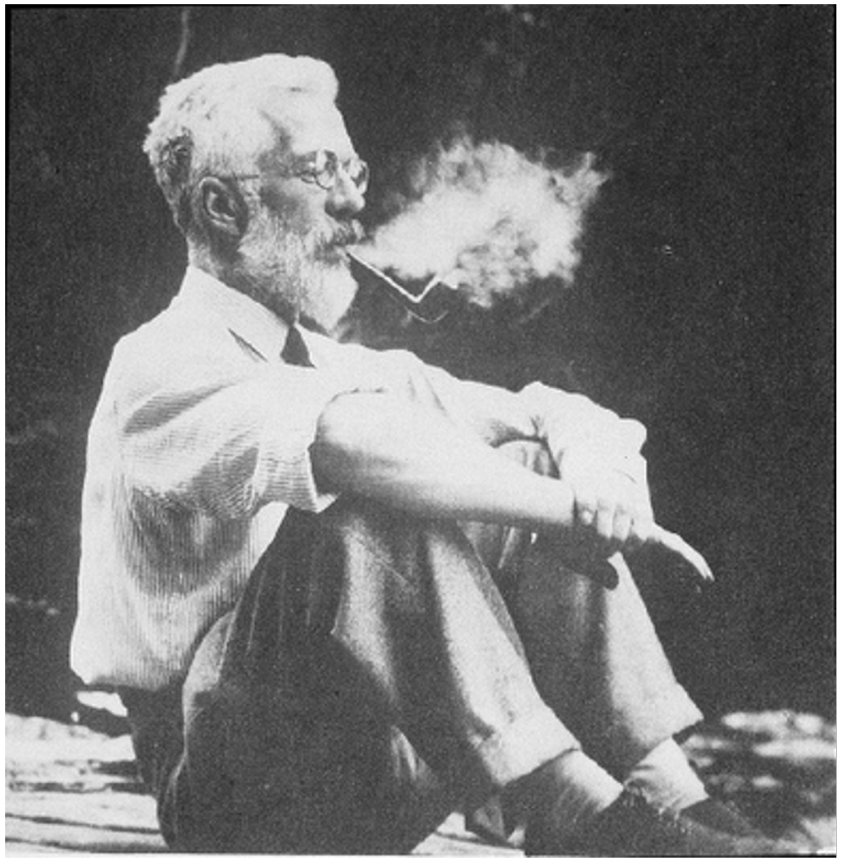
So we would need (Cornfield et al., 1959):

$RR_{UY}>10.73$

and $RR_{AU}>10.73$

This does not seem plausible
Almost certainly there is an effect

Cornfield conditions are useful for assessing if an effect could be completely explained away, but are less useful for assessing plausible effects sizes; sensitivity analysis is more useful for this



Sensitivity Analysis for Total Effects

The estimate adjusted for only measured covariates C is $\frac{P(Y=1|A=1,c)}{P(Y=1|A=0,c)}$

The estimate adjusted for C and U is $\frac{\sum_u P(Y=1|A=1,c,u)P(u|c)}{\sum_u P(Y=1|A=0,c,u)P(u|c)}$

Causal Effects: Let Y_a be the potential outcome or counterfactual outcome for each individual if, possibly contrary to fact, A were set to a

If the effect of A on Y is unconfounded conditional on C (i.e. $Y_a \perp\!\!\!\perp A \mid C$) then the first expression is $P(Y_1=1|c) / P(Y_0=1|c)$

If the effect of A on Y is unconfoudned conditional on (U,C) but not on C alone then the second expression is = $P(Y_1=1|c) / P(Y_0=1|c)$ but the first is not

We let the bias factor $Bias_{RR}$ be the ratio between the estimate and the "truth"

$$Bias_{RR} = \frac{P(Y=1|A=1,c)}{P(Y=1|A=0,c)} \Big/ \frac{\sum_u P(Y=1|A=1,c,u)P(u|c)}{\sum_u P(Y=1|A=0,c,u)P(u|c)}$$

We provide bounds for this bias based on sensitivity analysis parameters

Sensitivity Analysis for Total Effects

Numerous sensitivity analysis techniques exist for risk ratios (relating $Bias_{RR}$ to sensitivity analysis parameters for U-Y and U-A associations)

Many techniques also are available for differences in average outcomes

However many of these techniques make numerous assumptions e.g.

Binary Confounder: Cornfield et al. 1959; Bross 1966, 1967; Schlesselman 1978; Rosenbaum and Rubin 1983

No Interaction b/w Exposure and Confounder: Cornfield et al. 1959; Schlesselman 1978; Imbens, 2003

Only One Confounder: Almost all techniques

Techniques are sometimes criticized because they make assumptions to assess assumptions

Some techniques do not impose these assumptions but require specifying many parameters (Flanders and Khoury, 1990; VanderWeele and Arah, 2011) or parameters that make no reference to an unmeasured confounder (Robins et al, 2000; Diaz and van der Laan, 2013) but are difficult to interpret

Sensitivity Analysis w/o Assumptions

Here we give an approach that in some sense makes “no assumptions” (Ding and VanderWeele, 2016)

Assume we have one or more unmeasured confounders U

We define the following parameters (always >= 1)

$$RR_{UY} = \max\left(\frac{\max_u P(Y=1|A=0,c,u)}{\min_u P(Y=1|A=0,c,u)}, \frac{\max_u P(Y=1|A=1,c,u)}{\min_u P(Y=1|A=1,c,u)}\right)$$

$$RR_{AU} = \max_u \frac{P(u|A=1,c)}{P(u|A=0,c)}$$

These are essentially the maximum effect of U on Y across exposure groups and the maximum risk ratio relating A and any category of U

If U is binary these are just ordinary risk ratios, but apply more generally

Note that the parameters condition on the measured covariates C

They capture the confounding after having already controlled for C

Sensitivity Analysis w/o Assumptions

Let RR_{UY} be the maximum risk ratio relating any two categories of U to Y conditional on covariates C and across exposure A

Let RR_{AU} be the maximum risk ratio relating A to any category of U

The largest factor by which such a U could reduce an observed risk ratio estimate is given by (Ding and VanderWeele, 2016, Epidemiology):

$$B = \frac{RR_{UY} * RR_{AU}}{(RR_{UY} + RR_{AU} - 1)}$$

We can divide the estimate and CI by B to obtain a "corrected" estimate and CI (i.e. the maximum a confounder could shift the estimate)

The result holds without making any assumptions about the structure of U

If the association is protective we multiply rather than divide

Smoking Lung Cancer

Fisher thought the smoking-lung cancer relationship could be explained by a genetic variant U

But RR = 10.73 (95% CI: 8.02, 14.36)

Suppose we had:

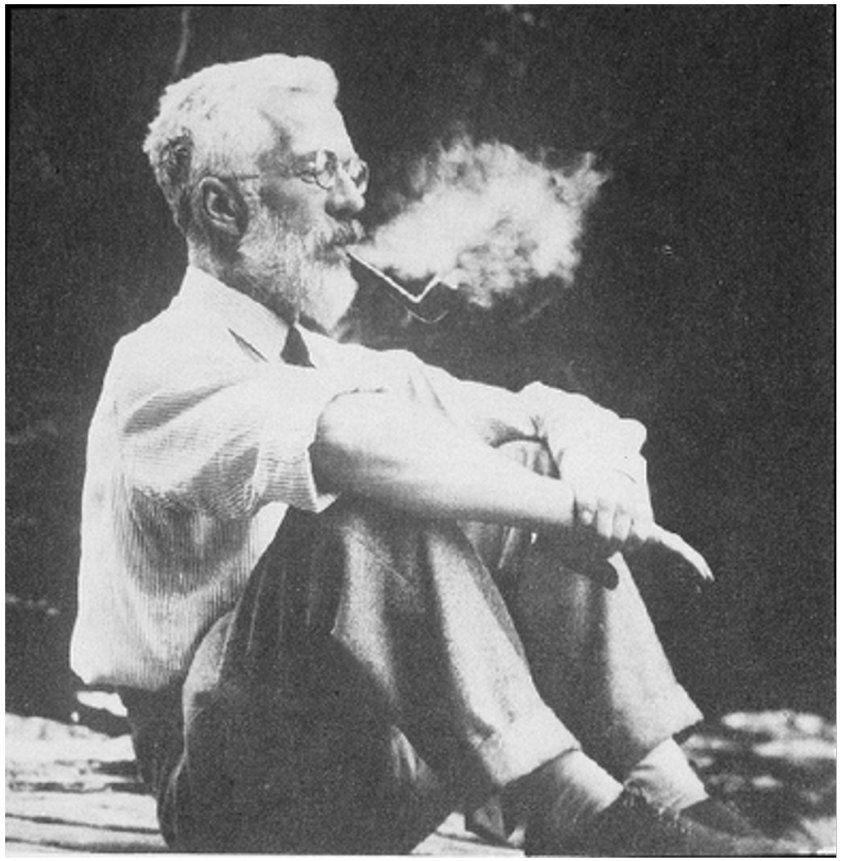
$RR_{UY}=2$

$RR_{AU}=3$

$$B = \frac{RR_{UY}*RR_{AU}}{(RR_{UY}+RR_{AU} - 1)} = \frac{2 * 3}{2+3-1} = 1.5$$

This would shift the estimate to

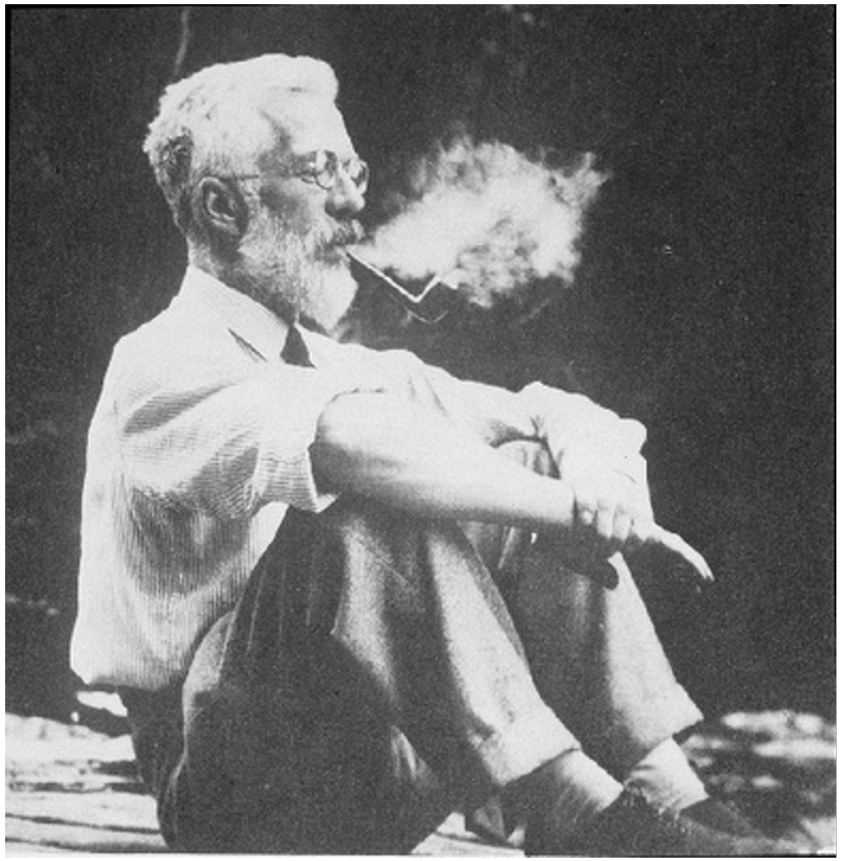
10.73/1.5 = 7.1 (95% CI: 5.3, 9.5)
Still a large effect



Smoking Lung Cancer

Fisher thought the smoking-lung cancer relationship could be explained by a genetic variant U

But RR = 10.73 (95% CI: 8.02, 14.36; Hammond and Horn, 1958)



Suppose we had:

$RR_{UY}=10.73$

$RR_{AU}=10.73$

$$B = \frac{RR_{UY}*RR_{AU}}{(RR_{UY}+RR_{AU} - 1)} = \frac{10.7 * 10.7}{10.7+10.7-1} = 5.6$$

Not even this amount of confounding would explain the association

We would still have an adjusted RR of

RR = 1.91 (95% CI: 1.43, 2.56)

Sensitivity Analysis w/o Assumptions

The Cornfield conditions are generally conservative

The largest factor by which such a U could reduce an observed risk ratio estimate is given by (take inverses for protective exposures):

$$B = \frac{RR_{UY} * RR_{AU}}{(RR_{UY} + RR_{AU} - 1)}$$

If we let RR_{UY} go to infinity we get $B = RR_{AU}$

If we let RR_{UA} go to infinity we get $B = RR_{UY}$

i.e. we recover the Cornfield Conditions (but without assuming a binary confounder); the Cornfield conditions essentially assume the other parameter is infinite; it is conservative in the other parameter

Table of Bias Factors

We can consider many different parameters and calculate the bias factor (i.e. maximum bias that can be generated)

		RR_{UD}								
bounding factor		1.3	1.5	1.8	2	2.5	3	3.5	4	5
	1.3	1.06	1.08	1.11	1.13	1.16	1.18	1.20	1.21	1.23
	1.5	1.08	1.12	1.17	1.20	1.25	1.29	1.31	1.33	1.36
	1.8	1.11	1.17	1.25	1.29	1.36	1.42	1.47	1.50	1.55
	2	1.13	1.20	1.29	1.33	1.43	1.50	1.56	1.60	1.67
RR_{EU}	2.5	1.16	1.25	1.36	1.43	1.56	1.67	1.75	1.82	1.92
	3	1.18	1.29	1.42	1.50	1.67	1.80	1.91	2.00	2.14
	3.5	1.20	1.31	1.47	1.56	1.75	1.91	2.04	2.15	2.33
	4	1.21	1.33	1.50	1.60	1.82	2.00	2.15	2.29	2.50
	5	1.23	1.36	1.55	1.67	1.92	2.14	2.33	2.50	2.78

Approaches to Sensitivity Analysis

Sensitivity analysis does not give one right answer but a range
It is sometimes objected that there is too much subjectivity in sensitivity analysis

Possible Approaches:
(1) Report how large the effects of the confounder would have to be to completely explain away (i) the effect and (ii) the CI
(2) Create a table with many values of all sensitivity analysis parameters; include those one thinks are too extreme; let the reader decide
(3) Find the most important measured confounder variable; examine if an unmeasured confounder of similar strength would change conclusions
(4) Use external data or expert opinion to inform sensitivity analysis parameters

We will focus on the first approach of “explaining away”

E-Values

The E-value (VanderWeele and Ding, 2017):

> "the minimum strength of association on the risk ratio scale that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to fully explain away a specific exposure-outcome association."

The E-value is thus a measure pertaining to… *evidence* for causality

With an observed risk ratio of RR, we have that if RR_{UY} and RR_{AU} are greater than (VanderWeele, and Ding, 2017):

$$\text{E-value} = RR + \text{sqrt}[\, RR*(RR-1)\,]$$

then this could suffice, but weaker joint confounder associations could not

E-Values

We can also ask how much confounding would explain away an estimate and its confidence interval; we apply the formula:

E-value = RR + sqrt[RR*(RR-1)]

We can apply this in a routine manner to both the estimate and the confidence interval limit closest to the null

If RR<1, take inverses first and apply the formula

We might call this the E-value (*evidence* for causality)

RR = 10.73 (95% CI: 8.02, 14.36)

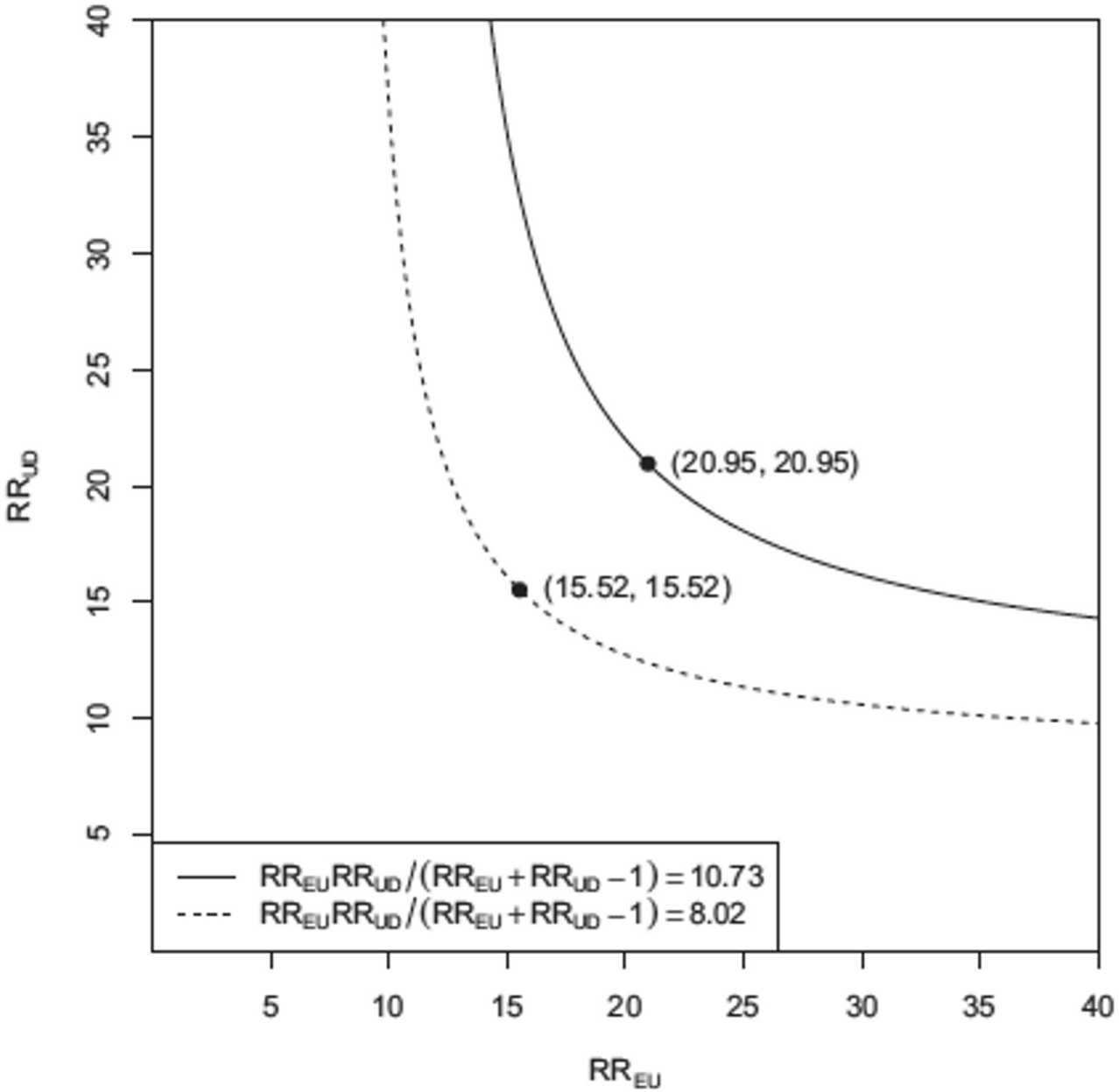
E-Value for Estimate: 20.9; E-value for CI: 15.5

"With an observed risk ratio of RR=10.7, an unmeasured confounder that was associated with both the outcome and the exposure by a risk ratio of 20.9-fold each, above and beyond the measured confounders, could explain away the estimate, but weaker joint confounder associations could not"

Sensitivity Analysis w/o Assumptions

Smoking and Lung Cancer (Hammond and Horn, 1958):

RR = 10.73
(95% CI: 8.02, 14.36)



Breast-feeding Example

Consider potential breast-feeding effects (concerns about confounding by SES)

Lower Respiratory Tract Infection (AHRQ, 2007): RR = 0.28 (95% CI: 0.14, 0.54)

To calculate E-value for estimate for LRTI: 1/0.28 = 3.57

E-value = RR + sqrt[RR*(RR-1)] = 3.57+sqrt{3.57×(3.57-1)} = 6.6

To calculate E-value for confidence interval 1/0.54 = 1.85

E-value = RR + sqrt[RR*(RR-1)] = 1.85+sqrt{1.85×(1.85-1)}=3.1

We can interpret the E-Value as e.g.

> "With an observed risk ratio of RR=0.28, an unmeasured confounder that was associated with both the outcome and the exposure by a risk ratio of 6.6-fold each, above and beyond the measured confounders, could explain away the estimate, but weaker joint confounder associations could not; to move the confidence interval to include the null, an unmeasured confounder that was associated with the outcome and the exposure by a risk ratio of 3.1-fold each could do so, but weaker joint confounder associations could not."

We could include statements like this in all empirical papers

Breast-feeding Example

Examples (all highly "statistically significant") (AHRQ, 2007; Moorman, 2008)

Lower Respiratory Tract Infection: RR = 0.28 (95% CI: 0.14, 0.54)

Maternal Ovarian Cancer: RR = 0.5 (95% CI: 0.3, 0.8)

Childhood Leukemia: RR = 0.81 (95% CI: 0.71, 0.91)

E-values:	Estimate	Upper CI
Lower Respiratory Tract Infection:	6.6	3.1
Maternal Ovarian Cancer:	3.4	1.8
Childhood Leukemia:	1.7	1.4

Strong evidence for LRTI; moderate for ovarian cancer; modest for leukemia

Just because an unmeasured confounder can explain an effect away does not mean that it does (effect might still be present, but evidence isn't definitive)

E-Values and P-values

In observational studies intended to assess causality, the E-value could be reported to supplement the p-value (or some other sensitivity analysis used)

P-value gives evidence for *association*

E-value gives evidence that the association is *causal*

The two can diverge:

E-values:	Estimate	Upper CI
Maternal Ovarian Cancer:	3.4	1.8
Childhood Leukemia:	1.7	1.4

P-value is more extreme for leukemia ($p<.001$) than ovarian cancer ($P=.006$)

E-value is more extreme for ovarian cancer than leukemia

- P-value (but not E-value) can be made more extreme by larger sample size
- The E-value for the CI will generally get larger as sample size increases, but converge to the E-value for the observed association
- In a different context Rosenbaum (2009) refers to this limit as “design sensitivity”; but what is most relevant is arguably the data available

E-Values and Interpretation

- What is a “large” E-value is relative to the outcome and exposure (we might report associations of these with measured covariates to see what is “large”)
- What is a “large” E-value depends on the measured covariates; the parameters are conditional on the measured covariates; if control is made for many measured covariates, it is harder to explain away a given estimate
- We can apply the approach with many unmeasured confounders but then large values of the parameters are more plausible
- E-value is conservative; it considers extreme scenarios; if the unmeasured confounder is very rare it will not generate the amount of bias suggested by the E-value or bounding factor
- If, after control for many covariates, we still have evidence for an effect after calculating the E-value, this can be very persuasive
- But we cannot really use the E-value in the same way to provide evidence for “no effect” because the bounding factor and E-value considers the maximum bias a confounder could induce; absence of evidence is not evidence of absence
- One can also calculate “non-null E-values” to see how much confounding is needed to move an estimate to any other value

Non-Null E-Values

If we want to see how much confounding would shift the estimate to a value other than 1, e.g. to RR^{true}_{AY}, we can do this as well

We just take the ratio of the ratio $RR' = RR^{true}_{ED} / RR$ and then again apply the E-value formula:

$$\text{E-value} = RR' + \text{sqrt}[\, RR'*(RR' -1)\,]$$

As before, if the ratio RR^{true}_{ED} / RR less is than 1, then take inverses before applying the formula

Other Scales

The same approach also works with:

- ✧ Odds Ratios with a rare outcome
- ✧ Hazard Ratios with a rare outcome at end of follow-up
- ✧ Ratios of count outcomes
- ✧ Continuous positive outcomes (using a ratio of expectations)

We just replace the RR's above with HR's, OR's, etc.

- For odds ratios with a common outcome, the OR is often much larger than the RR and these will not be comparable
- One can get an approximate E-value by using RR ≈ sqrt(OR) and then applying the E-value formulas (VanderWeele, 2017)
- Sensitivity analyses that present the sensitivity analysis parameters on an OR scale when using a common outcome are likely to exaggerate the robustness
- Other formulas for HR's with a common outcome

Difference Scales

For a difference in continuous outcomes, an approximate E-value can be obtained as follows:

Consider a standardized effect sizes "d" (the outcome variable should be divided by its standard deviation) and a standard error for this effect size s_d

Use approximation $RR \approx \exp(0.91 \times d)$

Use approximate CI: $(\exp\{0.91 \times d - 1.78 \times s_d\}, \exp\{0.91 \times d + 1.78 \times s_d\})$.

Then apply the E-value formula: RR + sqrt[RR*(RR-1)]

This approach relies on approximations from meta-analysis (Borenstein et al., 2009; Hasselblad and Hedges, 1995) to convert standardized effect sizes into OR's, and then approximations between OR and RR

Other sensitivity analysis approaches can be used for continuous outcomes but also make additional assumptions

Advantages of E-Values

Advantages of E-values:

(1) Easy to compute

- By hand, or with an online calculator, or R package or Stata: https://www.evalue-calculator.com/

(2) Easy to report and interpret

(3) Standardized metric across different scales

(4) Does not make functional form assumptions

(5) Removal of subjectivity in sensitivity analysis

(6) Moving away from OR parameters avoids exaggeration

When to avoid?

When the unmeasured confounder is known and rare since the E-value approach will be too conservative

It is best then to use other techniques (e.g. Schlesselman, 1978)

It is also always preferable to carry out a fuller sensitivity analysis but sometimes journals, editors, or even collaborators may make that difficult

Prevalence Specification

If we are willing to specify the prevalence (and make further assumptions) we can obtain the exact bias which will generally be smaller

Result (Schlesselman,1978) : If there is no unmeasured confounding given (C,U) i.e. $Y_a \perp\!\!\!\perp A \mid (C,U)$ and U is a binary unmeasured confounder with the same risk ratio, γ, on Y for both exposed and unexposed so that $$\gamma = \frac{E(Y|a,c,U=1)}{E(Y|a,c,U=0)}$$

Then we have that: $$B_{mult}(c) = \frac{1+(\gamma-1)P(U=1|a_1,c)}{1+(\gamma-1)P(U=1|a_0,c)}$$

We can use the bias formula by specifying:
(i) γ = the effect of U and
(ii) the prevalence of U amongst the exposed and unexposed

Prevalence Specification

$$B_{mult}(c) = \frac{1+(\gamma-1)P(U=1|a_1,c)}{1+(\gamma-1)P(U=1|a_0,c)}$$

Once we have calculated the bias term $B_{mult}(c)$, we can simply estimate our risk ratio controlling only for C (e.g. if the outcome is rare we just fit a logistic regression) and we divide our estimate by $B_{mult}(c)$ to get the corrected estimate for the causal risk ratio i.e. what we would have obtained if we had adjusted for U also

This is now exact, not conservative, but we must make assumptions, and also specify prevalences

We can obtain corrected confidence intervals by dividing both limits of the confidence interval by $B_{mult}(c)$

The result above is given by Schlesselman (1978) building on the work of Bross (1966)

Prevalence Specification

The estimates for breastfeeding for 6-12 months on ovarian cancer was OR of 0.5 (95% CI: 0.3, 0.8)

Suppose we thought that low SES increased the risk of ovarian cancer by 2 fold and that 10% of the 6-12 month breastfeeding group was low SES but 20% of the reference group (no breast-feeding) was low SES

The bias factor is then: $[1 + (2-1)*(0.1)] / [1 + (2-1)*(0.2)] = 0.91$
The corrected estimate would be 0.55 (95% CI: 0.33, 0.88)

With the prevalences specified, the risk ratio of 2-fold each do not explain away the CI (even though the E-value for the CI here was 1.8)
However, in specifying the prevalences there can be temptation to specify low prevalence to make the estimates “look good”
The formula also makes homogeneity assumptions

The E-Value gets around both of these issues
But if we know the prevalence is low, specifying it may be preferable

Difference Scale and Continuous Outcomes

The bias for the causal effect conditional on C on the difference scale:

$B_{add}(c) = \{E(Y|a_1,c) - E(Y|a_0,c)\} - \{E(Y_{a_1}|c) - E(Y_{a_0}|c)\}$

Result. Suppose that we have an unmeasured confounder U with the effect of U on Y on the difference scale the same for the exposed and unexposed with γ = $E(Y|a,c,U=1) - E(Y|a,c,U=0)$ then the conditional bias is:

$$B_{add}(c) = \gamma\delta$$

where δ = E(U|A=1,c)-E(U|A=0,c) is the difference in expectations of U between the exposed and unexposed conditional on C=c (prevalence difference if U is binary)

Note these parameters are conditional on / control for measured C

Difference Scale and Continuous Outcomes

Once we have calculated the bias term $B_{add}(c)$ we can simply estimate our causal effect conditional on C and then subtract the bias term to get the corrected estimate

To get corrected confidence intervals on can simply subtract γδ from both limits of the estimated confidence intervals for either conditional effects or for average causal effects if $B_{add}(c)$ is constant over c

The result was given in a regression context by Cochran (1938) and re-derived numerous times subsequently, though with increasingly weaker assumptions (Lin et al., 1998; VanderWeele and Arah, 2011)

It can be applied regardless of the estimation method and requires only that the effect of U on Y is the same in both exposure strata (no U*A interaction on the difference scale)
But this approach still does make this homogeneity assumption

Covariate Conditioning

- All of the results (and all sensitivity analysis parameters) are conditional on measured covariates
- Unmeasured confounding, to alter the estimate, must affect the exposure and outcome through pathways independent of the measured covariates

- Often, with extensive confounding control, an additional covariate (which may make a big difference when adjusting compared to the crude estimate) doesn't matter much under control for other covariate
- E.g. "Income" may not matter as much if we have also controlled for education, wealth, home ownership, and occupation
- It can be instructive examining this with one's data

- Sensitivity methods (e.g. Altonji et al., 2005) that assume unobserved confounding is comparable in magnitude to observed confounding are problematic b/c it looks like confounding gets worse with additional control
- Somewhat more sensible is to compare what proportion of the effect of measured confounding the unobserved confounding would have to have to explain away the effect (Altonji et al., 2005)
- Interpretation here as well is also with respect to the measured covariates

Conclusions

1) Sensitivity analysis is a powerful tool to assess robustness; and techniques are now available which effectively make “no assumptions”
2) We can use these to calculate the minimum amount of confounding that would suffice to explain away the effect (i.e. the E-value)
3) If the E-value suggests that the evidence is robust, this can be a very strong argument
4) It is more difficult to provide evidence that an effect is absent
5) Too many of our interpretative practices for observational studies are inherited from RCTs
6) The E-value approach corrects this; it is easy to use and is applicable with most effect measures
7) The E-value or other sensitivity analysis techniques could be used very broadly; if this were done our inferences about causation and science would likely be considerably improved; whether this is done depends in part on you…