CONFOUNDER RESISTANCE AND CONFOUNDER INTERVALS FOR A BINARY CONFOUNDER

Milo Schield, Augsburg College and Thomas V. V. Burnham, Cognitive Consulting
Dept. of Business Administration. 2211 Riverside Drive. Minneapolis, MN 55454

Abstract: The defining conditions for a binary confounder to nullify an association have been identified for a non-interactive model involving a binary predictor and a binary outcome. When the association involves a relative risk or prevalence, three values are required to specify the nature of the confounder. One goal of this paper is to identify a meaningful single-value that can specify the numerical properties of a binary confounder that would nullify a given association. Associations that can withstand a certain size confounder without being nullified are considered confounder resistant. A second goal is to identify conditions under which the influence of a confounder can be shown as confounder-intervals for an observed ratio and a given size confounder. Formulas for the upper and lower limits of confounder intervals are determined for relative prevalences. In order to highlight the influence of potential confounders, data analysts using relative risks or prevalences from observational data should be accompany these with some measure of their susceptibility to confounding using either the size confounder that would nullify the association or the interval for a given size confounder.

Keywords: Epidemiology

1. PROBLEM OF CONFOUNDING

Confounding is everywhere in observational studies. Things are tangled up and mingled together; everything seems to be connected to everything. An association between two variables is *confounded* by a third if the third is entangled with both these variables.

While random assignment can statistically break such entanglements, most studies can not (or do not) involve random assignment. Without random assignment, there is no known statistical test for confounding. (Pearl, 1998) Without knowing the distribution of confounders, there seems to be no way to say, "*there is a 20% chance that this observed association is due to confounding*" or "*If this association were entirely spurious, there is less than a 5% chance of seeing an association this big or bigger due to confounding*."

Finally, there seems to be no generally accepted way to talk about the nature or size of a confounder. We have no way to eliminate a variable in a regression by saying it is beneath some minimum threshold for susceptibility to confounding. There is nothing comparable to the "5% level of significance." As a result there is no way to determine what size relative risk constitutes strong evidence for saying the association is not spurious.

Operationally, epidemiologists tend to disregard relative risks of less than three as being generally inadequate to withstand the influence of confounding. Taubes (1995) noted the following: [emphasis added]

> Sir Richard Doll of Oxford University, who once co-authored a study erroneously suggesting that women who took the anti-hypertension medication reserpine had up to a fourfold increase in their risk of breast cancer, suggests that no single epidemiologic study is persuasive by itself unless the lower limit of its 95% confidence level falls above a **threefold** increased risk. Other researchers, such as Harvard's Trichopoulos, opt for a **fourfold** risk increase as the lower limit. Trichopoulos's ill-fated paper on coffee consumption and pancreatic cancer had reported a 2.5-fold increased risk. "*As a general rule of thumb*," says Angell of the New England Journal, "*we are looking for a relative risk of* ***three or more*** *[before accepting a paper for publication], particularly if it is biologically implausible or if it's a brand-new finding*." Robert Temple, director of drug evaluation at the Food and Drug Administration, puts it bluntly: "*My basic rule is if the relative risk isn't at least* ***three or four****, forget it*."

While a relative risk of three may be a rule of thumb in some areas, lower ratios are being used. In concluding that second-hand smoke caused health problems, the EPA relied on a relative risk of 1.2.[1] A relative prevalence of 1.25 is used to monitor adverse impact in hiring practices involving members of protected classes as identified by Title 7 of the 1964-1965 Civil Rights Act.[2]

But as John Bailar, an epidemiologist at McGill University and former statistical consultant for the NEJM, points out, "*there is no reliable way of identifying the dividing line*." Taubes (1995). Thus, any rule of thumb such as RR > 3 requires justification.[3]

[1] www.forces.org/evidence/ets-whop/index.htm

[2] On August 25, 1978, four federal agencies (Department of Labor, Equal Employment Opportunity Commission, Office of Personnel Management and Department of Justice) issued the Adoption by Four Agencies of Uniform Guidelines on Employee Selection Procedures (1978). The Uniform Guidelines provide standards for fair selection procedures for EEO protected classes. Adverse impact in the selection process is presumed when the pass rate of applicants from a protected class with a low pass rate is less than 80 percent of the pass rate of applicants from the group with the highest selection rate. This is also referred to as the "four-fifths" rule.

[3] If one had a distribution of confounders, then one might be able to make probabilistic statements. Of the 24 cases cited by Taubes (1995), 80% have RR ≤ 3.

An important goal of science is to quantity the properties of entities. Since unmeasured confounders are difficult to deal with, one approach is to identify assumptions under which the properties of a confounder are completely determined by a single value. Given the complete specifications of a confounder one can then determine its' effects on a given association. Schield and Burnham (2003) have shown that specifying a binary confounder involves three values when using relative prevalences.

The first goal of this paper is to identify a simple way to determine all the properties of a confounder by specifying just a single parameter: the confounder size.

A second goal is to identify what size confounder is required to nullify an observed association. Nullification is confounder-induced spuriosity.[4] An association is *spurious* – of no effect – if it vanishes after taking a confounder into account.

A third goal is to generate intervals for an observed relative risk based on the influence of a binary confounder of a given size.

2. NOTATION

The notation used in this paper is the same as that used in Schield and Burnham (2003). The predictor variable A, the outcome variable E and the confounder variable B are all binary. The variable name is used to indicate the values (e.g., *A* and *non-A*). *A'* designates non-*A*. If *E* is cancer and *A* is smoker, then *P(E|A')* is the prevalence of cancer for non-smokers.[5] In order to study differences between, and ratios of, prevalences, this notation is used:

1. $DP(Y:X) \equiv P(Y|X)-P(Y|X')$,
2. $RP(Y:X) \equiv P(Y|X)/P(Y|X')$, $XRP(Y:X) \equiv RP(Y:X)-1$,
3. $AFP(Y:X) \equiv DP(Y:X) \cdot P(X)/P(Y)$.

The colon indicates that the following value and its complement are involved. Consider cancer (*E*), smoking (*A*) and a cancer gene (*B*). *DP(B:A)* is the differential prevalence of the cancer gene for smokers vs. non-smokers. *RP(E:A)* is the relative prevalence, *XRP(E:A)* is the excess relative prevalence of cancer for smokers vs. non-smokers. *AFP(E:A)* is the fraction of cancer cases in the population that are attributed to smoking.

The selection of *A* vs. *A'*, and of *B* vs. *B'* is arbitrary. This paper assumes they are selected so $DP(E:A) > 0$ and $DP(E:B) > 0$.[6] These selections do not determine whether *DP(B:A)* is positive or negative in general.

[4] A spurious association can also be chance-based: due to sampling variability when there is no association in the population.

[5] Note that P(X) signifies prevalence or percentage – not probability.

[6] If DP(E:A) = 0 then reversal is not meaningful. If DP(E:B) = 0 or DP(B:A) = 0, then spuriousity and reversal are impossible.

3. DEFINING CONDITIONS FOR CONFOUNDER-INDUCED SPURIOSITY

Schield and Burnham (2003) obtained defining conditions under which an observed association would be made spurious by a confounder when using a non-interactive OLS model for binary data. The OLS non-interactive model has the form:

4. $E(A,B) = b0 + b1 \cdot A + b2 \cdot B$.

Recall that *E* is the outcome of interest, *A* is the binary predictor and *B* is the binary confounder. Note that b1 is the partial regression coefficient between the outcome (*E*) and the binary predictor (*A*) after taking into account the influence of the confounder (*B*) using a non-interactive model.

If b1 = 0 then any association between *A* and *E* is spurious. There are many forms of this spuriosity condition as shown in Appendix A of Schield and Burnham (2003). The main problem is that at least three values must be specified for a confounder in order to determine its influence on an observed association.

Can we summarize these characteristics in the same way that we summarize a distribution by its center and spread? A first step is to see how these characteristics interact in rendering a given association spurious. Hopefully this will help us identify a single value that might be used to determine more than one property of a confounder. The goal is to identify summary characteristics that will identify confounders having the same nullifying strength on a relative prevalence in ways that are meaningful and useful.

4. INFLUENCE ON RELATIVE PREVALENCE

Schield and Burnham (2003) showed that the condition needed for a binary confounder to nullify an observed relative prevalence, RP(E:A), is given by:

5. $$P(B) \bullet XRP(E:B) = \frac{XRP(E:A)\{1+[P(A) \bullet XRP(B:A)]\}}{[XRP(B:A) - XRP(E:A)]}$$

For an observed excess relative prevalence *XRP*(*E*:*A*) and predictor prevalence *P*(*A*), this condition involves three other factors: *P*(*B*), *XRP*(*B*:*A*) and *XRP*(*E*:*B*).

Notice how *XRP*(*E*:*B*) is directly influenced by *P*(*B*) for given values of *XRP*(*E*:*A*), *P*(*A*) and *XRP*(*B*:*A*). If *P*(*B*) is small, then *XRP*(*E*:*B*) must be large and vice versa. If we have no knowledge of *P*(*B*), then it seems unwarranted and opportunistic to pick values that yield smaller values for either the confounder size, *RP*(*E*:*B*) or the confounder linkage with the predictor, *RP*(*B*:*A*).

To avoid opportunism and to simplify things, suppose that *P*(*B*) = *P*(*A*). This restricts confounders to those in the same prevalence class as the exposure, just as the Attributable Fraction of Cases in the Population (*AFP*) measures the correlation between exposure and cases – relative to the maximum possible for exposures in the same prevalence class: e.g., *P(B)* = *P(A)*. See Schield and Burnham (2002).

Since the confounder is hypothetical, there is no claim that this assumption or stipulation is realistic. Only that it is one way of achieving the stated goal of specifying all the properties of a confounder given the observed data and a single value.

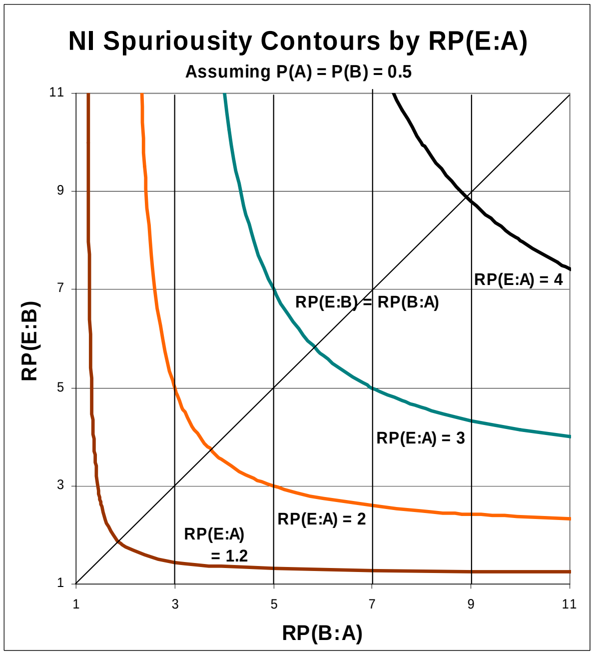
5. NULLIFICATION WHEN P(B)=P(A)

When P(B) = P(A), the nullification condition is:

6. $$XRP(E:B) = \frac{XRP(E:A)\{[1/P(A)] + XRP(B:A)\}}{XRP(B:A) - XRP(E:A)}$$

When *P*(*A*) = 0.5, we obtain the contours of equal strength shown in Figure 1. Although there are a wide range of combinations for *RP*(*E*:*B*) and *RP*(*B*:*A*) it can be shown that there is a symmetry around the line, *RP*(*E*:*B*) = *RP*(*B*:*A*).[7] When a function, y = f(x), has one point closest to the origin, that point is given by dy/dx = -x/y. Since these contours are symmetric about the diagonal, they are closest when their slope is -1, so that the closest point is x = y or *RP*(*E*:*B*) = *RP*(*B*:*A*).[8]

Figure 1: RP(E:B) vs RP(B:A) Spuriosity Contours



When P(B) = P(A), we can describe a strength contour using a single value, S, where RP(B:A) = PR(E:B) = S. For a given value of RP(E:A), this combination gives the point closest to the origin. This doesn't say that either *RP*(*E*:*B*) or *RP*(*B*:*A*) is smallest at this point. There are combinations where either is smaller, but this is the point at which the sum of their squares is smallest – they are jointly minimal.

All the other combinations can be derived given this one value of S since *P*(*B*) = *P*(*A*). One advantage of using this minimal Cartesian distance point, *XRP*(*B*:*A*) = *XRP*(*E*:*B*), is that it avoids extremes.

- A very weak confounder *XRP*(*E*:*B*), minimally more than XRP(E:A), can still nullify an association provided *XRP*(*B*:*A*) is very large. Focusing on the relatively small size of *XRP*(*E*:*B*) needed for nullification makes the observed association seem weak.
- A very strong confounder *XRP*(*E*:*B*) is required to nullify an association provided the excess confounder prevalence, *XRP*(*B*:*A*) is minimally greater than the observed association: *XRP*(*E*:*A*). Focusing on the large size of *XRP*(*E*:*B*) in this pair makes the observed association *XRP*(*E*:*A*) seem very strong.

6. S CONFOUNDER NULLIFICATION

An **S confounder** is hereby defined as a binary confounder where *P*(*B*) = *P*(*A*) and where *RP*(*B*:*A*) = *RP*(*E*:*B*) = S. Using equation 6, it follows that an S confounder will nullify the association *RP*(*E*:*A*) when

7. $$S - 1 = \frac{XRP(E:A)\{1 + [P(A) \bullet (S-1)]\}}{P(A)[(S-1) - XRP(E:A)]}$$

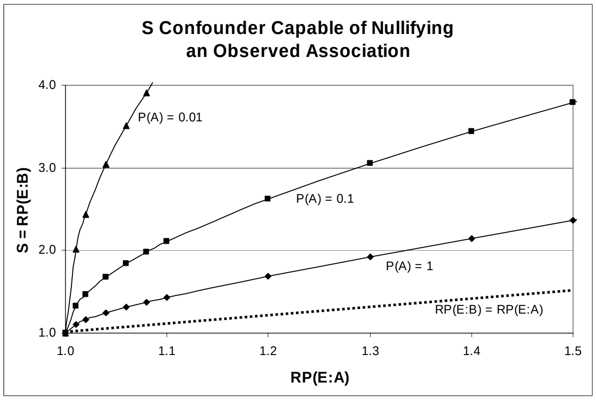
Collecting terms, solving the quadratic and taking the root for *S* > *RP*(*E*:*A*)[9] gives:

8. $$S = RP(E:A) + XRP(E:A)\sqrt{1 + 1/[P(A)XRP(E:A)]}$$

This equation identifies the size of an S confounder needed to nullify an observed association having a prevalence, *P*(*A*), and a relative prevalence, *RP*(*E*:*A*).

Figure 2 illustrates the size of an S confounder needed to nullify an observed association – given the prevalence *P*(*A*) and the relative prevalence, *RP*(*E*:*A*).

Figure 2: Minimum S Needed to Nullify Association



---

7 Let x = XRP(E:B), y = XRP(B:A), z = XRP(E:A) and k = 1/P(A). Equation 6 yields, x = z(k+y)/(y-z) so z = x•y/(k+x+y). The latter shows the symmetry between x and y for a given z.

8 This can be proven. Let $D^2 = x^2 + y^2 = y^2 + [z(k+y)/(y-z)]^2$. To minimize D for a given value of z, let dD/dy = 0. So, y = z + SQRT[z (z + k)] plus a negative and two imaginary roots. Substituting the positive solution into the equation for x gives: x = z + SQRT[z (z + k)]. So D is minimized when x = y, which means when XRP(B:A) = XRP(E:B) or RP(B:A) = RP(E:B)

9 RP(E:B) > RP(E:A) is a necessary condition for nullification. See Schield and Burnham (2003).

Consider those exposed to second hand smoke. If their prevalence is 25% and their relative risk of lung cancer is 1.2, then this association would be made spurious by an S confounder of size 2.1.

Note that the smaller the prevalence of the predictor, *P*(*A*), the larger the confounder size, S, needed to nullify an observed association, *RP*(*E*:*A*). For low-prevalence predictors, very large confounders are required to nullify the observed association.

Disciplines, not statisticians, must decide what size S confounder is considered small – just as with p-values.

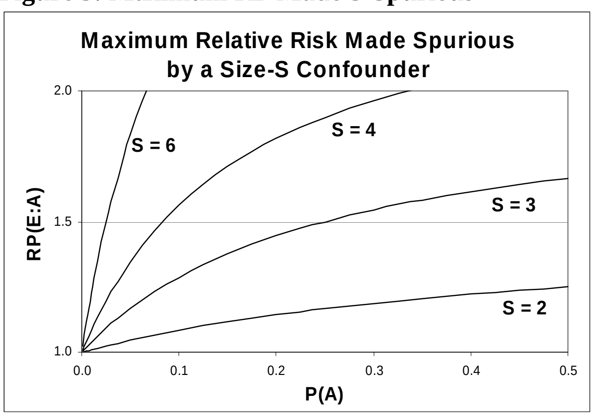
7. NULLIFICATION BY S CONFOUNDERS

The largest *XRP(E:A)* that is be made spurious by an S confounder as a function of *P(A)* is given by:[10]

9. $XRP(E:A) = P(A) \bullet (S-1)^2 / \{1 + [2P(A) \bullet (S-1)]\}$

As shown in Figure 3, relative prevalences under 1.5 are made spurious by S confounders with S < 4 when *P(A)* > 0.1.

Figure 3: Maximum RP Made S-Spurious



If S = 5, then *XRP*(*E:A*) = 16•*P(A)*/[1+8•P(*A*)].[11] If epidemiologists were to require that relative prevalences be able to withstand nullification by an S confounder of size 5, many relative risks would be noted as being vulnerable to confounding.[12]

8. IDEA OF CONFOUNDER INTERVALS

We now set aside the topic of nullification and turn to the question of influence. Suppose that an S confounder was tangled up in the observed association, *RP*(*E*:*A*). What value of *RP*(*E*:*A*) would be expected if that confounder were removed?

To repeat, note that we are not saying anything about nullification – just about influence – so we are not starting from the prior nullification equations. But certainly it seems useful to include the conditions *P*(*B*) = *P*(*A*) and *RP*(*B*:*A*) = *RP*(*E*:*B*) as determining the lower limit of a confounder interval. An S confounder may decrease without reversal, nullify or reverse an observed association. The latter is Simpson's Paradox. If the confounder interval for an observed relative risk included unity, we would say that for that size confounder the observed association was not 'confounder resistant'; otherwise it is 'confounder resistant.'

The next two sections (9 and 10) illustrate how the limits of a confounder interval are obtained using standardization for a particular case. Section 11 summarizes the general equations obtained when various single and double ratios are standardized.

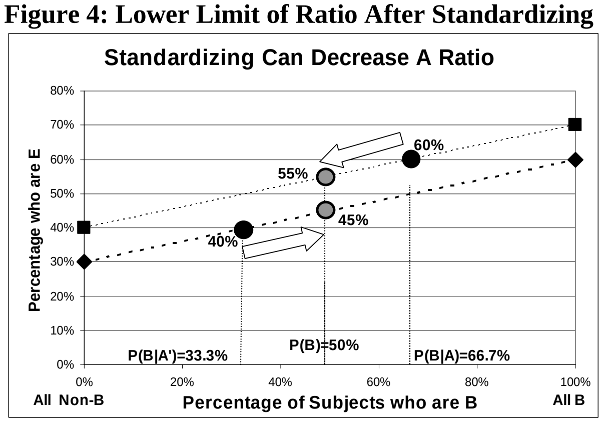
As presented in Schield (2004), standardization involves moving weighted averages along the lines connecting the actual data points: the rates. There is no necessity that these lines be parallel. But when viewing the results of a non-interactive model, there is typically no mention of the actual rates and the associated lines are necessarily parallel.

Standardizing using the four corner values from a non-interactive model (which are co-planar) give the same results as computing the expected values using the model. Using these four data points, the standardizing approach illustrates graphically what the non-interactive model does algebraically.

9. LOWER LIMIT

For example, what is the influence of an S confounder of size 2 on an observed association with *P*(*A*) = 50% and *RP*(*E*:*A*) = 1.5? To obtain the lower limit, the non-interactive model must fit four requirements: (1) *RP*(*E*:*A*) = 1.5, (2) *P*(*B*) = *P*(*A*) = 50%, (3) *RP*(*B*:*A*) = 2, and (4) *RP*(*E*:*B*) = 2. Figure 4 illustrates the non-interactive model fitting these values and the standardization from which one can obtain the lower limit.

Figure 4: Lower Limit of Ratio After Standardizing



To see (1) note that *P*(*E*:*A*) = 60% and *P*(*E*:*A*') = 40%, so *RP*(*E*:*A*) = 1.5. To see (2) note that *P*(*B*) = *P*(*A*) = 50% as specified. To see (3) note that *P*(*B*|*A*) = 66.7% and *P*(*B*|*A*') = 33.3%, so *RP*(*B*:*A*) = 2 (66.7% / 33.3%) as specified. Although this figure does not

10 Since RP is continuous, this also "equals" the minimum RP(E:A) that can withstand being made spurious by a size S confounder.

11 If S = 3 then XRP(E:A) = 4•P(A)/[1 + 4•P(A)]. If S = 10, then XRP(E:A) = 81•P(A)/[1+18•P(A)].

12 If P(A) = 0.5, RP(E:A) = 2.6; if P(A) = 0.1, RP(E:A) = 1.9.

show $P(E|B)$ or $P(E|B')$ directly, note that if $P(E|B) = 2/3$ and $P(E|B') = 1/3$, then $RP(E{:}B) = 2$ as specified. [13]

After taking into account the influence of this confounder, the standardized value of $P(E|A)$ is 55% while the standardized value of $P(E|A')$ is 45% so the standardized value of the relative prevalence of E for A, the lower limit of this confounder interval, is now 1.22.

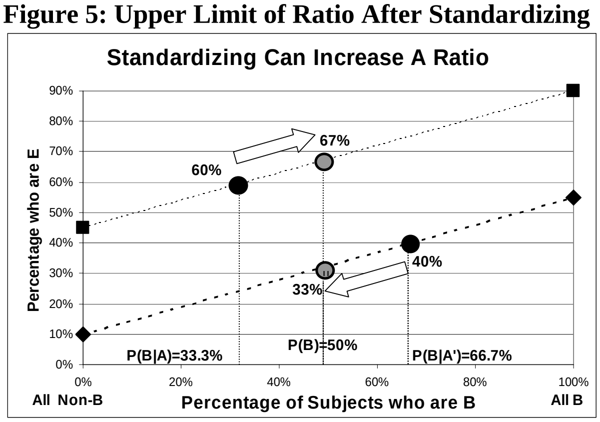
10. UPPER LIMIT

Now we turn to obtaining the upper limit of an S confounder interval. Recall that the removal of a confounder can increase the value of an association, $RP(E{:}A)$ as well as decrease the value. Since we know almost nothing about the confounder it seems inappropriate to assume that it must always decrease the observed association. Suppose that A and B are chosen so that $RP(E{:}A)$ and $RP(B{:}A)$ are both greater than unity. Schield and Burnham (2003) showed that in this case the direct effect is greater than the whole effect only if $0 < RP(B{:}A) < 1$.

What value of $RP(B{:}A)$ less than 1 can be readily determined given $RP(E{:}B)$? An obvious choice is $RP(B{:}A) = 1/RP(E{:}B)$. What is happening is that the confounder groups, B and B', are being exchanged – not in relation to the outcome E but in relation to the predictor groups, A and A'. In this sense, $RP(B{:}A) = 1/RP(E{:}B)$ used for the upper limit is closely related to $RP(B{:}A) = RP(E{:}B)$ used for the lower limit.

Using a non-interactive model, Figure 5 illustrates the standardization from which one can obtain the upper limit of a size S=2 confounder interval for $P(A) = 50\%$ and $RP(E{:}A) = 1.5$.

Figure 5: Upper Limit of Ratio After Standardizing



Note the four requirements this non-interactive model must fit: (1) $RP(E{:}A) = 1.5$, (2) $P(B) = P(A) = 50\%$, (3) $RP(B{:}A) = 1/2$, and (4) $RP(E{:}B) = 2$. To see (1) note that $P(E|A) = 60\%$ and $P(E|A') = 40\%$, so $RP(E{:}A) = 1.5$. To see (2) note that $P(B) = P(A) = 50\%$ as specified. To see (3) note that $P(B|A) = 33.3\%$ and $P(B|A') = 66.7\%$, so $RP(B{:}A) = 1/2$. Although $P(E|B)$ and $P(E|B')$ are not shown directly, note that if $P(E|B) = 2/3$ and $P(E|B') = 1/3$ then $RP(E{:}B) = 2$.[14]

After taking into account the influence of this confounder, the standardized value of $P(E|A)$ is 2/3 while the standardized value of $P(E|A')$ is 1/3 so the standardized value of the relative prevalence of E for A, the upper limit of this confounder interval, is 2.

So having obtained both the lower and upper limits of this S confounder interval, we can state the size of this particular confounder interval as follows. Given an observed relative prevalence of 1.5 and a predictor prevalence of 50%, the confounder interval due to a size 2 confounder is given by [1.22, 2.0].[15] [16]

To summarize, for the S confounder interval proposed herein, both lower and upper limits presume that $P(B) = P(A)$. The lower limit is that determined by $RP(B{:}A) = RP(E{:}B) = S$ while the upper limit is that determined by $RP(B{:}A) = 1/RP(E{:}B) = 1/S$.

11. CONFOUNDER INTERVAL FORMULAS

Appendix B derives the standardized values for $P(E|A)$ and $P(E|A')$ when $P(B|A) = P(B|A') = P(B)$ in terms of the slope b1 in the non-interactive model. Various combinations of these standardized values are also obtained. The spuriosity conditions obtained earlier can be obtained from these formulas.

It may be useful to see these standardized values in terms of the conditions specifying the predictor and the confounder – without including b1. The limits of S confounder intervals when $P(B) = P(A)$ are derived for $P(E)/P(E|A')$, in Appendix C and for the relative prevalence $RP(E{:}A)$ in Appendix D. In both cases, the formulas seem to conceal more than they reveal. Hopefully they contain analytical relationships that enable a better understanding of the underlying dynamics.

Figure 6 illustrates the lower limit of relative prevalence confounder intervals involving S confounders of size 2, 3 and 4. As a function of $RP(E{:}A)$, these lower limits are nearly linear.

---

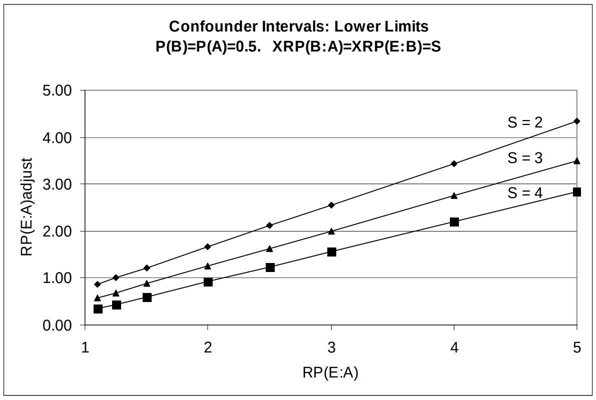
[13] P(E|B) is a weighted average on the right; P(E|B') on the left.
1. $P(E|B) = P(B|A) \cdot P(E|B,A) + P(B|A') \cdot P(E|B,A')$.
   $2/3 = (2/3)(70\%) + (1/3)(60\%) = +46.67\% + 20\%$
2. $P(E|B') = P(B'|A) \cdot P(E|B',A) + P(B'|A') \cdot P(E|B',A')$.
   $1/3 = (2/3)(30\%) + (1/3)(40\%) = 20\% + 13.3\%$

[14] P(E|B) is a weighted average on the right; P(E|B') on the left.
1. $P(E|B) = P(B|A) \cdot P(E|B,A) + P(B|A') \cdot P(E|B,A')$.
   $2/3 = (1/3)(90\%) + (2/3)(55\%) = 30\% + 36.67\%$
2. $P(E|B') = P(B'|A) \cdot P(E|B',A) + P(B'|A') \cdot P(E|B',A')$.
   $1/3 = (2/3)(45\%) + (1/3)(10\%) = 30\% + 3.33\%$

[15] The first confounder interval was obtained on 12/23/2003 using an Excel model with co-planar rates for RP(E:A) = 1.25, P(A) = 0.5.

[16] We avoid using ‘model’ in talking about an S confounder to emphasize that standardized values are based on a model – not the specifications of the S-confounder. We avoid using ‘adjusted’ to emphasize that the data itself is not being adjusted.

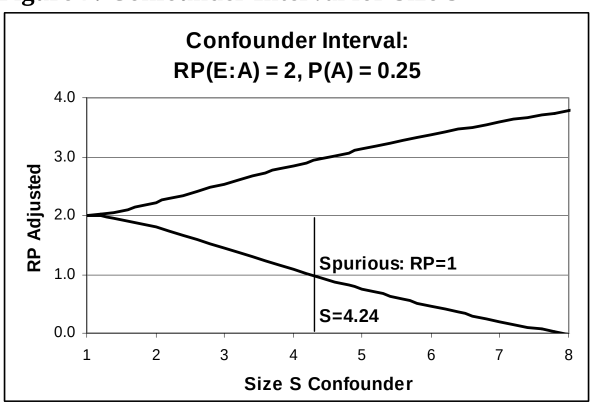
Figure 6: Confounder Lower Limits: P(B)=P(A)=0.5



Although the language of direct, whole and indirect effects is properly used only for differences, the terms can be appropriated for ratios provided the equation relating these items is set aside. Consider a whole effect given in terms of a ratio, *RP*(*E*:*A*), the observed relative prevalence association between *A* and *E*. Appendix B presents the direct effect in terms of a ratio given this whole effect and an S confounder.

Figure 7 shows the upper and lower limits of a confounder interval as a function of confounder size.

Figure 7: Confounder Interval for Size S



Some combinations of values may drive these equations into regions involving unacceptable values where the upper limit goes infinite or drops below 1, or where the lower limit goes below zero.

12. DISCUSSION

These relationships may be useful in educating data analysts and journalists on the possible influence of unobserved confounders even though these relationships just restate the size of the original association, RP(E:A), in different terms (since there is not yet any objective basis for selecting confounder sizes).

Furthermore, these relationships may be useful in setting rules or standards for publication by journal editors. They may even be useful in modeling to set a minimum criterion for including a predictor so as to avoid including predictors that may be statistically significant, but are too weak to withstand nullification by a confounder of a given size.

A deeper question involves the ability of relative prevalence to measure the causal status of the predictor. More work may be needed on this foundational issue.

13. RECOMMENDATIONS

The following are recommendations for handling count-based associations obtained from observational studies.

1. Those presenting relative risks or prevalences should indicate the minimum size S confounder that would nullify the observed association.
2. Those using relative risks or prevalences to make decisions for publication or for action should set minimum standards of the confounder size which an acceptable association must resist without being nullified or reversed – or what size confounder one should use in giving confounder intervals.
3. More analysis is needed on the use of a double ratio such as relative risk to measure the strength of evidence on the causal status of the predictor.

14. REFERENCES

Pearl, Judea (1998). *Why there is no statistical test for confounding, why many think there is, and why they are almost right.* UCLA Computer Science Department, Technical Report (R-256), July 1998.[17]

Schield, Milo and Thomas Burnham (2002). *Algebraic Relations Between Relative Risk, Phi and Measures of Necessity and Sufficiency in a 2x2 Table.* 2002 ASA Proceedings of the Section on Statistical Education.

Schield, Milo and Thomas Burnham (2003). *Confounder-induced Spuriosity and Reversal: Algebraic Conditions for Binary Data Using a Non-Interactive Model.* 2003 ASA Proceedings of the Section on Statistical Education. [CD-ROM], 3690-3697.

Schield, Milo (2004). *Three Graphs for Statistical Literacy.* International Conference on Mathematics Education (ICME-10) in Copenhagen.

Tabues, Gary (1995). *Epidemiology Faces Its Limits.* Science Magazine 269 (July 14): 164–69.[18]

Acknowledgments: This work was performed under a grant from the W. M. Keck Foundation "to support the development of statistical literacy as an interdisciplinary curriculum in the liberal arts." Dr. Schield is the project director of the W. M. Keck Statistical Literacy Project. He can be contacted at schield@augsburg.edu. This paper is at www.augsburg.edu/ppages/~schield and at www.augsburg.edu/statlit.

[17] On web at http://bayes.cs.ucla.edu/frl_papers.html.

[18] On web at www.agcom.purdue.edu/AgCom/homepages/tally/Science%20in%20Society%20web/96Taubesarticle.html

APPENDIX A: REGRESSION COEFFICIENTS

The following are four forms of the slope $b_1(E|A,B)$ in a non-interactive OLS regression model involving binary data. These four forms are taken from Appendix E in Schield and Burnham (2003).

The first form involves differences of double ratios.

A1. $$b_1 = \frac{DP(E:A) - [DP(B:A) \bullet DP(E:B)]}{1 - [DP(B:A) \bullet DP(A:B)]}.$$

The denominator is never negative.

The second form involves the attributable fraction in the population. This is closely related to Phi. See Schield and Burnham (2002).

A2 $$b_1 = \frac{P(E) \bullet [AFP(E:A) - AFP(B:A) \bullet AFP(E:B)]}{P(A) \bullet \{1 - [DP(B:A)(DP(A:B)]\}}.$$

The third and fourth forms involve double ratios.

A3. Double-ratio form with P(B|A') in numerator:

$$b_1 = \frac{P(E)\{XRP(E:A)[P(B \mid A') \bullet XRP(E:B)+1] - [P(B \mid A') \bullet XRP(B:A)(XRP(E:B)]\}}{\{1-[DP(B:A) \bullet DP(A:B)]\}[P(A) \bullet XRP(E:A)+1][P(B) \bullet XRP(E:B)+1]}$$

A4. Double-ratio form, $P(A)$ and $P(B)$ in numerator:

$$b_1 = \frac{P(E)\{XRP(E:A)[P(A) \bullet XRP(B:A) + P(B) \bullet XRP(E:B)+1] - [P(B) \bullet XRP(E:B) \bullet XRP(B:A)]\}}{\{1-[DP(B:A) \bullet DP(A:B)\}[P(A) \bullet XRP(B:A)+1][P(A) \bullet XRP(E:A)+1][P(B) \bullet XRPB(E:B)+1]}$$

Cases with zero denominators are ignored. Non-zero denominators are always positive when *XRP(B:A)*, *XRP(E:B) and XRP(E:A)* are positive.

To allow validation using Derive, the following appendices involve a proper-name notation described in Appendix B of Schield and Burnham (2003).

Note that in Appendix B, relationships are worked out two ways. First, by treating P(B) as a variable, B. There is no rule saying that standardization must be done using the common prevalence, P(B). Second by using the common prevalence, P(B). This gives both subgroups the same mixture as the combined group so P(B|A) = P(B|A') = P(B).

Thus B2d comes from B2c, B3d from B3c, B4e from B4d, B5e from B5d, B6b from B6a, B6e from B6d, B7b from B7a, B8b from B8a and B9b from B9a.

APPENDIX B: EXPECTED (b1)

Values expected when XP=XQ=B (using an 'eB' suffix) and when XP=XQ=XF (using an 'eXF' suffix).

B1. E(A, B) = AF + b1(A-AH) + b2(B-BH)[19]

B2a. AP = E(A=1, B=XP)
B2b. AP = AF + b1(1-AH) + b2(XP-BH)
B2c. APeB = AF + b1·AG + b2(B-BH)
B2d. APeXF = AF + b1·AG

B3a. AQ = E(A=0, B=XQ)
B3b. AQ = AF + b1(0-AH) + b2(XQ-BH)
B3c. AQeB = AF - b1·AH + b2(B-BH)
B3d. AQeXF = AF - b1·AH

B4a. AFeB = APeB·AH + AQeB·AG
B4b. AFeB = [AF + b1·AG + b2(B-BH)] AH
+ [AF - b1·AH + b2(B-BH)] AG
B4c. AFeB = [AF + b1(1-AH) + b2(B-BH)] AH
+ [AF - b1·AH + b2(B-BH)] (1-AH)
B4d. AFeB = AF + b2(B-BH)
B4e. AFeXF = AF

B5a. AP-AQ = [AF + b1(1-AH) + b2(XP-BH)]
- [AF + b1(0-AH) + b2(XQ-BH)]
B5b. AP-AQ = + b1 + b2(XP-XQ)
B5c. APeB - AQeb = [AF + b1·AG + b2(B-BH)]
- [AF - b1·AH + b2(B-BH)]
B5d. APeB - AQeB = b1(AG+AH) = b1
B5e. APeXF - AQeXF = b1

B6a. ARREeB = APeB / AQeB =
[AF+b1·AG+b2(B-BH)]/[AF-b1·AH + b2(B-BH)]

B6b. ARREeXF= APeXF / AQeXF
= [AF + b1·AG] / [AF - b1·AH]

B6c. ARREeB -1 = {[AF + b1·AG + b2(B-BH)]
/ [AF - b1·AH + b2(B-BH)]} – 1
B6d. ARREeB -1 = b1 / [AF - b1·AH + b2(B-BH)]
B6e. ARREeXF-1 = b1 / (AF - b1·AH)

B7a. APeB/AFeB=[AF+b1·AG+b2(B-BH)]
/[AF+ b2(B-BH)]
B7b. APeXF/AFeXF = [AF + b1(1-AH)] / AF

B8a. AFeB/AQeB= [AF+ b2(B-BH)]
/ [AF - b1·AH + b2(B-BH)]
B8b. AFeXF/AQeXF = AF/[AF - b1·AH)]

B9a. AAFPeB = (AFeB-AQeB)/AFeB
= b1·AH /[AF+ b2(B-BH)]
B9b. AAFPeXF = b1·AH/AF

If B has no effect on the A-E relationship (b2 = 0), then ARREeB = ARREeXF and the choice of B=BH makes no difference. But if B has an effect (b2 ≠ 0) on E, then the choice of B does make a difference.

[19] E(A,B) = b0 + b1·A + b2·B. AF = b0 + b1·AH + b2·BH

APPENDIX C: AF/AQeXF

Expanding Eq B6b using b1 from Eq. A4 gives:

C1a. T0 = 1 + T1(T2 + T3·T4 - T3 - T4)

C1b. T0 = 1/(AH·ARREeXF+AG)
C1c. T1 = $1/(1-XPhi^2)$
C1d. T2 = 1/(AH·ARRE + AG)
C1e. T3 = 1/(BH·BRRE + BG)
C1f. T4 = 1/(AH·XRPB + AG)
C1g. $XPhi^2 = \frac{AH(1-AH)BH(XRPB-1)^2}{(1-BH)(AH(XRPB-1)+1)^2}$

C2a. AF/AQ = AH·ARRE + AG[20]
C2b. BF/BQ = BH·ARRE + BG
C2c. XF/XQ = AH·XRPB + AG = BH/XQ
C2d. AF/AQeXF = AH·ARREeXF1 + AG

C3a. T0 = AQeXF/AF
C3b. T2 = AQ/AF
C3c. T3 = BQ/AF
C3d T4 = XQ/BH

Only T2 depends on AP. Define K1.
C4a. K1 = - [1 + T1(T3·T4 - T3 - T4)][21]
C4b. T0 = (T1·AQ/AF) - K1
C4c. 1/T0 = 1/[(T1·AQ/AF) - K1]
C4d. 1/T0 = (AF/AQ)/ [T1 – (K1·AF/AQ)]
C4e. AF/AQeXF = (AF/AQ)/[T1 - (K1·AF/AQ)]

Special Cases:

If K1 = 0, then
C5a. AF/AQeXF = (AF/AQ) / T1
C5b. AH·ARREeXF = [(AH·ARRE+AG)/T1] - AG
C5c. ARREeXF = [(ARRE+AG/AH)/T1] – AG/AH

If T1 = 0, then
C6a. (AH·ARREeXF + AG) = - 1/K1
C6b. AH·ARREeXF = (-1/K1) - AG
C6c. ARREeXF = - [(1/K1) + AG]/AH

If ARRE=1 (spurious independence where AF=AQ),
C7a. (AH·ARREeXF + AG) = 1 / (T1 - K1)
C7b. ARREeXF = {[1/(T1 - K1)] - AG}/AH

If ARRE = 1, AH=1/2 and XPhi=0 so T1=1, then
C8a. ARREeXF = (1+K1)/(1 - K1)
ARREeXF = 0 for K1=-1, 1 for K1=0 and ∞ for K1= 1.

If ARREeXF = ARRE so AF/AQeXF = AF/AQ, then B has no influence on the ratio association between A and E as evaluated at B = XF. Thus,
C9a T1 - (K1·AF/AQ) = 1 From C4e
C9b (T1-1)/K1 = AF/AQ

[20] AH(AP/AQ)+AG) = [AH·AP + AQ·AG]/AQ = AF/AQ
[21] K1 = -{1 + [BP/AF + XP/BH - 1]/[(1-XPhi2)(BP·XP)/(AF·BH)]}

APPENDIX D: ARREeXF

Let S be the size of the confounder where BH = AH.
D1a Let U = 1/(AH·S + AG)
D1b. Let V = 1/(AH/S + AG)
D1c. $XPhi^2 = [AH(XRPB-1)]^2/[AH(XRPB-1)+1]^2$

Lower Limit: Let BRRE = S, XRPB=S and BH = AH,
D2a. T3 = U = T4.
From Equation C4a:
D2b. K1= - [1 + T1(U^2-U-U)] = - [1 + (T1·U)(U-2)]
D2c. AF/AQeXF = (AF/AQ)
/ {T1 + [1+ (T1·U)(U-2)](AF/AQ)}
D2d. AH·ARREeXF+ AG = {(AH·ARRE+AG)
/ {T1 + [1+ (T1·U)(U-2)](AH·ARRE+AG)}}

Upper Limit: Let BRRE=S, XRPB=1/S and BH = AH,
D3a. T3 = U, and T4 = V
D3b. K1 = - [1 + T1(U·V – U – V)]
D3c. AF/AQeXF = (AF/AQ)
/ {T1 + [1+T1(U·V - U - V)](AF/AQ)}
D3d. AH·ARREeXF + AG = (AH·ARRE + AG)
/ {T1+[1+T1(U·V-U-V)](AH·ARRE + AG)}

Special Cases:
Let AH = 0.5:
D4a. $XPhi^2 = (XRPB-1)^2 / (XRPB+1)^2$
D4b. $1-XPhi^2 = 4 \cdot XRPB/[(XRPB+1)^2]$
D4c. $T1 = 1/(1-XPhi^2) = [(XRPB+1)^2]/ (4 \cdot XRPB)$
Lower Limit:
D5a. $T1Low = 1/(1-XPhi^2) = [(S+1)^2]/ 4S$
D5b. ARREeXF + 1 = {(ARRE+1)
/{T1Low + [1+T1Low·U(U-2)][(ARRE+1)/2]}}
Upper Limit:
D6a. $T1High = 1/(1-XPhi^2) = [(1/S+1)^2]/ (4/S)$
D6b. ARREeXF +1 = (ARRE+1)
/ {T1Hi + [1 + T1Hi(U·V-U-V)][(ARRE+1)/2]}

Let AH = 0.5 and S = 2 so U = 2/3 and V = 4/3.
Lower Limit:
D7a. $T1Low = [(S+1)^2]/ 4S = 9/8$
D7b. ARREeXF + 1 = (ARRE+1)
/ {(9/8) + [1+(9/8)(2/3)(-4/3)][(ARRE+1)/2]}
D7c. ARREeXF + 1 = (ARRE+1)
/ {(9/8) + (0)[(ARRE + 1)/2]}
So when ARRE = 2,
D7d. ARREeXF = {3/(9/8)} -1 = (8/3)-1 = 1.67

Upper Limit:
D8a. $T1High = [(1/S+1)^2]/ (4/S) = (9/4)/(4/2) = 9/8$
D8b. ARREeXF +1 = (ARRE+1)
/ {(9/8) + [1 + (9/8)(8/9-2/3-4/3)][(ARRE+1)/2]}
D8c. ARREeXF +1 = (ARRE+1)
/ {(9/8) - (1/4)[(ARRE+1)/2]}
So when ARRE = 2,
D8d. ARREeXF = {3/[9/8-3/8]} -1 = 4 - 1 = 3.0

For a size S=2 confounder, the confounder interval for ARRE = 2 with AH = 0.5 is [1.67, 3.0].